

# Quantum Information Theory 1

EPFL 2025-2026 学年第一学期课程笔记

X. D. H.

2025 年 12 月 29 日

## 摘要

这是 EPFL 2025 秋季学期的 QI 的课程笔记。

## 目录

<b>1</b>	<b>Lecture 2: Quantum 2 bit System</b>	<b>4</b>
1.1	Take home message . . . . .	4
1.2	Questions and thoughts . . . . .	6
<b>2</b>	<b>Lecture 3: Purification and POVM measurement</b>	<b>7</b>
2.1	Take home messages . . . . .	7
2.1.1	基础量子计算 . . . . .	7
2.1.2	Purification 相关 . . . . .	8
2.1.3	Fundamental POVM measurement . . . . .	9
2.2	Questions and thoughts . . . . .	11
<b>3</b>	<b>Lecture 4: Kraus Representation of Quantum Channel</b>	<b>13</b>
3.1	Kraus Representation of Quantum Channels . . . . .	13
3.2	POVM as a Quantum Channel . . . . .	14
3.3	Questions and thoughts . . . . .	16
<b>4</b>	<b>Lecture 5: Stinespring representation and Choi representation</b>	<b>17</b>
4.1	Stinespring Representation . . . . .	17
4.2	Choi Representation . . . . .	19
4.3	Questions and Thoughts . . . . .	20
4.4	Bell State 以及 2Qubit 的基 . . . . .	21
<b>5</b>	<b>Lecture 6: Choi representation and Complete Positivity</b>	<b>23</b>
5.1	Choi Representation 继续 . . . . .	23
5.2	Representation of Quantum Channel . . . . .	25
5.3	Questions and Thoughts . . . . .	26

<b>6</b>	<b>Lecture 7: Vectorization 的数学技巧</b>	<b>27</b>
6.1	Vectorization 和 Quantum Channel 结合 . . . . .	27
6.2	Average Over Unitary . . . . .	28
6.3	General map Vectorization . . . . .	30
6.4	Pauli Transfer Matrix . . . . .	31
6.5	Questions and Thoughts . . . . .	32
<b>7</b>	<b>Lecture 8: Shot Noise</b>	<b>33</b>
7.1	Shot Noise 讨论 . . . . .	33
7.1.1	Estimator 的基本概念 . . . . .	33
7.1.2	测量方法的影响 . . . . .	34
7.1.3	概率上界与收敛 . . . . .	35
7.2	Question and thoughts . . . . .	37
<b>8</b>	<b>Lecture 9: 区分量子态与矩阵模长</b>	<b>38</b>
8.1	区分量子态 . . . . .	38
8.2	矩阵模长 . . . . .	41
<b>9</b>	<b>Lecture 10: Entropy</b>	<b>44</b>
9.1	Classical Entropy . . . . .	44
9.1.1	Single Random Variable Entropy . . . . .	44
9.1.2	2 Random Variables Entropy . . . . .	46
9.2	Quantum Entropy . . . . .	48
9.2.1	Single Quantum System Entropy . . . . .	48
9.2.2	2 Quantum Systems Entropy . . . . .	49
9.3	Fidelity . . . . .	51
<b>10</b>	<b>Lecture 11: Resouce Theory and Entanglement Resource</b>	<b>53</b>
10.1	Set Up for Resource Theory . . . . .	53
10.2	Resource Theory of Entanglement . . . . .	53
<b>11</b>	<b>Lecture 12: Entanglement Measure</b>	<b>54</b>
11.1	Entanglement Measure . . . . .	54
11.2	Mixed State Entanglement Measure . . . . .	54
<b>12</b>	<b>Sup 1: QI 之中常用的数学技巧</b>	<b>55</b>
12.1	基础数学技巧 . . . . .	55
12.1.1	张量积 . . . . .	55
12.1.2	Unitary . . . . .	55
12.1.3	Trace . . . . .	56
12.2	Pauli 矩阵的特殊性质 . . . . .	56
12.2.1	基本定义 . . . . .	57

12.2.2 基本乘法性质 . . . . .	57
12.2.3 对易反对易关系 . . . . .	57
12.2.4 互相作用, pauli 矩阵作用在 pauli 基上 . . . . .	57
12.2.5 指数形式的 Pauli 矩阵 . . . . .	57
12.2.6 Pauli 矩阵的本征值与本征基 . . . . .	57
12.3 常用不等式 . . . . .	58
<b>13 Scratch book</b>	<b>59</b>

# 1 Lecture 2: Quantum 2 bit System

## 1.1 Take home message

### 2 bit 经典系统

对于 2bit 经典系统我们使用一个向量描述这个系统处于各个状态的概率：

$$v = (p, 1 - p) \quad (1)$$

并且向量要满足所有项的和为 1，并且每一项都是非负实数。所以其参数空间也很好找就是一个线段!! 线段上面的每一个点都对应一个经典的 2 bit system。

### 2bit 量子系统的空间

推广到量子系统我们下面使用一个  $2 \times 2$  Hermite 复矩阵进行描述，我们称为 density matrix，这个矩阵需要满足下面的性质：

- Hermite Matrix 也就是  $\rho = \rho^\dagger$
- Trace 1 也就是  $\text{Tr}(\rho) = 1$
- Positive Semi-definite 也就是  $\text{Eigen}\rho \geq 0$

### 参数空间：Density Matrix的Pauli基展开

对于所有的  $2 \times 2$  Hermite 矩阵我们都可以使用 Pauli 矩阵进行展开，对于 density matrix 我们发现三个实数自由度写在 pauli 基下面就是：

$$\rho = \frac{1}{2}(I + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z) = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma}) \quad (2)$$

可以确定我们只需要三个实参数就可以描述。

### Bloch Sphere以及量子态

对于这三个实数自由度的选取我们存在限制条件的根据【density matrix 的所有特征值都大于 0】我们发现：

$$\lambda_1 = \frac{1 + |\vec{r}|}{2} \geq 0, \quad \lambda_2 = \frac{1 - |\vec{r}|}{2} \geq 0 \quad \Rightarrow \quad |\vec{r}| \leq 1 \quad (3)$$

所以我们如果在  $\mathbb{R}^3$  之中画出来，所有可能的量子态就在一个半径为 1 的球之内，这个球我们称为 Bloch Sphere。

对于经典的 2 bit system 其实参数空间就是一个线段，但是量子的参数空间是一个球!! 所以量子系统的参数空间更大！并且其实经典的参数空间也就是 bloch 球的 y 轴这个直径。

### Pure State 和 Mixed State

我们定义 pure state 和 mix state 如下。

### Definition 1. Pure State and Mixed State

如果一个态不能写作  $\rho = p\sigma_a + (1-p)\sigma_b$  的形式，那么我们称这个态为 *pure state*。否则称为 *mixed state*。

如果使用 bloch 球的表示，pure state 就是球面上的点，mixed state 就是球面之内的点。

### Ensemble Ambiguity paradox

这并不是一个悖论而是量子信息上一个很神奇的点，也就是  $\rho = I/2$  这个 Maximally Mixed State。这个态可以有无数种 ensemble realization。比如：

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -| \quad (4)$$

这说明我们不能通过实验测量区分这个态是怎么进行制备的，不论是经典的制备还是量子的制备。

### Multi Qubit System

对于一般的  $\rho$  我们会发现满足了三个条件之后我们有  $N = d^2 - 1$  的自由度。需要这么多个实数参数才能描述一个  $d$  维的量子系统。

但是我们有一个点就是，更大的量子系统我们可以通过 tensor product 进行构造也就是：

$$\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i \quad (5)$$

### 纠缠态

我们进行构造的时候我们会发现有一些态并不能够写作上面的子系统张量积的形式「至少对于某两个特定子系统来说」我们认为这个子系统存在 entanglement 并且这个态是 entangled state。

### Definition 2. Entangled State

如果一个态不能写作  $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$  的形式，那么我们称这个态为 *entangled state*。

### Partial Trace of System

如果考虑一个多系统体系，但是我们仅仅关心一个子系统的行为我们可以 focus on 子系统，同时取一个 partial trace 这样子我们期望的情况是不变的。我们得到子系统 A 上面的 reduced density matrix：

$$\rho_A = \text{Tr}_B(\rho_{AB}) = \sum_i \langle Id \otimes i | \rho_{AB} | Id \otimes i \rangle \quad (6)$$

这个 reduced density matrix 可以让我们在子系统 A 上面进行所有的测量，并且得到整体系统一样的结果。

我个人理解这就是一个方便的工具。让我们等效的只看一个子系统!!! 在纠缠下的行为。

## 1.2 Questions and thoughts

**Question 1.0** 对于一个  $\rho$  表示的 system, 到底有多少参数是 basis invariant 的, 还有哪些参数其实是因为 basis transformation 引入的? 到底什么是 basis??

注意, 这些东西都是抽象的。其中的 coordinate 并不是我们日常的。很可能是很抽象的事件。但是同时, 我们也可以考虑进行 Similar Transformation。那么得到的就是另一个新的体系, 可以理解为过了一个量子门 (!!) □

## 2 Lecture 3: Purification and POVM measurement

### 2.1 Take home messages

#### 2.1.1 基础量子计算

##### Pauli 矩阵分解

Pauli 矩阵分解可以通过矩阵的 norm 来描述，我们用  $\text{Tr}$  来定义矩阵的内积：

$$\langle A|B \rangle = \text{Tr}(A^\dagger B) \quad (7)$$

于是 Pauli 基下面就可以写作：

$$\rho = \frac{1}{2}(\text{Tr}(\rho)I + \sum_{i=1}^3 \text{Tr}(\rho \sigma_i) \sigma_i) \quad (8)$$

##### Partial Tr 的一些 comment

Partial Trace 的操作其实是让对于 A 系统的求和概率 dept on B 系统的所有可能性！！这样可以让只看 A 系统的时候还能描述 AB 系统的纠缠行为！

##### 量子计算基础

对于量子计算来说，我们需要 manipulate qubit system。我们可以通过量子门来进行 manipulate。

**量子门是一堆 Unitary Operator** 我们可以通过一系列的量子门来进行量子计算。

##### 单比特门

下面介绍几个比较重要的 single qubit gate：

- **Pauli Gate:**  $X = \sigma_x, Y = \sigma_y, Z = \sigma_z$

- **Hadamard Gate:**  $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

Hadamard Gate 的作用是把  $|0\rangle$  变成  $|+\rangle$ ，把  $|1\rangle$  变成  $|-\rangle$ 。也就是把 Z 基变成 X 基。我们有：

$$H X H = Z, \quad H Y H = -Y, \quad H Z H = X \quad (9)$$

这样的关系让我们可以通过这个 Gate 改变测量方向。

##### 单比特一般演化

单比特态的演化我们可以把一个 Unitary Operator 写作这样的形式：

$$U = \exp(-i\theta \vec{s} \cdot \vec{\sigma}) \quad (10)$$

也就是我们把指数上面的 Hermite 矩阵进行展开！我们研究这个矩阵对于 Bloch 球向量的作用：

$$U \rho U^\dagger = \frac{1}{2}(I + \vec{r}' \cdot \vec{\sigma}), \quad (11)$$

其中  $\vec{r}' = R\vec{r}$ ， $R$  是一个绕着  $s$  轴旋转  $2\theta$  的旋转矩阵。

### 多比特门

多比特门最重要的是 CNOT 门，这个门的矩阵表示是：

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (12)$$

他的作用是第一个 bit 决定我们是不是反转第二个 bit。

**Theorem 1.** 所有 *Unitary Operator* 都可以通过 *CNOT* 和单比特门进行构造!!

### 2.1.2 Purification 相关

#### Purification 基本操作

对于一个 mix state 我们可以通过引入辅助系统来把这个 mix state 变成 pure state。这个过程我们称为 purification。

**Theorem 2.** 任何一个 *Generic Mix State*  $\rho_A$  都存在一个 state  $|\psi_{AB}\rangle$  使得：

$$\rho_A = \text{Tr}_B(|\psi_{AB}\rangle\langle\psi_{AB}|) \quad (13)$$

purification 并不是唯一，但是存在一个 general 的构造方法：

- 首先把  $\rho_A$  进行谱分解：

$$\rho_A = \sum_i p_i |i_A\rangle\langle i_A| \quad (14)$$

- 然后引入一个辅助系统 B，构造一个纯态：

$$|\psi_{AB}\rangle = \sum_i \sqrt{p_i} |i_A\rangle |i_B\rangle \quad (15)$$

#### Remark:

注意 purification 前后存在一个  $\sqrt{p_i}$  的区别，记得开根号!!

这个方法下我们的 B 系统的维度至少和 A 系统的 rank 一样大（也就是 A 系统有非 0 本征值的维度!）并且这是进行 purification 的最小维度!! 不会比这个小，否则就会丢失信息。

#### Purification After Evolve

我们考虑对于 B 系统进行一个 unitary 演化，我们会发现演化后的 purification 仍然是一个 purification!!

#### Proper and Improper Mixture

对于同一个 density matrix 表达的 state 我们很可能有不同的实现或者理解方式，比如对于一个 maximally mixed state  $\rho = \frac{1}{2}I$ ，我们可以有两种理解：



- **Proper Mixture**: 我们可以把这个态理解为  $|0\rangle$  和  $|1\rangle$  我们各自有 50% 的概率进行制备
- **Improper Mixture**: 我们也可以把这个态理解为一个 entangled 态  $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$  的一个子系统。

两者的物理完全不一样的!

- Proper Mixture 是经典的, 他是源于我们对于制备情况的无知
- Improper Mixture 是对于一个纠缠态的子系统的表述

### 2.1.3 Fundamental POVM measurement

#### 回顾怎么测量经典系统

对于一个经典系统, 所谓**测量**就是我们去问这个系统一个问题, 然后得到一个答案。

- 每一个问出的问题可以通过一组向量进行表达:  $\{\vec{m}_k\}$  每一个向量  $\vec{m}_k$  代表一种可能测量结果。这个向量需要满足两个条件:
  - $m_k$  的所有分量都是非负的【半正定矩阵】
  - $\sum_k m_k = 1$ , 也就是所有可能结果
- 测量得到  $k$  结果的概率是  $p_k = m_k \cdot p$ , 其中  $p$  是用来描述经典状态的向量

#### POVM Measurement 的定义

POVM Measurement 就是对于上面的情况的推广, 下面给出详细的定义要求:

#### Definition 3. POVM 测量

对于一个量子系统我们可以用一个密度矩阵进行描述  $\rho$ 。下面我们对其进行测量:

1. 对这个量子系统问出的一个问题可以数学上对应这一组矩阵  $M_k$  这组矩阵需要满足下面的条件:
  - 每一个  $M_k$  都是半正定矩阵, 也就是本征值  $\geq 0$
  - $\sum_k M_k = I$  【注意是矩阵直接求和出来  $I$ 】
2. 对于某一个系统  $\rho$ , 测量得到  $k$  结果的概率是:

$$p_k = \text{Tr}(\rho M_k) \quad (16)$$

#### Projective Measurement as a special POVM

我们定义 projective measurement 是一个特殊的 POVM 测量。需要满足其中矩阵  $M_k$  是正交归一化的投影矩阵:  $M_i M_j = \delta_{ij} M_i$  对于一个  $d$  维度的系统, 我们需要有  $d$  个这样的投影矩阵。

对于测量结果，我们之前讨论的 expectation value 可以这样得到。我们为每一个测量结果  $k$  赋予一个数值  $\lambda_k$  作为一个「具体测量出来的数」，下面给出 expectation value 的计算：

$$\sum_k \lambda_k p_k = \sum_k \lambda_k \text{Tr}(\rho M_k) = \text{Tr}(\rho \sum_k \lambda_k M_k) \quad (17)$$

我们定义  $\sum_k \lambda_k M_k$  也就是一个 Hermite 矩阵，我们把他叫做 observable。这就和量子力学里面讨论对应上了。

此外可以研究测量后系统的状态，根据量子力学我们会发现 projective measurement 后系统的状态变成：

$$\rho' = \frac{M_k \rho M_k}{\text{Tr}(\rho M_k)} \quad (18)$$

这是在测量到  $k$  结果之后系统的状态。【其中我们使用了一个归一化捏！保证我们的  $\rho'$  密度矩阵是  $\text{Tr} = 1$  的】

#### State Discrimination Problem

下面思考一个问题，我们现在手里有一个量子态，可能是  $|0\rangle$  也可能是  $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ 。使用【唯一一次测量】试图确认我们到底是哪一个态！！

**答案是：不可能！！我们永远不能仅仅一次测量就能保证确认是某一个态！！**

但是我们可以设计一个测量方案来最大化我们正确识别的概率！！这个问题就是 State Discrimination Problem。我们选择下面 POVM：

$$M_1 = a|1\rangle\langle 1|, \quad M_2 = b|-\rangle\langle -|, \quad M_3 = I - M_1 - M_2 \quad (19)$$

并且给出约束条件  $a, b \geq 0$  同时选择的  $a, b$  需要保证  $M_3$  是半正定的！！【这个约束条件很重要！！】在这个测量方案下：

- 得到结果 1 我们知道手里是  $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ ；
- 得到结果 2 我们知道手里是  $|0\rangle$ ；
- 得到结果 3 我们就知道了！！【这个结果 3 我们称为 inconclusive result】。

一个值得讨论的问题是：我们怎么选择  $a, b$  来最大化我们正确识别的概率呢？

首先我们需要计算正确识别的概率，在上面的构造之中我们知道给出正确结论的概率可以如下进行计算：

#### Definition 4. 正确概率

由于两个态都是随机  $1/2$  概率出现的，我们定义正确识别概率为：

$$\begin{aligned} P_{\text{correct}} &= \frac{1}{2}P(\text{correct}|\psi_1) + \frac{1}{2}P(\text{correct}|\psi_2) \\ &= \frac{1}{2}\langle \psi_1 | M_2 | \psi_1 \rangle + \frac{1}{2}\langle \psi_2 | M_1 | \psi_2 \rangle \end{aligned} \quad (20)$$

然后我们进行一个假设：【假设  $a = b$ 】 $M_1, M_2$  这两个 POVM 有一样的系数。于是我们计算  $M_3$  的本征值，根据半正定的条件我们可以得到最终的结果。

**Theorem 3.** 一般两个量子态的 *State Discrimination Problem* 的最优解

对于一般的两个量子态  $|\psi\rangle, |\phi\rangle$ ，我们可以通过下面的 POVM 测量来最大化正确识别概率：

$$\begin{aligned} F_\psi &= \frac{1}{1 + |\langle\phi|\psi\rangle|} |\phi^\perp\rangle\langle\phi^\perp| \\ F_\phi &= \frac{1}{1 + |\langle\phi|\psi\rangle|} |\psi^\perp\rangle\langle\psi^\perp| \\ F_\gamma &= I - F_\psi - F_\phi = \frac{2|\langle\phi|\psi\rangle|}{1 + |\langle\phi|\psi\rangle|} |\gamma\rangle\langle\gamma|, \end{aligned} \quad (21)$$

其中量子态是：

$$|\gamma\rangle = \frac{1}{\sqrt{2(1 + |\langle\phi|\psi\rangle|)}} (|\psi\rangle + e^{i \arg(\langle\phi|\psi\rangle)} |\phi\rangle). \quad (22)$$

比如对于上面区分  $|0\rangle, |+\rangle$  这样的两个态，我们可以选择：

$$M_1 = \frac{\sqrt{2}}{1 + \sqrt{2}} |1\rangle\langle 1|, \quad M_2 = \frac{\sqrt{2}}{1 + \sqrt{2}} |-\rangle\langle -|, \quad M_3 = I - M_1 - M_2. \quad (23)$$

这样的一个解！

#### Convex Mixture of Projective Measurement

我们发现如果把一些 projective measurement 进行 convex mixture，我们会得到一个 POVM 测量！！那么这个 POVM 的意义是什么呢？

其实是说我们有一定概率进行某一个 projective measurement，同时也有一定概率进行另一个 projective measurement！

#### Extra Note on informationally complete measurement

[待补充，以及参考第三次的作业题！]

## 2.2 Questions and thoughts

### Question 2.1 怎么理解 projective measurement 是一个特殊的 POVM?

我目前的理解是 Projective Measurement 仅仅允许测量某一个密度矩阵表示的态，【到一个纯态上面的概率】。因为我们的 projective operator 其实就是一个纯态的投影。

但是一般的 POVM 是允许我们讨论一个混态是某个混态的概率的。这个样子，我们量子体系可以问的问题就更多了！！能测量的问题就更多了！ □

### Important: Pauli 矩阵的海量结论

由于 Pauli 矩阵在 QI 之中应用太频繁了，我下面列出海量重要的结论，方便查阅!!

$$\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (24)$$

1.

### Question 2.2 State Discrimination Problem 的最优解是怎么样的?

这个 wiki 上面有明确讨论: [wiki](#) 最优解，建议参考这个页面。同时结论我也写在正文内容里面了。 □

### 3 Lecture 4: Kraus Representation of Quantum Channel

#### 3.1 Kraus Representation of Quantum Channels

##### Quantum Channels

对于一个 Quantum Channel 来说需要满足下面四个性质：

1. Trace Preserving  $Tr(\mathcal{E}(\rho)) = Tr(\rho)$
2. Preserve positivity  $\mathcal{E}(\rho) \geq 0$  也就是本征值是都需要非负的
3. Linear  $\mathcal{E}(a\rho_1 + b\rho_2) = a\mathcal{E}(\rho_1) + b\mathcal{E}(\rho_2)$
4. Completely Positive

$$(\mathcal{I} \otimes \mathcal{E})(\rho) \geq 0 \quad (25)$$

也就是对于任意的辅助系统 A,  $\mathcal{I}_A \otimes \mathcal{E}_S$  作用在系统 S 上, 结果依然是一个正定矩阵。这是一个特别重要的性质!!

##### Kraus Representation Theorem

对于这样的 Quantum Channel, 我们可以使用 Kraus Operators 进行描述, 存在下面的定理:

##### Theorem 4. Kraus Representation

满足上面四个条件的 Quantum Channel 可以通过一组 Kraus Operators  $\{A_i\}$  进行实现, 满足下面的两个要求:

- $\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger$
- $\sum_i A_i^\dagger A_i = I$

我们可以很容易证明这个演化的构造是满足前面三个条件的!!

##### Quantum Channel基础例子

下面给出几个比较基础的 Quantum Channel 例子:

##### • Unitary Operator

显然一个 Unitary Operator 也就是一个 Quantum Channel。像是 Pauli Gate, Hadamard Gate... 等等

##### • Dephasing Channel

这个是我们如果把一组基下面的所有 projector 放在一起作为一组 Kraus Operator 那么可以得到一个 Dephasing Channel。比如说:

$$A_0 = |0\rangle\langle 0|, \quad A_1 = |1\rangle\langle 1|. \quad (26)$$

对于 2bit 上面这样的情况, 我们相当于只考虑 bloch 球的 z 方向投影。

### Remark:

Dephasing Channel 的一个作用我们可以理解为「把一个量子的态变成经典的」。因为经典的态其实就等价于在  $z$  方向的一个概率分布。量子和经典的区别就在于两字的态可以有  $x$  和  $y$  方向的分量。

这样的一个过程我们称之为 decoherence。也就是退相干。

- **Convex combination of unitary channels**

也就是按照一定概率进行一组 unitary channel 的选择:  $\mathcal{E}(\rho) = \sum p_i U_i \rho U_i^\dagger$ 。存在一定的概率进行每一个 Unitary Channel 的作用。

- **Complete Depolarizing Channel**

这个 channel 的作用是消灭一个态并直接变成另外的一个态:  $\mathcal{E}(\rho) = \sigma$  我们可以按照如下方法构造 Kraus Operator:

$$A_{ij} = \sqrt{\lambda_j} |j\rangle \langle i| \quad (27)$$

其中  $\sigma = \sum_j \lambda_j |j\rangle \langle j|$  是  $\sigma$  的谱分解。我们可以验证这个 channel 的 Kraus Operator 满足  $\sum_{ij} A_{ij}^\dagger A_{ij} = I$ , 并且  $\mathcal{E}(\rho) = \sigma$ 。

- **Partial Trace**

这个显然也是一个 channel, 这个 channel 的  $A_k$  Operator 比较古怪因为它是一个从大空间到小空间的 operator。我们写作这个样子  $A_k = I_A \otimes \langle k|_B$

### Measure record and Update

下面我们讨论一个特殊的 channel, 存在一个特殊的测量记录和更新的功能:

$$\mathcal{E}(\rho) = \sum_k |k\rangle \langle k| \otimes B_k \rho B_k^\dagger \quad (28)$$

其中  $B_k$  本身就是一组第二个子系统上的 Kraus Operator。

对于这个 channel 我们把测量的结果「每一种结果的概率」记录在第一个子系统, 并让第二个系统「变成测量后坍缩的状态」。

$$Tr_s(\mathcal{E}(\rho)) = \sum_k |k\rangle \langle k| Tr(B_k \rho B_k^\dagger) = \sum_k p_k |k\rangle \langle k| \quad (29)$$

$$Tr_a(\mathcal{E}(\rho)) = \sum_k B_k \rho B_k^\dagger \quad (30)$$

## 3.2 POVM as a Quantum Channel

### POVM测量以及坍缩

并且我们会发现, 这个 channel 其实给出了一组 POVM 测量:

$$M_k = B_k^\dagger B_k \quad (31)$$

这样子定义的  $M_k$  矩阵是一个良定的 POVM 测量！那么我们可以理解  $\mathbf{a}$  系统其实就是一个 POVM 测量之后的坍缩结果。

**Theorem 5.** *POVM 测量后坍缩结果*

进行一个 POVM 测量之后，测量结果坍缩到某一个态是：

$$\rho \rightarrow \rho_k = \frac{B_k \rho B_k^\dagger}{\text{Tr}(B_k \rho B_k^\dagger)} \quad (32)$$

其中  $B_k = \sqrt{M_k}$ 。「注意，这是一个记号，我们写出来意思其实是：  $M_k = B_k^\dagger B_k$ 」分母是一个归一化条件。因为所有的 *density matrix* 是需要保证 *trace* 为 1 的。

**Naimark Dilation Theorem**

我们之前讨论 POVM 之于 projective measurement 很像是 mix state 和 pure state 的关系。确实，我们也可以通过选择一个辅助系统对于一个 POVM 进行 purification 变成一个 projective measurement。下面定理给出了操作：

**Theorem 6.** *Naimark Dilation Theorem*

如果  $\{M_i\}$  是一个良定的 POVM 测量作用在一个系统  $A$  上面。那么我们存在一个 *isometry* 「一个可逆的向更大的系统的映射」保证：

$$M_i = V^\dagger (\Pi_i) V \quad (33)$$

其中  $\Pi_i$  是一组 *projective measurement* 作用在更大的系统  $A \otimes B$  上面，并且  $V : \mathcal{H}_A \rightarrow \mathcal{H}_A \otimes \mathcal{H}_B$  是一个 *isometry*。

我们可以严格构造这个 *isometry* 是：

$$V = \sum_k |k\rangle \otimes B_k \quad (34)$$

并且这个 projective Operator 需要满足：

$$\Pi_k = |k\rangle \langle k| \otimes I \quad (35)$$

在这样的构造下面，我们可以证明，这个 projective measurement 等价于 POVM 测量。我们的操作是，先把所有的系统  $A$  的态映射到整个更大的系统上面：

$$\rho_A \rightarrow V \rho_A V^\dagger = \sum_{k,k'} |k\rangle \langle k'| \otimes B_k \rho_A B_{k'}^\dagger \quad (36)$$

然后我们对其进行 projective measurement，可以证明最后的结果就是 POVM 测量的结果。

$$\text{Tr}(\Pi_i V \rho_A V^\dagger) = \text{Tr}(M_i \rho_A) \quad (37)$$

### Remark:

注意，如果我们希望在一个更大的系统上用 Projective Measurement 实现这个 POVM 的话我们需要引入一个辅助系统 B，并且使用两个操作：

1. 先把系统 A 的态通过 isometry 映射到更大的系统上面
2. 然后在更大的系统上进行 Projective Measurement

其中最关键的不是构造 Projective Measurement，而是构造这个 isometry!!

### 一般POVM的Channel

对于一般的 POVM 测量，我们可以构造一个 Channel 来实现 POVM 之后 Update 的效果。给定一个合理的测量的 POVM  $\{M_k\}$ ，我们可以定义一个 Channel：

$$M_i = \sum_{\alpha} K_{i\alpha}^{\dagger} K_{i\alpha} \quad (38)$$

其中  $K_{i\alpha}$  是一组 Kraus Operator。如果测量到了结果 i，那么系统坍缩到的态是：

$$\rho \rightarrow \rho_i = \frac{\sum_{\alpha} K_{i\alpha} \rho K_{i\alpha}^{\dagger}}{\text{Tr}(\sum_{\alpha} K_{i\alpha} \rho K_{i\alpha}^{\dagger})} \quad (39)$$

### 3.3 Questions and thoughts

没有什么问题耶！好耶!!



## 4 Lecture 5: Stinespring representation and Choi representation

{sec:Lect

### 4.1 Steinspring Representation

#### 环境相互作用下的Unitary演化

我们考虑一个量子体系  $S$  和一个环境  $E$  的相互作用。假设环境  $E$  的初始态是  $|0\rangle_E$ ，那么我们可以把体系  $S$  和环境  $E$  看成一个联合体系。我们把这个体系记作：

$$\rho_{SE} = \rho_S \otimes |0\rangle_E \langle 0|. \quad (40)$$

下面我们把这个联合体系进行一个整体的 unitary 演化  $U$ ，我们使用这样的演化算符：

$$U = |0\rangle \langle 0| \otimes U_0 + |1\rangle \langle 1| \otimes U_1. \quad (41)$$

下面我们计算这个整体的演化之后只看  $S$  体系的状态，我们得到：

$$\text{Tr}_E \left( U(\rho \otimes |0\rangle \langle 0|) U^\dagger \right) = \rho_{00}|0\rangle \langle 0| + \rho_{11}|1\rangle \langle 1| + \rho_{01}|0\rangle \langle 1|C + \rho_{10}|1\rangle \langle 0|C^*, \quad (42)$$

其中我们引入的  $C$  常数是： $C = \langle 0|U_1^\dagger U_0|0\rangle$ 。这个是由我们的 channel 以及环境的初始态共同决定的。

#### Partial Dephasing Channel结果

观察这个结果，我们会发现这个更大系统的 Unitary 配上合适的环境的出事状态，我们可以得到一个 partial dephasing 的效果下面我们证明一下，上面的结果  $\text{Tr}_E (U(\rho \otimes |0\rangle \langle 0|) U^\dagger) = \sigma = p\rho + (1-p)Z\rho Z$  成立，我们直接进行计算：

$$p\rho + (1-p)Z\rho Z = \rho_{00}|0\rangle \langle 0| + \rho_{11}|1\rangle \langle 1| + (2p-1)(\rho_{01}|0\rangle \langle 1| + \rho_{10}|1\rangle \langle 0|), \quad (43)$$

我们对比系数会发现如果选取合适的 Unitary  $U_1, U_2$  满足关系：

$$2p-1 = C = \langle 0|U_1^\dagger U_0|0\rangle, \quad (44)$$

那么之前更大系统的 Unitary 就相当于一个 partial dephasing channel。这给我们一个很重要的启示：任何一个 Quantum Channel 都可以通过引入一个环境系统，然后对联合系统进行 Unitary 演化，最后对环境系统进行 Partial Trace 来实现！！

#### Steinspring Dilation Thm

于是在上面的启示下面我们给出下面的定理：

#### Theorem 7. Steinspring Dilation Theorem

对于任意的一个 quantum channel  $\mathcal{E}$ ，都存在一个辅助系统  $E$  的 Hilbert 空间  $\mathcal{H}_E$ ，一个环境的初始态  $|0\rangle_E \in \mathcal{H}_E$ ，以及一个联合系统的 unitary 演化  $U$  使得对于任意的输入态  $\rho_S$ ，都有：

$$\mathcal{E}(\rho_S) = \text{Tr}_E \left( U(\rho_S \otimes |0\rangle \langle 0|_E) U^\dagger \right). \quad (45)$$

下面我们将进行证明，我们使用 Kraus 表示进行证明，假设有一个 quantum channel 使用 Kraus 表示写出来是：

$$\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger, \quad \sum_i A_i^\dagger A_i = I. \quad (46)$$

下面我们希望能够使用 Steinspring 表示来表示这个 channel。我们试图构造一个 Unitary 满足下面的条件：

$$U(|0\rangle \otimes |\psi\rangle) = \sum_i |i\rangle \otimes A_i |\psi\rangle. \quad (47)$$

这样的话如果我们 tr 掉 E，就会得到： $\sum_i A_i |\psi\rangle$  我们用张量积的形式写开来，我们知道始末两个量子态可以写作

$$|0\rangle \otimes |\psi\rangle = \begin{pmatrix} |\psi\rangle \\ 0 \\ \vdots \end{pmatrix}, \quad \sum_i |i\rangle \otimes A_i |\psi\rangle = \begin{pmatrix} A_0 |\psi\rangle \\ A_1 |\psi\rangle \\ \vdots \end{pmatrix}. \quad (48)$$

所以我们最终呈现的 Unitary 矩阵一定是，第一列是  $A_i$  后面的列待定。我们可以根据 Unitary 的条件确定后面列的数值。

当我们写出来  $U(|0\rangle \otimes |\psi\rangle) = \sum_i |i\rangle \otimes A_i |\psi\rangle$  的时候，我们考虑作用在一般的一个密度矩阵上面：

$$\text{Tr}_E \left( U(|0\rangle\langle 0| \otimes |\psi\rangle\langle \psi|) U^\dagger \right) = \sum_{i,j} \text{Tr}_E \left( |i\rangle\langle j| \otimes A_i |\psi\rangle\langle \psi| A_j^\dagger \right) \quad (49)$$

我们 tr 掉 E 之后就会发现，non-diagonal 的部分自然就消失了，剩下的就是我们需要的 Kraus Operator 作用的 Channel 的形式：

$$\sum_i A_i |\psi\rangle\langle \psi| A_i^\dagger \quad (50)$$

对于纯态的 Channel 我们证明完毕，对于混态其实就是一个直接的推广，上面的操作和线性组合是兼容的，我们只需要推广：

$$\rho = \sum_i p_i |\psi_i\rangle\langle \psi_i| \quad (51)$$

然后对于纯态使用上面的操作就可以了！！

#### 补充内容：如何寻找Steinspring Operator?

我们给定一组 Kraus Opertaor，我们在想如何具体的构造 Steinspring Operator  $U$ 。上面给出了理论的方法，下面我们讨论实操的每一个步骤。对于具体的计算，我们比较方便的有两种操作：

**方法 1: Gram-Schmidt Process**

**方法 2: Quantum Computation Method**

[我懒了写了，具体看第五次作业的 1.(c)。这里给出了极其具体的操作方法]

## 4.2 Choi Representation

### Vectorization数学技巧

为了介绍这个 representation, 我们需要先引入 vectorization 的技巧, 我们定义:

#### Definition 5. Vectorization

也就是把一个矩阵  $A$  变成一个向量  $|Vec(A)\rangle$ , 具体的定义是:

$$A = \sum_{ij} A_{ij} |i\rangle\langle j| \rightarrow |Vec(A)\rangle = \sum_{ij} A_{ij} |i\rangle \otimes |j\rangle. \quad (52)$$

在这样的定义下面我们可以证明一个很重要的恒等式:

**Lemma 1.** 对于任意的矩阵  $A, B, C$ , 都有:

$$|Vec(ABC)\rangle = (A \otimes C^T) |Vec(B)\rangle. \quad (53)$$

这个等式有很多的应用:

1. Ricochet Identity:

$$(A \otimes I) |Vec(I)\rangle = (I \otimes A^T) |Vec(I)\rangle. \quad (54)$$

值得注意的是  $|Vec(I)\rangle$  其实就是 bell state  $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$

证明方法就是写作  $|Vec(ABC)\rangle$  的形式, 但是由于有的矩阵就是 Identity 所以可以随便换顺序, 因此给出了上面的结果。

2. Usage of Ricochet Identity:

比如我们可以推导出来这样的一个结果:

$$(UV^\dagger \otimes I) |\varphi^+\rangle = (U \otimes V^*) |\varphi^+\rangle. \quad (55)$$

这个给出了两个量子线路是等价的, 也就是说我们可以简化量子线路。

### Choi Representation定义

上面 vectorization 的数学技巧的基础上我们发现我们还可以很好的使用量子态来表示一个 operator。我们这个思路下, 我们或许可以使用量子态来表示一个 quantum channel。我们给出下面的定义:

### Definition 6. Choi Representation

对于任意的一个 Quantum Channel  $\mathcal{E}(\rho)$ , 我们定义一个与之对应的量子态 **Choi State**  $J(\mathcal{E})$ , 这个量子态定义为:

$$J(\mathcal{E}) = (\mathcal{E} \otimes I)(|Vec(I)\rangle\langle Vec(I)|), \quad (56)$$

### Kraus Operator表示的Channel的Choi State

显然对于 Kraus Channel 我们可以计算出来 Choi State 的形式, 来考虑下面这样的 Kraus Operators:

$$(UV^\dagger \otimes I)|\varphi^+\rangle = (U \otimes V^*)|\varphi^+\rangle. \quad (57)$$

通过计算我们知道的结论是:

### Theorem 8. Choi State 和 Kraus Operators 的关系

$$J(\mathcal{E}) = \sum_i |Vec(A_i)\rangle\langle Vec(A_i)|. \quad (58)$$

### Remark:

所以 Choi State 某种意义上可以理解为 Kraus Operators 的 Vectorization 之后的结果。

一个例子, 我们可以计算出 dephasing channel 的 Choi State:

$$J(\mathcal{E}) = (\mathcal{E} \otimes I)(|00\rangle\langle 00| + |11\rangle\langle 11| + |00\rangle\langle 11| + |11\rangle\langle 00|) \quad (59)$$

$$= |00\rangle\langle 00| + |11\rangle\langle 11| \quad (60)$$

$$= |Vec(A_0)\rangle\langle Vec(A_0)| + |Vec(A_1)\rangle\langle Vec(A_1)|. \quad (61)$$

## 4.3 Questions and Thoughts

**Question 4.3** 为什么 Steinspring Operator 我们证明作用在一个量子态上面给出需要的结果就可以了? 不需要在密度矩阵吗? 感觉不够 general?

其实是足够的, 我在正文内容补充了解释看上面吧!

□

**Question 4.4** 一个 general 的 choi state 我们怎么找到一组基进行对角化?

我们如果想把一个 Choi State 给出一组 Kraus Operator 的话那么需要对于其进行对角化, 下面我们讨论对于一个 general 的 Choi State 我们怎么计算对角化!

在 computational Basis 下面一个 General 的 Choi State 可以写作下面的形式:

$$\rho = \sum_{i,j,k,l \in \{0,1\}} \rho_{ij,kl} |ij\rangle\langle kl| \quad (62)$$

我们首先需要对于 Computational Basis 的向量形式有一个 convention (因为对角化都是用矩阵进行 mma 操作的)。我们写：

$$|00\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad |01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad |10\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad |11\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (63)$$

在这样的设定基础上我们可以把  $\rho$  写成一个 4x4 的矩阵：

$$\rho = \begin{bmatrix} \rho_{00,00} & \rho_{00,01} & \rho_{00,10} & \rho_{00,11} \\ \rho_{01,00} & \rho_{01,01} & \rho_{01,10} & \rho_{01,11} \\ \rho_{10,00} & \rho_{10,01} & \rho_{10,10} & \rho_{10,11} \\ \rho_{11,00} & \rho_{11,01} & \rho_{11,10} & \rho_{11,11} \end{bmatrix} \quad (64)$$

然后我们就可以直接对于这个矩阵进行对角化，得到一组本征值和本征向量。然后我们把本征向量重新写成  $|Vec(A_i)\rangle$  的形式，然后使用 vectorization 的反操作就可以得到 Kraus Operator 了!!  $\square$

**Question 4.5** 一个 general 的密度矩阵我们怎么使用 bell basis 和 computational basis 进行展开？

我们知道 2qubit 这两个 basis 构成了 2qubit 系统的完备基，那么我们可以使用这两个 basis 进行展开。我们具体计算的方法就是逐个分量进行计算：

$$\rho_{ij,kl} = \langle ij|\rho|kl\rangle \quad (65)$$

$\square$

#### 4.4 Bell State 以及 2Qubit 的基

我们常用的一组 state 就是 bell state，我们定义：

$$\begin{aligned} |\Phi^+\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \\ |\Phi^-\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \\ |\Psi^+\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \\ |\Psi^-\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \end{aligned} \quad (66)$$

这组 state 有很多特殊的性质。

- 四个 Bell State 互相正交，并且构成了 2-bit 系统的一个完备基础。
- Bell State 是最大纠缠态，代表着两个 bit 能够有的最大的纠缠

- Bell State 其实都是正比于 Pauli Matrix 的 Vectorization 形式:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}|Vec(I)\rangle, \quad |\Phi^-\rangle = \frac{1}{\sqrt{2}}|Vec(Z)\rangle, \quad (67)$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}|Vec(X)\rangle, \quad |\Psi^-\rangle = \frac{i}{\sqrt{2}}|Vec(Y)\rangle. \quad (68)$$

对于 2Qubit 的系统我们还有一组基也就是 computational basis, 写作:

$$\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}. \quad (69)$$

这组基和 Bell State 基是等价的, 我们可以通过下面的变换进行互相转换。

## 5 Lecture 6: Choi representation and Complete Positivity

### 5.1 Choi Representation 继续

我们之前研究知道通过 Vectorization 我们可以将一组 Kraus Operators 映射为一个 Choi State。下面我们希望反过来通过 Choi State 来得到 Channel 的 Kraus Operators。

例子：一个特殊channel的choi state

首先我们计算一个特别的 channel 的 Choi State。我们考虑下面的 channel：

$$\mathcal{E}(\rho) = \frac{1}{3} (\text{Tr}(\rho)I + \rho^T). \quad (70)$$

我们计算这个 channel 的 Choi State，也就是对于  $|Vec(I)\rangle$  进行  $\mathcal{E}(\rho)$  的作用我们计算的时候会把前一个和后一个 qubit 进行区分开来计算：

$$J(\mathcal{E}) = \sum_{i,j} \mathcal{E}(|i\rangle\langle j|) \otimes |i\rangle\langle j| \quad (71)$$

于是我们给出结果是：

$$J(\mathcal{E}) = \frac{1}{3} \left( \sum_i I \otimes |i\rangle\langle i| + \sum_{i,j} |j\rangle\langle i| \otimes |i\rangle\langle j| \right) \quad (72)$$

$$= \frac{1}{3} \left( \sum_i I \otimes |i\rangle\langle i| + \sum_{i,j} |j\rangle\langle i| \otimes |i\rangle\langle j| \right) \quad (73)$$

我们会发现这个 Choi State 的第二个部分是一个 SWAP Operator。我们下面定义：

- SWAP Operator:

$$S = \sum_{i,j} |i\rangle\langle j| \otimes |j\rangle\langle i| \quad (74)$$

- SWAP Operator 的本征态，就是所有的 Bell State 并且除了最后一个本征值都是 1，我们可以 decompose 成为下面的形式：

$$SWAP = |\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+| - |\psi^-\rangle\langle\psi^-| \quad (75)$$

- SWAP Operator 相当于对于一个 2qubit 系统进行「转置」也就是：

$$SWAP|i\rangle\langle j| = |j\rangle\langle i| \quad (76)$$

根据这个信息我们发现 Choi State 可以写作：

$$J(\mathcal{E}) = \frac{1}{3} (I \otimes I + SWAP) \quad (77)$$

$$= \frac{2}{3} (|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+|) \quad (78)$$

于是很显然我们发现这个 Choi State 可以写作三个 vectorization 的矩阵的形式：

$$\begin{aligned} A_0 &= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{\sqrt{3}} = \frac{1}{\sqrt{3}}I, \\ A_1 &= \frac{|0\rangle\langle 0| - |1\rangle\langle 1|}{\sqrt{3}} = \frac{1}{\sqrt{3}}Z, \\ A_2 &= \frac{|0\rangle\langle 1| + |1\rangle\langle 0|}{\sqrt{3}} = \frac{1}{\sqrt{3}}X. \end{aligned} \quad (79)$$

并且我们可以验证这些 Kraus Operators 满足完备关系：

$$\sum_i A_i^\dagger A_i = \frac{1}{3}(I^2 + Z^2 + X^2) = \frac{1}{3}(3I) = I. \quad (80)$$

所以我们成功地通过 Choi State 得到了 Kraus Operators。

#### 不唯一的Kraus Operators表示

但是另一个观察会发现这个 Kraus Operators 的表示并不是唯一的。我们可以进行一个 unitary 变换，可以得到：

$$J(\mathcal{E}) = \frac{2}{3}|00\rangle\langle 00| + \frac{2}{3}|11\rangle\langle 11| + \frac{2}{3}|\psi^+\rangle\langle \psi^+|. \quad (81)$$

我们分别验证这个 Choi State 对应的 Kraus Operators：

$$A'_0 = \sqrt{\frac{2}{3}}|0\rangle\langle 0|, \quad A'_1 = \sqrt{\frac{2}{3}}|1\rangle\langle 1|, \quad A'_2 = \frac{1}{\sqrt{3}}X. \quad (82)$$

发现他们也满足完备关系。

#### 通过Choi State得到Kraus Operators的一般方法

我们，嗯分析上面的方法发现可以一套完备的方法通过 Quantum Channel 得到 Choi State 再得到 Kraus Operators：

1. **计算量子态的 Choi State**：通过  $J(\mathcal{E}) = (\mathcal{E} \otimes I)(|Vec(I)\rangle\langle Vec(I)|)$  计算 Choi State。
2. **Choi State 的谱分解**：计算 Choi State 的谱分解  $J(\mathcal{E}) = \sum_i \lambda_i |\psi_i\rangle\langle \psi_i|$ 。
3. **得到 Kraus Operators**：通过  $A_i = \sqrt{\lambda_i} |\psi_i\rangle\langle \psi_i|$  得到 Kraus Operators。
4. **检查完备关系**：检查  $\sum_i A_i^\dagger A_i = I$  是否成立。

#### Unitary Mixing Freedom

上面已经发现 Kraus Operators 的表示并不是唯一的。我们可以通过一个 unitary 变换得到另外一组 Kraus Operators。我们总结一下这个结论：



**Theorem 9. Unitary Mixing**

对于任意的一个量子态 (显然 Choi State 是一个量子态) 的两种不同的谱分解:

$$\rho = \sum_k \lambda_k |\lambda_k\rangle\langle\lambda_k|. \quad (83)$$

$$\rho = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|. \quad (84)$$

那么这两组本征态之间存在一个 unitary 变换  $U$  使得:

$$\sqrt{p_i}|\varphi_i\rangle = \sum_k U_{ik} \sqrt{\lambda_k} |\lambda_k\rangle. \quad (85)$$

下面我们证明这个结论。首先根据两种谱分解我们给出两种将  $\rho$  purification 的方式:

$$|\psi\rangle = \sum_k \sqrt{\lambda_k} |\lambda_k\rangle \otimes |k\rangle, \quad |\varphi\rangle = \sum_i \sqrt{p_i} |\varphi_i\rangle \otimes |i\rangle. \quad (86)$$

Schmidt Decomposition 定理告诉我们这两种 purification 方式之间存在一个 unitary 变换  $U$  使得:

$$|\varphi\rangle = (I \otimes U)|\psi\rangle. \quad (87)$$

我们写开会发现:

$$\begin{aligned} \sqrt{p_i}|\varphi_i\rangle &= (I \otimes \langle i|)(I \otimes U)|\psi\rangle \\ &= \sum_k \sqrt{\lambda_k} (I \otimes \langle i|U)(|\lambda_k\rangle \otimes |k\rangle) \\ &= \sum_k \sqrt{\lambda_k} |\lambda_k\rangle \cdot (\langle i|U|k\rangle) \\ &= \sum_k \sqrt{\lambda_k} U_{ik} |\lambda_k\rangle. \end{aligned} \quad (88)$$

□然后我们会发现上面的 general 量子态的定理如果使用在 Choi State 上面就得到了我们下面的结论:

**Theorem 10. Unitary Mixing Freedom for Kraus Operators**

对于任意的一个 Quantum Channel  $\mathcal{E}$  的两组不同的 Kraus Operators 表示  $\{A_i\}$  和  $\{B_j\}$ , 那么这两组 Kraus Operators 之间存在一个 unitary 变换  $U$  使得:

$$B_j = \sum_i U_{ji} A_i. \quad (89)$$

其中就是使用了 Vectorization 线性的性质, 可以把外面的 unitary 拉进去!

**5.2 Representation of Quantum Channel****Completely Positive**

我们最后回来讨论一下什么样子的映射才是一个 Quantum Channel。这里特别强调一下 completely positive 的概念。

- **Positive Map:** 一个映射  $\mathcal{E}$  是 positive map 如果对于任意的正定矩阵  $\rho \geq 0$  都有  $(I \otimes \mathcal{E})(\rho) \geq 0$ 。

例子: Positive Map but not Completely Positive Map

一个最基本的例子就是转置映射  $T(\rho) = \rho^T$ 。我们验证这个映射是 positive map 但是不是 completely positive map。

我们考虑转置映射  $\otimes$  Identity 映射作用在一个 bell state 上面:

$$(T \otimes I)(|\phi^+\rangle\langle\phi^+|) \quad (90)$$

发现会给出 SWAP Operator。我们知道 SWAP Operator 不是正定的所以这个映射不是 completely positive map。

Representation Thm

[我懒得写了真的 (((]

### 5.3 Questions and Thoughts

## 6 Lecture 7: Vectorization 的数学技巧

### 6.1 Vectorization 和 Quantum Channel 结合

#### Vectorization的定义

基本定义我重新 copy paste 在这里。

#### Definition 7. Vectorization

也就是把一个矩阵  $A$  变成一个向量  $|Vec(A)\rangle$ ，具体的定义是：

$$A = \sum_{ij} A_{ij} |i\rangle\langle j| \rightarrow |Vec(A)\rangle = \sum_{ij} A_{ij} |i\rangle \otimes |j\rangle. \quad (91)$$

一个很重要的恒等式：

**Lemma 2.** 对于任意的矩阵  $A, B, C$ ，都有：

$$|Vec(ABC)\rangle = (A \otimes C^T) |Vec(B)\rangle. \quad (92)$$

#### Vectorization以及内积

对于 Matrix 我们可以定义一个内积：

$$\langle A, B \rangle_{\text{HS}} := \text{Tr}(A^\dagger B). \quad (93)$$

然后通过计算会发现这个内积和 vectorization 的内积是等价的：

$$\text{Tr}(A^\dagger B) = \langle \text{vec}(A) | \text{vec}(B) \rangle \quad (94)$$

然后我们会发现。【期望值可以理解为 Hermite Observable】和 density matrix 的内积，所以我们有：

$$\text{Tr}[O\mathcal{E}(\rho)] = \langle \text{vec}(O) | \text{vec}(\mathcal{E}(\rho)) \rangle \quad (95)$$

所以，我们就可以使用内积的方法来计算量子态的期望值了！

#### Quantum Channel的Vectorization表示

下一个问题就是怎么一个 density matrix 的 vectorization 和其经过 Quantum Channel 作用之后的 vectorization 联系起来。对于一个 Unitary Channel 我们有：

$$|\text{vec}(U\rho U^\dagger)\rangle = (U \otimes U^*) |\text{vec}(\rho)\rangle. \quad (96)$$

因此对于一般的 Kraus Operator 定义的 Quantum Channel 我们有：

$$|\text{vec}(\mathcal{E}(\rho))\rangle = \sum_K (K \otimes K^*) |\text{vec}(\rho)\rangle. \quad (97)$$

### 复合的channel

我们考虑一个 Quantum Channel 是由一系列 Unitary Operator 的乘积构成的：

$$U = U_n U_{n-1} \cdots U_1. \quad (98)$$

$$\rho \longrightarrow U \rho U^\dagger \quad (99)$$

那么我们考虑一般的复合 Channel 定义，考虑两组 Kraus Operator 给出的 Channel：

$$\mathcal{E}^{(1)}(\rho) = \sum_j K_j^{(1)} \rho \left(K_j^{(1)}\right)^\dagger, \quad \mathcal{E}^{(2)}(\rho) = \sum_k K_k^{(2)} \rho \left(K_k^{(2)}\right)^\dagger. \quad (100)$$

我们考虑复合的 Channel 我们发现 Vectorization 我们有：

$$|\text{vec}(\mathcal{E}(\rho))\rangle = \left[ \sum_k \left( K_k^{(2)} \otimes \left( K_k^{(2)} \right)^* \right) \right] \left[ \sum_j \left( K_j^{(1)} \otimes \left( K_j^{(1)} \right)^* \right) \right] |\text{vec}(\rho)\rangle. \quad (101)$$

## 6.2 Average Over Unitary

### 基本概念

我们考虑一个客观测量的期望值，但是我们考虑如果量子态进行了一些随机的 Unitary 变换之后的期望值。也就是说我们考虑下面这个量子态：

$$E_\theta \text{Tr} \left[ U_\theta \rho U_\theta^\dagger O \right], \quad (102)$$

我们发现使用 Vectorization 的语言我们有：

$$\text{Tr} \left[ U_\theta \rho U_\theta^\dagger O \right] = \langle \text{vec}(O) | (U_\theta \otimes U_\theta^*) | \text{vec}(\rho) \rangle. \quad (103)$$

#### Definition 8. Average Over Unitary

对于进行随机 Unitary 演化之后的期望值，满足关系：

$$E_\theta \text{Tr} \left[ U_\theta \rho U_\theta^\dagger O \right] = \langle \text{vec}(O) | (E_\theta [U_\theta \otimes U_\theta^*]) | \text{vec}(\rho) \rangle. \quad (104)$$

其中我们定义 **Moment Operator** 为：

$$M \equiv E_\theta [U_\theta \otimes U_\theta^*], \quad (105)$$

### 例子：随机Z方向旋转

我们考虑一个随机 Unitary 的在 Bloch 球上沿 Z 方向转动的例子，每一个转动角度的概率都是一样的。我们定义 Unitary 算符是：

$$U_z(\theta) = e^{-i\theta Z/2} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}, \quad (106)$$

我们计算这个算符的  $U_\theta \otimes U_\theta^*$  然后写成矩阵形式！并且对于  $\theta$  进行均匀分布的 average，我们得到：

$$\begin{aligned} M &= \frac{1}{2\pi} \int_0^{2\pi} \text{diag}(1, e^{-i\theta}, e^{i\theta}, 1) d\theta. \\ \frac{1}{2\pi} \int_0^{2\pi} e^{\pm i\theta} d\theta &= 0, \\ M &= \text{diag}(1, 0, 0, 1). \end{aligned} \quad (107)$$

中间两个 0 是因为我们平均的时候 average 掉了所有的震动。所以写作 bra 和 ket 的形式有：

$$M = |00\rangle\langle 00| + |11\rangle\langle 11|. \quad (108)$$

我们再把这个 Moment Operator 写作 Kraus Operator 作用在 vectorization 之后的 density matrix 上面：

$$|\text{vec}(\mathcal{E}(\rho))\rangle = M|\text{vec}(\rho)\rangle \quad (109)$$

$$= (|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|) |\text{vec}(\rho)\rangle \quad (110)$$

$$(111)$$

我们从中读出两个 Kraus Operator 于是发现这个随机 channel 就是 complete dephasing channel！

### 高阶的 Moment Operator

我们考虑客观测量的「更高阶」平均值，类比  $\langle x^2 \rangle$  我们有：

$$E_\theta \left( \text{Tr} [U_\theta \rho U_\theta^\dagger]^k \right). \quad (112)$$

**注意！** 我们是先取 Tr 再取 k 次方！我们知道 Tr 的 k 次方可以写作下面的形式：

$$\text{Tr}[X]^2 = \text{Tr}[X \otimes X] \quad (113)$$

因此我们有

$$E_\theta \left( \text{Tr} [U_\theta \rho U_\theta^\dagger]^k \right) = E_\theta \left( \text{Tr} [(U_\theta \rho U_\theta^\dagger)^{\otimes k}] \right) = E_\theta \left( \text{Tr} [(U_\theta^{\otimes k} \rho^{\otimes k} (U_\theta^\dagger)^{\otimes k}) O^{\otimes k}] \right). \quad (114)$$

第三步之中我们意识到，矩阵的乘法和张量积的顺序是可以交换的，因为两个操作互不干扰！所以我们使用 Vectorization 的语言就可以有：

$$E_\theta \left( \text{Tr} [U_\theta \rho U_\theta^\dagger]^k \right) = E_\theta \left( \langle \text{vec}(O^{\otimes k}) | U_\theta^{\otimes k} \otimes U_\theta^{*\otimes k} | \text{vec}(\rho^{\otimes k}) \rangle \right) = \langle \text{vec}(O^{\otimes k}) | M_k | \text{vec}(\rho^{\otimes k}) \rangle \quad (115)$$

其中我们定义高阶的 Moment Operator 为：

### Definition 9. 高阶 Moment Operator

我们定义高阶的 Moment Operator 为：

$$M_k = E_\theta [U_\theta^{\otimes k} \otimes U_\theta^{*\otimes k}]. \quad (116)$$

### Remark:

注意！张量积是不能定义顺序的！所以我们高阶 moment operator 一定不可以写作  $(E_\theta [U_\theta \otimes U_\theta^*])^{\otimes k}$  这种形式！这是完全错误的。

### 张量积和Vectorization不可换

我们需要注意的是，有两个很像的东西是完全不一样的：

$$|\text{Vec}(A \otimes B)\rangle \neq |\text{Vec}(A)\rangle \otimes |\text{Vec}(B)\rangle. \quad (117)$$

但是这两个东西长得很像但是有一点点特别需要注意的差距。我们分别展开计算得到：

$$|\text{vec}(A)\rangle \otimes |\text{vec}(B)\rangle = \sum_{i,j,k,l=1}^n a_{ij} b_{kl} |ij\rangle \otimes |kl\rangle = \sum_{i,j,k,l=1}^n a_{ij} b_{kl} |ijkl\rangle \quad (118)$$

但是，计算  $|\text{vec}(A \otimes B)\rangle$  我们有：

$$|\text{vec}(A \otimes B)\rangle = \sum_{i,j,k,l=1}^n a_{ij} b_{kl} |ik\rangle \otimes |jl\rangle = \sum_{i,j,k,l=1}^n a_{ij} b_{kl} |ikjl\rangle. \quad (119)$$

其中唯一的区别是 **ket 的张量积的顺序是不一样的**。导致这个区别的主要问题是，vectorization 的定义是把所有的 bra 变成的 ket 张量积在原有的 ket 后面的。这需要特别注意。

所以我们还有一个超级重要的结论就是：

$$\langle \text{vec}(C \otimes D) | \text{vec}(A \otimes B) \rangle = \langle \text{vec}(C) | \text{vec}(A) \rangle \langle \text{vec}(D) | \text{vec}(B) \rangle, \quad (120)$$

### 例子：计算方差

我们计算一个 random unitary 变换之后的期望值的方差：

$$\begin{aligned} \text{Var}_\theta[f_\theta] &= E_\theta \left( \text{Tr} [U_\theta \rho U_\theta^\dagger O] \right)^2 - \left( E_\theta \left( \text{Tr} [U_\theta \rho U_\theta^\dagger O] \right) \right)^2 \\ &= \langle \text{vec}(O^{\otimes 2}) | M_2 | \text{vec}(\rho^{\otimes 2}) \rangle - (\langle \text{vec}(O) | M_1 | \text{vec}(\rho) \rangle)^2. \end{aligned} \quad (121)$$

其中就是使用了方差的计算公式。并且二阶的 Moment Operator 帮助我们完成了计算。

## 6.3 General map Vectorization

### General map的Vectorization表示

我们考虑一个一般的 map 作用在一个 state 上面，前后 vectorization 的关系。我们考虑作用在一个 density matrix 上的一般 map：

$$\mathcal{E}(|i\rangle\langle j|) = \sum_{k,l} c_{kl}^{ij} |k\rangle\langle l|. \quad (122)$$

前后的 state 我们可以使用 vectorization 进行表示：

$$|\text{vec}(\mathcal{E}(\rho))\rangle = \sum_{i,j} \sum_{k,l} \rho_{ij} c_{kl}^{ij} |k\rangle |l\rangle. \quad (123)$$

为此我们可以定义一个 Operator 进行描述这个 general 的操作：

$$M = \sum_{i,j,k,l} c_{kl}^{ij} |k\rangle\langle l| \langle i|j|. \quad |\text{vec}(E(\rho))\rangle = M |\text{vec}(\rho)\rangle, \quad (124)$$

#### Transpose的Vectorization表示

我们知道 transpose 并不是一个 Quantum Channel，因为它不满足 complete positivity。但是我们依然可以考虑 transpose map 的 vectorization 表示。我们有：

$$\mathcal{E}(|i\rangle\langle j|) = |j\rangle\langle i|. \quad (125)$$

然后进行 vectorization 我们有：

$$M = \sum_{i,j} |j\rangle\langle i| \langle i|j| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \text{SWAP} \quad (126)$$

## 6.4 Pauli Transfer Matrix

对于 1 qubit 上面的任意 quantum operation 我们前面一直在使用其对于 computational basis 的作用描述这个 operation。但是我们也可以使用 Pauli basis 来描述这个 quantum operation。这种描述方式叫做 Pauli Transfer Matrix (PTM) representation。

我们首先定义一个 quantum operation 作用在 Pauli Matrix 上面，得到的结果是：

$$\mathcal{E}(P_i) = \sum_j C_j^i P_j, \quad (127)$$

其中  $P_i, P_j \in \{I, X, Y, Z\}$ 。我们可以为每一个 Pauli 矩阵赋予一个 ket 进行表示：

$$I \equiv |0\rangle, \quad X \equiv |1\rangle, \quad Y \equiv |2\rangle, \quad Z \equiv |3\rangle. \quad (128)$$

下面我们就可以把 quantum operation 写作一个矩阵作用在这些 ket 上面，这个矩阵我们定义为 **Pauli Transfer Matrix**，我们有：

$$M = \sum_{ij} C_j^i |j\rangle\langle i|, \quad (129)$$

#### Hadamard Gate的PTM表示

考虑下面这样的 Hadamard Gate:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (130)$$

我们计算其作用在 Pauli 矩阵之后的结果，使用 Pauli 矩阵进行展开：

$$H I H^\dagger = I, \quad H X H^\dagger = Z, \quad H Y H^\dagger = -Y, \quad H Z H^\dagger = X. \quad (131)$$

因此我们可以写出 Hadamard Gate 的 PTM 表示：

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \quad (132)$$

#### PTM表示与Vectorization后的表示

对于一个 Channel 我们发现 PTM 的表示和 Vectorization 的表示是有联系的。我们考虑 Hadmand Gate 作用前后的 Vectorization：

$$|\text{vec}(E(\rho))\rangle = M |\text{vec}(\rho)\rangle, \quad (133)$$

可以求出其中矩阵是：

$$M = H \otimes H^* = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}. \quad (134)$$

我们发现只要对其做一个相似变换，换一组基就可以得到 PTM 表示。并且换的基正好就是 Pauli 矩阵 Vectorization 之后的基，也就是 Bell Basis。

## 6.5 Questions and Thoughts

### Question 6.6 Vectorization 可以跟那些操作进行互换？

我们发现求和是和 vectorization 可以互换的！！所以数乘和求和对于 vectorization 是可以互换的。□

### Question 6.7 对于一个系统，我们对于一部分进行 Unitary 演化和对于另一部分进行 Tr 是可以交换的吗？

是的!!! 啊啊啊啊啊啊啊啊!!

□

### Question 6.8 张量积的 Tr 和之前是什么关系捏？

我们考虑关系：

$$\text{Tr}(A \otimes B) = (\text{Tr}A)(\text{Tr}B) \quad (135)$$



## 7 Lecture 8: Shot Noise

### 7.1 Shot Noise 讨论

测量一个量子系统对于某个客观测量  $M$  的期望值的时候我们从不可以只测一次，需要进行多次测量然后取平均得到。我们称这个结果是 estimator，也就是一个期望值的估量，同时也是一个随机变量。

我们下面希望研究这个随机变量在什么程度上可以近似真实的期望值，同时也希望定量的描述其误差。

#### 7.1.1 Estimator 的基本概念

##### Estimator 的定义

我们根据实际情况进行下面的定义：

##### Definition 10. Estimator

对于一个系统  $\rho$ ，我们希望测量某个观测量  $M$  的期望值  $\langle M \rangle = \text{Tr}(\rho M)$ 。我们在  $M = \sum_n \lambda_n |n\rangle\langle n|$ 。我们制备  $N$  份这样的系统并进行独立测量。对于第  $s$  次测量给出的结果为  $\lambda_{ms}$ ，那么我们定义 *Estimator* 为：

$$M_N = \frac{1}{N} \sum_{s=1}^N \lambda_{m(s)} . \quad (136)$$

##### Estimator 概率论计算

对于 Estimator 本身就是一个随机变量，其意义是测量  $N$  次的平均结果。并且我们可以计算其期望值和方差：

##### Theorem 11. Estimator 的期望值和方差

对于上面的定义，我们有：

$$\mathbb{E}(M_N) = \langle M \rangle = \text{Tr}(\rho M) , \quad (137)$$

并且

$$\text{Var}(M_N) = \frac{\text{Var}_\rho(M)}{N} \quad (138)$$

其中：

$$\text{Var}_\rho(M) = \text{Tr}(\rho M^2) - (\text{Tr}(\rho M))^2 \quad (139)$$

其具体证明的核心在于，我们知道每一次测量结果是独立的。所以可以把每一次测量当作一个随机过程，然后使用概率论的独立随机变量的性质进行计算。我们可以写作一个定理：

**Theorem 12.** 独立随机变量的期望值和方差

对于  $N$  个独立随机变量  $X_1, X_2, \dots, X_N$ , 我们有:

$$\mathbb{E} \left( \sum_{i=1}^N X_i \right) = \sum_{i=1}^N \mathbb{E}(X_i), \quad \text{Var} \left( \sum_{i=1}^N X_i \right) = \sum_{i=1}^N \text{Var}(X_i) . \quad \text{Var}(\alpha X) = \alpha^2 \text{Var}(X) \quad (140)$$

**Remark:**

直观上, 后两个条件是有点矛盾的。但是中间的条件说的是独立随机变量可以直接求和; 后面一个条件说的是如果是两个不独立的随机变量相加, 但是正好两个一模一样, 那么方差就会变成  $4\text{Var}(X)$ 。所以并不矛盾。

因此对于 Estimator 的情况我们就有:

$$\mathbb{E} \left( \hat{M}_N \right) = \frac{1}{N} \sum_{s=1}^N \mathbb{E} \left( \lambda_{m(s)} \right) \quad \text{Var} \left( \hat{M}_N \right) = \frac{1}{N^2} \sum_{s=1}^N \text{Var} \left( \lambda_{m(s)} \right) \quad (141)$$

下面分别计算每一个独立随机变量的期望值和方差。我们知道每一个随机变量的取值为  $\lambda_{m(s)}$ , 并且其概率以及平均值为:

$$\mathbb{E} \left( \lambda_{m(s)} \right) = \sum_{m(s)} \lambda_{m(s)} \mathbb{P} \left( \lambda_{m(s)} \right) . \quad \mathbb{P}(\lambda_{m(s)}) = \langle \lambda_{m(s)} | \rho | \lambda_{m(s)} \rangle \quad (142)$$

$$\mathbb{E} \left( \lambda_{m(s)}^2 \right) = \sum_{m(s)} \lambda_{m(s)}^2 \mathbb{P}(m(s)) = \sum_{m(s)} \lambda_{m(s)}^2 \langle \lambda_{m(s)} | \rho | \lambda_{m(s)} \rangle = \text{Tr}(\rho M^2). \quad (143)$$

带入最初的公式也就得到了 estimator 的方差和期望值!!

**Remark:**

对于测量一个系统, 我们设计一个 Estimator, 我们需要保证其是 unbiased 的 (也就是期望值等于真实值)。同时我们希望其方差尽可能小, 这样才能保证测量结果的可靠性。因此我们需要验证期望值以及尽量减小方差。

### 7.1.2 测量方法的影响

我们会发现对于同一个物理量, 我们可以有不同的测量方法。我们分析一下不同的测量方法的性质区别。然后下一小节我们会意识到, 选择不同的测量方法会影响 Estimator 的收敛性质。

#### 两个纯态 Fidelity 测量

对于两个纯态的重叠度 Fidelity 我们存在两种测量方法:

- 对于两个态寻找一个 Unitary 的演化保证  $|\psi\rangle = U_\psi|0\rangle$ ,  $|\varphi\rangle = U_\varphi|0\rangle$ . 然后我们进行测量:

$$M = |0\rangle\langle 0| \quad \rho = U_\psi^\dagger |\psi\rangle\langle\psi| U_\psi \quad (144)$$

- 直接进行测量, 使用 SWAP 算符作为测量算符:

$$M = \text{SWAP} \quad \rho = |\psi\rangle\langle\psi| \otimes |\varphi\rangle\langle\varphi| \quad (145)$$

对于第一种情况我们计算：

$$\text{Var}_\rho(M) = |\langle \varphi | \psi \rangle|^2 - |\langle \varphi | \psi \rangle|^4. \quad (146)$$

而对于第二种情况我们计算：

$$\text{Var}_\rho(M) = 1 - |\langle \varphi | \psi \rangle|^4. \quad (147)$$

因此明显对于第一种方法方差更小，也就预示着我们的 Estimator 的方差更小。

#### 多比特Z测量

这种情况下可观测量的谱更加小!! 但是方差一样 [\[回头补充\]](#)

### 7.1.3 概率上界与收敛

下面我们使用随机过程的方法，研究 Estimator 的收敛性质。对于一个随机变量的收敛情况我们可以用下面的概率进行描述：

$$\mathbb{P}(|Z - \mathbb{E}(Z)| \geq t) \quad (148)$$

如果随机变量距离期望很远的概率很小，那么我们就认为这个随机变量是收敛的。下面我们将给出不等式约束这个取值。

#### Markov不等式

##### Theorem 13. Markov 不等式

对于一个非负随机变量  $Z \geq 0$ ，我们有对于任意  $t > 0$ ：

$$\mathbb{P}(Z \geq t) \leq \frac{\mathbb{E}(Z)}{t}. \quad (149)$$

具体证明也就是把随机变量的取值分为大于和小于  $t$  的两个部分然后进行估算。

#### Chebyshev不等式

##### Theorem 14. Chebyshev 不等式

对于一个随机变量  $Z$ ，我们有对于任意  $t > 0$ ：

$$\mathbb{P}(|Z - \mathbb{E}(Z)| \geq t) \leq \frac{\text{Var}(Z)}{t^2}. \quad (150)$$

证明暂且省略。但是我们发现如果这个随机变量  $Z = M_N$  的情况下，我们有：

$$\mathbb{P}\left(\left|\hat{M}_n - \text{Tr}(M\rho)\right| \geq t\right) \leq \frac{\text{Var}(\hat{M}_N)}{t^2} = \frac{\text{Var}_\rho(M)}{Nt^2}. \quad (151)$$

我们考虑我们希望误差小于  $\epsilon$ ，概率大于  $1 - \delta$ 。【其中  $\epsilon, \delta$  都是小量】。数学上就是满足：

$$\mathbb{P}\left(\left|\hat{M}_n - \text{Tr}(M\rho)\right| \geq \epsilon\right) \leq \delta. \quad (152)$$

那么我们带入上面的不等式就知道，充分条件「并不必要」：

$$\frac{\text{Var}_\rho(M)}{N\epsilon^2} \leq \delta \implies N \geq \frac{\text{Var}_\rho(M)}{\delta\epsilon^2} \quad (153)$$

也就是说：

- 测量次数多余于  $\frac{\text{Var}_\rho(M)}{\delta\epsilon^2}$  次就可以！那么显然如果  $\text{Var}_\rho(M)$  比较小，那么测量次数就可以少一些。

所以我们倾向于选择方差比较小的测量方法进行测量，这样可以减少测量次数。

### Hoeffding不等式

我们知道  $M_N$  是  $N$  个独立随机变量的平均值，如果我们知道每一个随机变量的取值范围，那么我们可以使用 Hoeffding 不等式进行更紧的约束。

#### Theorem 15. Hoeffding 不等式

对于  $N$  个独立随机变量  $X_1, X_2, \dots, X_N$ ，并且每一个随机变量的取值范围为  $a_i \leq X_i \leq b_i$ 。那么考虑另一个随机变量  $S_N = \sum_{i=1}^N X_i$ ，我们有对于任意  $t > 0$ ：

$$\mathbb{P}(|S_n - \mathbb{E}(S_n)| \geq t) \leq 2 \exp\left(-\frac{2t^2}{\sum_i (b_i - a_i)^2}\right). \quad (154)$$

我们省略证明，讨论对于 Esitimator 的情况。Esitimator 可以写作：

$$\hat{M}_N = \frac{1}{N} \sum_{s=1}^N \lambda_{m(s)} = \sum_{s=1}^N \frac{\lambda_{m(s)}}{N}. \quad (155)$$

我们考虑测量的客观测量的谱的上下界是：

$$\frac{\lambda_{\min}}{N} \leq \frac{\lambda_{m(s)}}{N} \leq \frac{\lambda_{\max}}{N}. \quad (156)$$

我们记：

$$\sum_i (b_i - a_i)^2 = \sum_{s=1}^N \left(\frac{\lambda_{\max} - \lambda_{\min}}{N}\right)^2 = \frac{\Delta\lambda^2}{N}. \quad (157)$$

也就是说：

$$\mathbb{P}\left(\left|\hat{M}_N - \text{Tr}(\rho M)\right| \geq t\right) \leq 2 \exp\left(-\frac{2Nt^2}{\Delta\lambda^2}\right). \quad (158)$$

如果我们希望误差小于  $\epsilon$ ，概率大于  $1 - \delta$ 。【其中  $\epsilon, \delta$  都是小量】。数学上就是满足：

$$2 \exp\left(-\frac{2N\epsilon^2}{\Delta\lambda^2}\right) \leq \delta \implies N \geq \frac{\Delta\lambda^2}{2\epsilon^2} \ln\left(\frac{2}{\delta}\right). \quad (159)$$

这告诉我们：

- 可观测量的谱范围  $\Delta\lambda$  越小，测量次数就可以越少。

## 7.2 Question and thoughts

### Question 7.9 什么叫一个 Estimator 是 unbiased 的?

也就是我们计算这个 Estimator 的期望值, 会发现等于我们希望测量的真实数值。一般我们都是希望使用 Estimator 测量一个系统的可观测量的期望值, 所以我们就是需要证明:

$$\mathbb{E}(M_N) = \text{Tr}(\rho M) \quad (160)$$

左边的证明也就用到我们最初求解 Estimator 期望值的过程。  $\square$

### Question 7.10 Unitary 可以把一个纯态变成混态吗

不可以。因为 Unitary 演化保持态的纯度不变。也就是说如果  $\rho^2 = \rho$ , 那么  $U\rho U^\dagger U\rho U^\dagger = U\rho U^\dagger$  也是成立的。  $\square$

## 8 Lecture 9: 区分量子态与矩阵模长

### 8.1 区分量子态

#### 量子态区分概率最大化

给定两个量子态  $\rho_0$  以及  $\rho_1$  我们希望通过 POVM 对其进行区分。假设我们使用的 POVM 是有两个元素  $\{M_0, M_1\}$ 。其中  $M_i$  回答的问题是是不是  $\rho_i$  态。我们给出下面定理：

#### Theorem 16. 量子态区分概率

给定两个量子态  $\rho_0$  以及  $\rho_1$ ，其先验概率分别假设为  $1/2$ 。我们使用  $POVM\{M_0, M_1\}$  对其进行区分，我们【主观上认为】测量到  $M_0$  结果量子态就是  $\rho_0$ ，测量到  $M_1$  结果量子态就是  $\rho_1$ 。那么我们成功区分的概率为：

$$p = \frac{1}{2} + \frac{1}{2} \text{Tr}(M_0(\rho_0 - \rho_1)) . \quad (161)$$

Proof: 我们计算成功概率：

$$\mathbb{P}(\text{succes}) = \mathbb{P}(\rho = \rho_0)\mathbb{P}(\text{measure } 0 \mid \rho = \rho_0) + \mathbb{P}(\rho = \rho_1)\mathbb{P}(\text{measure } 1 \mid \rho = \rho_1) \quad (162)$$

$$= \frac{1}{2} \text{Tr}(M_0\rho_0) + \frac{1}{2} \text{Tr}(M_1\rho_1). \quad (163)$$

根据 POVM 的完备性我们有  $M_1 = I - M_0$ ，可以直接给出上面的结论。

#### Remark:

注意，我们是主观上认为测量到  $M_0$  结果量子态就是  $\rho_0$ ，测量到  $M_1$  结果量子态就是  $\rho_1$ 。实际上如果不小心设计了一个  $M_0$  导致  $\rho_0$  永远测不到这个结果，我们只能说这个测量设定很不合理，但是也不是不行。

#### Theorem 17. 量子态区分最优 POVM

给定两个量子态  $\rho_0$  以及  $\rho_1$ ，其先验概率分别假设为  $1/2$ 。我们使用  $POVM\{M_0, M_1\}$  对其进行区分，我们主观上认为测量到  $M_0$  结果量子态就是  $\rho_0$ ，测量到  $M_1$  结果量子态就是  $\rho_1$ 。那么使得成功区分概率最大的 POVM 为：

$$p_{opt} = \max_{0 \leq M_0 \leq I} \frac{1 + \text{Tr}(M_0(\rho_0 - \rho_1))}{2} \quad (164)$$

其中  $M_0$  是任意的正半定矩阵且  $M_0 \leq I$ 。

#### 例：纯态区分

我们考虑区分两个纯态  $|\psi_0\rangle$  以及  $|\psi_1\rangle$ 。我们不妨构建一个 POVM 保证  $M_0 = |0\rangle\langle 0|$  以及  $M_1 = |1\rangle\langle 1|$ ，其中  $|0\rangle$  以及  $|1\rangle$  是两个态的正交归一化基底。我们计算成功概率：

$$p = \frac{1 + \text{Tr}(M_0(\rho_0 - \rho_1))}{2} = \frac{1 + 1}{2} = 1. \quad (165)$$

因此我们可以完美区分两个纯态。

#### 例：混态区分

对于 single qubit 的混态我们考虑区分下面两个态：

$$\rho_0 = \frac{1}{2}(I + \vec{r}_0 \cdot \vec{\sigma}), \quad \rho_1 = \frac{1}{2}(I + \vec{r}_1 \cdot \vec{\sigma}). \quad (166)$$

我们不妨更一般的构建下面一个 POVM:

$$M_0 = a(I + \vec{x} \cdot \vec{\sigma}), \quad M_1 = I - M_0. \quad (167)$$

但是 generally 这两个矩阵需要满足一些约束条件才能是一个合理的 POVM，我们下面给出定理：

#### Theorem 18. Single qubit POVM 的约束条件

给定一个 single qubit 的 POVM  $\{M_0, M_1\}$ ，其中  $M_0 = a(I + \vec{x} \cdot \vec{\sigma})$  以及  $M_1 = I - M_0$ 。那么其需要满足下面的约束条件：

$$\|\vec{x}\| \leq 1 \quad a \in \left[0, 1/(1 + \|\vec{x}\|^2)\right] \quad (168)$$

Proof: 我们干的事情就是计算两个的本征值，保证本征值非复即可。

在这个基础上我们计算这两个混态的区分概率：

#### Theorem 19. Single qubit 混态区分概率

给定两个 single qubit 混态  $\rho_0 = \frac{1}{2}(I + \vec{r}_0 \cdot \vec{\sigma})$  以及  $\rho_1 = \frac{1}{2}(I + \vec{r}_1 \cdot \vec{\sigma})$ ，其先验概率分别假设为  $1/2$ 。我们使用 POVM  $\{M_0, M_1\}$  对其进行区分，我们主观上认为测量到  $M_0$  结果量子态就是  $\rho_0$ ，测量到  $M_1$  结果量子态就是  $\rho_1$ 。那么成功区分的概率为：

$$p = \frac{1}{2} + \frac{a}{4}(2\vec{x} \cdot (\vec{r}_0 - \vec{r}_1)). \quad (169)$$

这个计算之中我们使用了 Pauli 矩阵的 trace 性质：

$$\text{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}, \quad \text{Tr}(\sigma_i) = 0. \quad (170)$$

我们使用 Pauli Basis 展开 single qubit 的量子态的原因正是因为这个性质。

#### Theorem 20. Single qubit 混态区分最优概率

对于上方的 set up，使得成功区分概率最大的 POVM 为：

$$\vec{x}_{\text{opt}} = \frac{\vec{r}_0 - \vec{r}_1}{\|\vec{r}_0 - \vec{r}_1\|}, \quad M_{0,\text{opt}} = \frac{1}{2}(I + \vec{x}_{\text{opt}} \cdot \vec{\sigma}). \quad (171)$$

此时成功区分的最优概率为：

$$p_{\text{opt}} = \frac{1}{2} + \frac{1}{4} \left( \frac{(\vec{r}_0 - \vec{r}_1) \cdot (\vec{r}_0 - \vec{r}_1)}{\|\vec{r}_0 - \vec{r}_1\|} \right) = \frac{1}{2} + \frac{1}{4} \|\vec{r}_0 - \vec{r}_1\|. \quad (172)$$

Proof: 在上面约束条件下面进行放缩就好了!! 这里我们最后一步放缩的时候会使用:

$$p = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{1 + \|\vec{x}\|^2} \|\vec{x}\| \|\vec{r}_0 - \vec{r}_1\| \cos(\vec{x}; \vec{r}_0 - \vec{r}_1). \quad (173)$$

然后我们对此考虑到函数:

$$f(t) = \frac{t}{1 + t^2}, \quad t \geq 0 \quad (174)$$

在  $t = 1$  处取得最大值  $1/2$ 。然后我们带回去就得到了最终结果!!

### 数学工具: 1-Norm

上面讨论的都是单个 bit 的情况, 我们考虑更高维度的量子态区分。为此我们先要定义 1-norm 的数学工具:

#### Definition 11. Matrix 1-Norm

对于一个矩阵  $X$  我们定义其 1-Norm 为:

$$\|X\|_1 = \text{Tr}(\sqrt{X^\dagger X}) = \sum_i \sqrt{\lambda_i(X^\dagger X)}, \quad (175)$$

其中  $\lambda_i(X^\dagger X)$  是矩阵  $X^\dagger X$  的本征值。

显然对于 Unitary 的矩阵  $U$  我们有  $\|U\|_1 = \text{Tr}(I) = d$ , 其中  $d$  是矩阵的维度。对于 Hermitean 矩阵  $H$  我们有  $\|H\|_1 = \sum_i |\lambda_i(H)|$ , 其中  $\lambda_i(H)$  是矩阵  $H$  的本征值。

### higher dimension state区分

我们考虑区分两个高维度量子态  $\rho_0$  以及  $\rho_1$ 。我们有下面定理:

#### Theorem 21. Higher dimension 量子态区分概率

给定两个量子态  $\rho_0$  以及  $\rho_1$ , 其先验概率分别假设为  $1/2$ 。我们使用  $POVM\{M_0, M_1\}$  对其进行区分, 我们主观上认为测量到  $M_0$  结果量子态就是  $\rho_0$ , 测量到  $M_1$  结果量子态就是  $\rho_1$ 。那么成功区分的概率为:

$$p_{opt} = \frac{1}{2} \left( 1 + \max_{0 \leq M \leq I} \text{Tr}(M(\rho_0 - \rho_1)) \right) = \frac{1}{2} + \frac{1}{4} \|\rho_0 - \rho_1\|_1. \quad (176)$$

Proof: 其中我们显然使用了一个恒成立的数学关系:

$$\frac{1}{2} \|\rho_0 - \rho_1\|_1 = \max_{0 \leq M \leq I} \text{Tr}(M(\rho_0 - \rho_1)). \quad (177)$$

证明的核心是构造一个只有  $\rho_0 - \rho_1$  本征值为正的本征子空间的投影算符作为 POVM 的一个元素。



## 8.2 矩阵模长

### 一般定义

上面我们给出了 1-Norm 的定义，下面我们逐渐给出一些常用的定义与性质。

#### Definition 12. 一般矩阵 Norm

对于一个矩阵如果我们希望赋予一个 Norm 的概念  $\|\cdot\| : K^{m \times n} \mapsto \mathbb{R}$ ，那么需要满足一些必要的性质：

- 非负性： $\|A\| \geq 0$
- 零点对应： $\|A\| = 0 \Leftrightarrow A = 0$
- 三角不等式： $\|A + B\| \leq \|A\| + \|B\|$
- 齐次性： $\|\alpha A\| = |\alpha| \|A\|$

辅助定义我们还需要定义一些概念：

#### Definition 13. Normal Matrix

对于一个矩阵  $A \in K^{n \times n}$ ，如果其满足  $AA^\dagger = A^\dagger A$ ，那么我们称其为 Normal Matrix。

显然 Hermitean 矩阵、Unitary 矩阵都是 Normal Matrix。我们考虑之前对于这两种 normal 矩阵的 1-Norm 的讨论可以有下面的定理：

#### Theorem 22. Normal Matrix 的 1-Norm

对于一个 Normal Matrix  $A \in K^{n \times n}$ ，其 1-Norm 可以表示为：

$$\|A\|_1 = \sum_i |\lambda_i(A)|, \quad (178)$$

其中  $\lambda_i(A)$  是矩阵  $A$  的本征值。

我们一般都只考虑 Normal Matrix 的情况，因为演化算子是 Unitary 的量子态和测量算子是 Hermitean 的。所以基本上都是 Normal Matrix。

#### Remark:

我们需要注意，对于 Unitary 矩阵来说  $\|U\|_1 = d$ ，其中  $d$  是矩阵的维度。这意味着：

- Unitary 矩阵的本征值模长之和等于矩阵的维度。
- Unitary 矩阵的本征值求和不一定等于矩阵的维度。

请区分模长求和与直接求和！！

## Schatten p-Norm

### Definition 14. Schatten p-Norm

对于一个矩阵  $A \in K^{m \times n}$ ，我们定义其 Schatten p-Norm 为：

$$\|A\|_p = \left( \sum_{i=1}^d \left| \sqrt{\lambda_i(A^\dagger A)} \right|^p \right)^{1/p}. \quad (179)$$

其中  $\lambda_i(A^\dagger A)$  是矩阵  $A^\dagger A$  的本征值。

对于一般的矩阵我们可以使用奇异值分解来分析其 Schatten p-Norm 的性质。我们有定理：

### Theorem 23. 矩阵的奇异值分解与 Schatten p-Norm

对于一个矩阵  $A \in K^{m \times n}$ ，其奇异值分解为  $A = U\Sigma V^\dagger$ ，其中  $U \in K^{m \times m}$  以及  $V \in K^{n \times n}$  是 Unitary 矩阵， $\Sigma \in K^{m \times n}$  是一个对角矩阵，其对角线元素为矩阵  $A$  的奇异值。我们有：

$$\|A\|_p = \left( \sum_{i=1}^d \sigma_i(A)^p \right)^{1/p}, \quad (180)$$

Proof: 这是因为矩阵的奇异值和  $A^\dagger A$  的本征值有关系：

$$\sigma_i = \sqrt{\lambda_i(A^\dagger A)} \quad (181)$$

同样的考虑对于 Normal Matrix 的情况，我们存在一个特例：

### Theorem 24. Normal Matrix 的 Schatten p-Norm

对于一个 Normal Matrix  $A \in K^{n \times n}$ ，其 Schatten p-Norm 可以表示为：

$$\|A\|_p = \left( \sum_i |\lambda_i(A)|^p \right)^{1/p}, \quad (182)$$

其中  $\lambda_i(A)$  是矩阵  $A$  的本征值。

- 我们从此只考虑 Normal Matrix 的情况。

## Examples of Schatten p-Norm

如果  $p = 1$  那么 Schatten 1-Norm 就是我们之前定义的 1-Norm。对于 Normal Matrix 来说：

### Theorem 25. Normal Matrix 的 Schatten 1-Norm

对于 Normal Matrix  $A \in K^{n \times n}$ ，其 Schatten 1-Norm 可以表示为：

$$\|A\|_1 = \sum_i |\lambda_i|. \quad (183)$$

对于  $p = 2$  我们有 Schatten 2-Norm:

$$\|A\|_2 = \left( \sum_i |\lambda_i|^2 \right)^{1/2}. \quad (184)$$

但是我们可以证明对于 Normal Matrix 其存在另一种更漂亮的形式

**Theorem 26.** *Normal Matrix 的 Schatten 2-Norm*

对于 Normal Matrix  $A \in K^{n \times n}$ , 其 Schatten 2-Norm 可以表示为:

$$\|A\|_2 = \sqrt{\text{Tr}(A^\dagger A)}. \quad (185)$$

等价的也可以表示为:

$$\|A\|_2 = \sqrt{\sum_{i,j} |A_{ij}|^2}. \quad (186)$$

我们可以使用量子计算机来计算两个量子态区别的 Schatten 2-Norm, 因为数学上我们有:

$$\begin{aligned} \|\rho_0 - \rho_1\|_2 &= \sqrt{\text{Tr}((\rho_0 - \rho_1)(\rho_0 - \rho_1))} \\ &= \sqrt{\text{Tr}(\rho_0^2) + \text{Tr}(\rho_1^2) - 2 \text{Tr}(\rho_0 \rho_1)}. \end{aligned} \quad (187)$$

其中的信息可以通过 SWAP 测试来获得。

$$\text{Tr}(\rho_i^2) = \text{Tr}((\rho_i \otimes \rho_i) \text{SWAP}), \quad \text{Tr}(\rho_1 \rho_2) = \text{Tr}((\rho_1 \otimes \rho_2) \text{SWAP}). \quad (188)$$

### Infinity Norm

我们考虑 Schatten p-Norm 在  $p \rightarrow \infty$  的极限情况, 我们有下面定理:

**Theorem 27.** *Normal Matrix 的 Infinity Norm*

对于 Normal Matrix  $A \in K^{n \times n}$ , 其 Schatten Infinity-Norm 可以表示为:

$$\|A\|_\infty = \max_i |\lambda_i(A)|, \quad (189)$$

其中  $\lambda_i(A)$  是矩阵  $A$  的本征值。

也就是说这个给出了矩阵的最大本征值模长。一个很现实的意义就是他给出了一个算子的最大期望值:

$$|\text{Tr}(M\rho)| \leq |\lambda_{\max}| = \|M\|_\infty. \quad (190)$$

## 9 Lecture 10: Entropy

### 9.1 Classical Entropy

#### 9.1.1 Single Random Variable Entropy

##### Shannon Entropy

首先我们定义经典信息理论之中的熵的概念：

##### Definition 15. Shannon Entropy

对于一个随机变量  $X$  其可能取值为  $\{x_i\}$ ，对应的概率分布为  $\{p_i\}$ ，我们定义这个随机变量的 *Shannon Entropy* 为：

$$H(X) = - \sum_i p_i \log p_i. \quad (191)$$

注意！对于经典信息理论这里的对数我们一般取以 2 为底！

对于经典的 Shannon Entropy 我们一般有三个 Interpretation：

- Uncertainty of Random Variable before knowing its value
- Learning  $X$  所需要的「知识量」
- 平均的“Surprisal” 也就是  $-\log p_i$  的平均值

数学上我们会需要完全不可能的事件  $p_i = 0$ ，我们定义  $0 \log 0 = 0$  作为一个解析延拓。

##### 存储信息的最少bit数

Shannon Entropy 存在一个重要的 interpretation 就是考虑下面情况：

- 假设有一组数据  $X_1, X_2, \dots, X_n$ ，其中  $X_i$  在某一个有限范围内取值，并且这一组数据出现每一个取值的比例是已知为  $p_a$ 。

下面我们的问题是：为了储存这一组数据，我们最少需要多少 bit？

##### 1234数据例子

假设现在存在一个存有  $n$  个数的数组，其中每一个数据都是 1, 2, 3, 4 中的一个，并且出现的比率分别为  $1/2, 1/4, 1/8, 1/8$ 。那么我们需要多少 bit 来储存这个数据呢？我们可以设计下面的编码方法：

**普通编码方案：**我们使用 2 bit 来储存每一个数据， $1 \rightarrow 00, 2 \rightarrow 01, 3 \rightarrow 10, 4 \rightarrow 11$ 。这样我们总共需要  $2n$  bit 来储存这组数据。

**优化编码方案：**我们使用 1 bit 来储存 1, 2 bit 来储存 2, 3 bit 来储存 3 和 4。也就是  $1 \rightarrow 0, 2 \rightarrow 10, 3 \rightarrow 110, 4 \rightarrow 111$ 。我们可以证明这样的编码方法是无歧义的，并且我们总共需要的 bit 数为：

$$n \left( \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3 \right) = \frac{7n}{4} < 2n \quad (192)$$

这个例子说明了，如果我们知道一个数组之中的数据的出现的比例「一种概率分布」，我们可以通过优化编码方式来减少储存数据所需要的 bit 数。

### Shannon Entropy作为最少bit数的下界

下面我们会给出 Shannon Entropy 一个终于要的性质，也就是 Shannon Entropy 给出了储存数据所需要的最少 bit 数的下界。

#### Theorem 28. Shannon's Noiseless Coding Theorem

假设有一个数组有  $n$  个数据，数据的取值为  $\{x_i\}, i = 1, \dots, m$ ，对应的概率分布为  $\{p_i\}$ 。那么为了储存这个数组，我们至少需要  $nH(X) + 1$  bit，其中  $H(X)$  为对应的 Shannon Entropy。

我们下面给出证明的思路。

**Binary Case:** 一个简单的情况就是整个数组只能取 0 和 1 两种值，并且 0 和 1 出现的概率分别为  $1 - p$  和  $p$ 。因此我们知道，长成这样子的  $n$  个数据的数组只有：

$$\binom{n}{pn} = \frac{n!}{(pn)!(n-pn)!} \quad (193)$$

种可能性。我们假设一个二进制数来记录这么多个数字，我们考虑需要多少位数，也就是：

$$L \geq \log_2 \binom{n}{pn} \quad (194)$$

$$= \log(n!) - \log((np)!) - \log((n(1-p))!) \quad (195)$$

$$\approx n \log(n) - n - (np \log(np) - np) - (n(1-p) \log(n(1-p)) - n(1-p)) \quad (196)$$

$$= -np \log(p) - n(1-p) \log(1-p) \quad (197)$$

$$= nH(p), \quad (198)$$

中间第二步我们使用了 Stirling 公式  $\ln(n!) \approx n \ln n - n + \frac{1}{2} \ln(2\pi n)$  也是一个重要的近似估计公式。

**General Case:** 对于一般的情况我们可以使用类似的方法进行估计。假设数据的取值为  $\{x_i\}, i = 1, \dots, m$ ，对应的概率分布为  $\{p_i\}$ 。我们考虑这样的序列有多少种构型：

$$\#config = \frac{n!}{(p_1 n)!(p_2 n)! \cdots (p_m n)!} \quad (199)$$

我们同样使用 Stirling 公式进行近似可以得到：

$$L \geq \log_2 \#config \quad (200)$$

$$= \log(n!) - \sum_{i=1}^m \log((p_i n)!) \quad (201)$$

$$\approx n \log(n) - n - \sum_{i=1}^m (p_i n \log(p_i n) - p_i n) \quad (202)$$

$$= -n \sum_{i=1}^m p_i \log(p_i) = nH(X). \quad (203)$$

### 9.1.2 2 Random Variables Entropy

更有趣的是当我们考虑两个随机变量的时候，我们可以定义更多的熵的概念来描述两个随机变量之间的关系。

#### Joint Entropy

我们首先定义 Joint Entropy 的概念：

##### Definition 16. Joint Entropy

对于两个随机变量  $X$  和  $Y$ ，其联合概率分布为  $p(x_i, y_j)$ ，我们定义 Joint Entropy 为：

$$H(X, Y) = - \sum_{i,j} p(x_i, y_j) \log p(x_i, y_j). \quad (204)$$

我们知道如果这两个随机变量完全独立，那么  $H(X, Y) = H(X) + H(Y)$ 。否则  $H(X, Y) \leq H(X) + H(Y)$ 。这个 Entropy 的 interpretation 其实就是把两个随机变量看作一个联合的随机变量，然后计算这个联合随机变量的 Entropy。

#### Conditional Entropy

在 Joint Entropy 的基础上我们可以定义 Conditional Entropy 的概念：

##### Definition 17. Conditional Entropy

对于两个随机变量  $X$  和  $Y$ ，其联合概率分布为  $p(x_i, y_j)$ ，我们定义 Conditional Entropy 为：

$$H(X|Y) = H(X, Y) - H(Y) \quad (205)$$

这个 Entropy 的意义是：

- 在知道  $Y$  的情况下  $X$  的不确定性

因此这个意义已经说明了  $H(X|Y) \neq H(Y|X)$ 。并且满足一个不等式： $0 \leq H(X|Y) \leq H(X)$

这个 interpretation 可以通过一个等价的定义进行理解：

##### Definition 18. Conditional Entropy (equivalent definition)

对于两个随机变量  $X$  和  $Y$ ，其联合概率分布为  $p(x_i, y_j)$ ，我们定义 Conditional Entropy 为：

$$H(X|Y) = \sum_j p(y_j) H(X|Y = y_j) = - \sum_j p(y_j) \sum_i p(x_i|y_j) \log p(x_i|y_j) \quad (206)$$

其中  $p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)}$  被 interpret 为在知道  $Y = y_j$  的情况下  $X$  的条件概率分布。

#### Mutual Information

我们可以定义 Mutual Information 的概念：

**Definition 19. Mutual Information**

对于两个随机变量  $X$  和  $Y$ ，其联合概率分布为  $p(x_i, y_j)$ ，我们定义 *Mutual Information* 为：

$$I(X : Y) = H(X) + H(Y) - H(X, Y) \quad (207)$$

这个 Entropy 的意义是：

- 这相当于知道  $Y$  之前的  $X$  的位置的信息量  $H(X)$  减去知道  $Y$  之后  $X$  的位置的信息量  $H(X|Y)$ ，也就是  $I(X : Y) = H(X) - H(X|Y)$
- 可以被理解为：知道  $Y$  之后就知道了多少关于  $X$  的信息；由于其对称性也可以被理解为：知道  $X$  之后就知道了多少关于  $Y$  的信息；也就是两者之间的共同 share 信息量

特别的我们会发现，如果  $X$  和  $Y$  独立，那么  $I(X : Y) = 0$ 。而如果  $X = Y$ ，那么  $I(X : Y) = H(X) = H(Y)$ 。我们可以证明其满足不等式：

$$0 \leq I(X : Y) \leq \min\{H(X), H(Y)\} \quad (208)$$

上面几个 entropy 如果被 interpret 为「未知的信息量」的话，我们可以使用 venn 图表示如下：

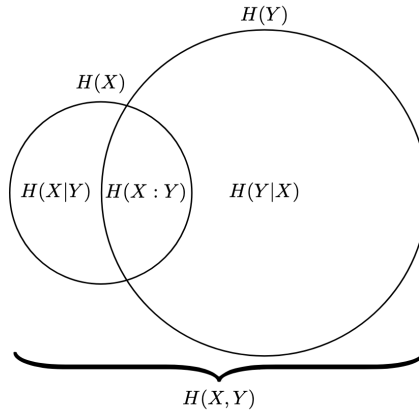


图 1: Classical Entropy Venn Diagram

{fig:venn

**Relative Entropy**

我们可以定义 Relative Entropy 的概念：

**Definition 20. Relative Entropy**

我们考虑对于一个随机变量  $X$ ，其概率分布为  $p(x_i)$ ，我们考虑另一个概率分布  $q(x_i)$ ，我们定义 *Relative Entropy* 为：

$$D(p||q) = \sum_i p(x_i) \log \frac{p(x_i)}{q(x_i)} = -H(X) - \sum_i p(x_i) \log q(x_i) \quad (209)$$

这个 Entropy 的意义是：

- 表示两个概率分布之间的相似性！

我们可以证明其满足不等式：

$$D(p||q) \geq 0 \quad \text{with equality iff } p = q \quad (210)$$

#### Shannon Entropy的另一种interpretation

我们研究 Relative Entropy，不妨选择  $q(x_i) = \frac{1}{d}$  为均匀分布，那么我们发现：

$$D(p||q) = \sum_i p(x_i) \log p(x_i) - \sum_i p(x_i) \log \frac{1}{d} \quad (211)$$

$$= -H(X) + \log d \quad (212)$$

$$\Rightarrow H(X) = \log d - D(p||q) \quad (213)$$

也就是说，Shannon Entropy 表达了一个概率分布和均匀分布之间的距离！越接近均匀分布 Entropy 越大，越远离均匀分布 Entropy 越小！

#### Remark:

对于经典信息论里面所有的  $\log$  我们都取以 2 为底。

## 9.2 Quantum Entropy

### 9.2.1 Single Quantum System Entropy

#### Von Neumann Entropy

下面我们考虑把经典信息论的内容推广到量子信息论之中。我们可以定义 Von Neumann Entropy 的概念：

#### Definition 21. Von Neumann Entropy

对于一个量子态  $\rho$ ，我们定义 Von Neumann Entropy 为：

$$S(\rho) = -\text{Tr}(\rho \log \rho). \quad (214)$$

这里的对数我们一般取以 2 为底！如果我们可以进行谱分解  $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$ ，那么我们有：

$$S(\rho) = -\sum_i \lambda_i \log \lambda_i. \quad (215)$$

一个 interpretation 就是，这个 entropy 可以描述一个量子系统存放  $N$  个量子态所需要最少的 Hilbert Space 的维度。如果我们希望用一个量子系统来存放  $\rho^{\otimes N}$  量子态，那么这个量子系统的 Hilbert Space 的维度至少需要为  $2^{NS(\rho)}$ 。

#### Properties of Von Neumann Entropy

我们可以证明 Von Neumann Entropy 满足下面的性质：



- 对于 pure state 我们有  $S(\rho) = 0$ ; 对于 maximally mixed state 我们有  $S(\rho) = \log d$ , 其中  $d$  为 Hilbert Space 的维度。这是一个量子系统 entropy 的极限情况。
- 如果我们的 State 发生了 Unitary 的演化  $\rho \rightarrow U\rho U^\dagger$ , 那么 Entropy 不变。也就是说  $S(U\rho U^\dagger) = S(\rho)$

### Von Neumann Entropy与Measurement

我们考虑对于一个物理量的  $M = \sum_i m_i |i\rangle\langle i|$  的测量。这个测量可以为我们定义另一个随机变量  $M$ , 其概率分布为  $p_i = \langle i|\rho|i\rangle$  给出  $m_i$  的取值。那么我们可以定义这个随机变量的 Shannon Entropy 为:

$$H(M) = -\sum_i p_i \log p_i. \quad (216)$$

我们可以发现下面的结论:

$$S(\rho) = -\sum_i \lambda_i \log \lambda_i \leq H(M) \quad (217)$$

这个结论有两种说明的方法:

- 使用 Eigen Basis 之外的任意 basis 然后 decoherence(Killing the off-diagonal), 则 entropy 增加; 也可以理解为让这个 state 更加的 mixed
- Measuring Uncertainty 在 Eigen Basis 是最小的

### Properties of Von Neumann Entropy II

我们研究对于多个系统以及组合系统的 entropy 性质:

- 对于两个系统  $A$  和  $B$ , 我们有  $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$
- Von Neumann Entropy 满足三角不等式:  $S(\rho_A) + S(\rho_B) \geq S(\rho_{AB}) \geq |S(\rho_A) - S(\rho_B)|$
- 对于一个系统可以写作很多 state 的 convex combination, 比如  $\rho = \sum_i p_i \rho_i$ , 那么我们有  $S(\rho) \geq \sum_i p_i S(\rho_i)$ 。也就是说 entropy 是一个 concave 函数。或者说: convex mixing 会增加 entropy。

## 9.2.2 2 Quantum Systems Entropy

### Joint Quantum Entropy

很自然我们对于一个组合系统  $AB$  定义 Joint Quantum Entropy, 也就是这个系统的 Von Neumann Entropy:

$$S(\rho_{AB}) = -\text{Tr}(\rho_{AB} \log \rho_{AB}). \quad (218)$$

### Conditional Quantum Entropy and Mutual Quantum Information

我们可以定义 Conditional Quantum Entropy 的概念:

**Definition 22. Conditional Quantum Entropy**

对于一个组合系统  $AB$ ，我们定义 *Conditional Quantum Entropy* 为：

$$S(A|B) = S(\rho_{AB}) - S(\rho_B). \quad (219)$$

我们可以定义 *Mutual Quantum Information* 的概念：

**Definition 23. Mutual Quantum Information**

对于一个组合系统  $AB$ ，我们定义 *Mutual Quantum Information* 为：

$$I(A : B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \quad (220)$$

我们考虑一个系统， $AB$  处于一个 pure state 是一个 bell state  $|\phi^+\rangle$  我们知道纯态的 entropy 为 0，也就是说  $S(\rho_{AB}) = 0$ 。但是我们知道每一个子系统的 entropy 为 1，也就是说  $S(\rho_A) = S(\rho_B) = 1$ 。因此我们发现： $S(A|B) = -1$ ，也就是说在知道  $B$  的情况下  $A$  的 entropy 变成了-1！这说明量子信息论 conditional entropy 不一定需要满足大于 0 的条件！这和经典信息论有很大的不同！

**Relative Quantum Entropy****Definition 24. Relative Quantum Entropy**

对于两个量子态  $\rho$  和  $\sigma$ ，我们定义 *Relative Quantum Entropy* 为：

$$S(\rho||\sigma) = \text{Tr}(\rho \log \rho) - \text{Tr}(\rho \log \sigma). \quad (221)$$

如果我们使用对角化的 basis 进行  $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$  和  $\sigma = \sum_j \mu_j |\phi_j\rangle\langle\phi_j|$  的谱分解，那么我们有：

$$S(\rho||\sigma) = \sum_i \lambda_i \log \lambda_i - \sum_{i,j} \langle\psi_i|\phi_j\rangle\langle\phi_j|\psi_i\rangle \lambda_i \log \mu_j, \quad (222)$$

我们可以证明其满足下面的性质：

- $S(\rho||\sigma) \geq 0$ ，并且当且仅当  $\rho = \sigma$  时取等号
- $S(\rho||\sigma) \neq S(\sigma||\rho)$ ，也就是说这个量不是对称的
- Unitary Invariance：对于任意的 unitary operator  $U$ ，我们有  $S(U\rho U^\dagger||U\sigma U^\dagger) = S(\rho||\sigma)$

**Data Processing Inequality**

对于 Relative Entropy 我们可以证明其在一个量子信道  $\mathcal{E}$  下满足 Data Processing Inequality：

$$S(\mathcal{E}(\rho)||\mathcal{E}(\sigma)) \leq S(\rho||\sigma). \quad (223)$$

我们回顾 Matrix Norm 的概念，发现 1-norm 也同样满足这个关系！！

### 9.3 Fidelity

#### Classical Fidelity

我们对于同一个随机变量  $X$ ，其概率分布分别为  $p(x_i)$  和  $q(x_i)$ ，我们定义 Classical Fidelity 为：

#### Definition 25. Classical Fidelity

对于同一个随机变量  $X$ ，其概率分布分别为  $p(x_i)$  和  $q(x_i)$ ，我们定义 Classical Fidelity 为：

$$F(p, q) = \left( \sum_i \sqrt{p(x_i)q(x_i)} \right). \quad (224)$$

#### Properties of Classical Fidelity

我们可以证明 Classical Fidelity 满足下面的性质：

- $0 \leq F(p, q) \leq 1$ ，并且当且仅当  $p = q$  时取  $= 1$ 。
- $F(p, q) = 0$  当且仅当完全没有重叠也就是  $\forall i, p(x_i)q(x_i) = 0$

注意，Fidelity 不能够描述两个经典 vector 之间的距离，因为其不满足三角不等式。

#### Quantum Fidelity

对于两个量子态  $\rho$  和  $\sigma$ ，我们定义 Quantum Fidelity 为：

#### Definition 26. Quantum Fidelity

对于两个量子态  $\rho$  和  $\sigma$ ，我们定义 Quantum Fidelity 为：

$$F(\rho, \sigma) = \min_{\{M_i\}} \sum_i \sqrt{\text{Tr}(M_i \rho) \text{Tr}(\sigma M_i)}, \quad (225)$$

其中  $\{M_i\}$  是一组 POVM。也就是说我们选择一组测量让 Quantum State 给出一组概率分布，然后选择那个让这个概率分布的 Classical Fidelity 最小的测量给出的数值。

对于 Quantum Fidelity 我们存在另一个等价形式：

#### Theorem 29. 对于两个量子态 $\rho$ 和 $\sigma$ ，我们 Quantum Fidelity 等价于：

$$F(\rho, \sigma) = \text{Tr} \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}}. \quad (226)$$

#### Properties of Quantum Fidelity

我们考虑两个 state  $\rho$  和  $\sigma$ ，假设他们可以在同一组基上对角化， $\rho = \sum_i r_i |i\rangle\langle i|$ ， $\sigma = \sum_i s_i |i\rangle\langle i|$ 。那么我们有：

$$F(\rho, \sigma) = \sum_i \sqrt{r_i s_i}. \quad (227)$$

也就退化为这一组基上面的 classical fidelity。这个可以直接通过等价形式证明！并且我们注意  $i = |i\rangle\langle i|$  是一个 projective operator 所以  $i^2 = i = \sqrt{i}$ 。

### Pure State Fidelity

如果  $\rho$  是一个 pure state  $\rho = |\psi\rangle\langle\psi|$ ，而  $\sigma$  是任意的一个量子态，那么我们有：

$$F(\rho, \sigma) = \sqrt{\langle\psi|\sigma|\psi\rangle}. \quad (228)$$

这个可以直接通过等价形式证明！并且我们注意  $\rho^2 = \rho = \sqrt{\rho}$ 。

而如果  $\sigma$  也是一个 pure state  $\sigma = |\phi\rangle\langle\phi|$ ，那么我们有：

$$F(\rho, \sigma) = |\langle\psi|\phi\rangle|. \quad (229)$$

也就是说 fidelity 描述了两个 pure state 之间的 overlap。

### Uhlmann's theorem

我们知道 pure state 的 fidelity 可以很好的计算并且有实际意义，因此我们希望考虑 mix state 纯化之后的 fidelity 和 mix state 的 fidelity 之间的关系。我们有下面的定理：

#### Theorem 30. Uhlmann's theorem

对于两个量子态  $\rho$  和  $\sigma$ ，我们定义他们在同一个更大的 Hilbert Space 上的 purification 分别为  $|\psi_\rho\rangle$  和  $|\phi_\sigma\rangle$ ，那么我们有：

$$F(\rho, \sigma) = \max_{|\psi_\rho\rangle, |\phi_\sigma\rangle} |\langle\psi_\rho|\phi_\sigma\rangle|, \quad (230)$$

也就是说 mix state 的 fidelity 等于所有纯化之后的 pure state fidelity 的最大值。

### Data Processing Inequality of Fidelity

我们可以证明 Fidelity 在量子信道  $\mathcal{E}$  下满足 Data Processing Inequality：

$$F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma). \quad (231)$$

也就是说我们不能通过量子信道让两个量子态变得更不相似！量子信道只会让两个态更加相似。

## 10 Lecture 11: Resource Theory and Entanglement Resource

{sec:Lect

### 10.1 Set Up for Resource Theory

### 10.2 Resource Theory of Entanglement

## 11 Lecture 12: Entanglement Measure

{sec:Lect

### 11.1 Entanglement Measure

### 11.2 Mixed State Entanglement Measure

## 12 Sup 1: QI 之中常用的数学技巧

这里我们汇总一下做题之中涉及的数学技巧以及容易混淆的常用结论捏

### 12.1 基础数学技巧

#### 12.1.1 张量积

##### 张量积和矩阵加法

张量积和矩阵加法可以互换，满足关系：

$$\sum_i (A_i \otimes B) = \left( \sum_i A_i \right) \otimes B. \quad (232)$$

##### 张量积和矩阵乘法

张量积和矩阵乘法是可以互换的，矩阵乘法可以找到自己对应的张量积子空间对应乘起来就好：

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD) \quad (233)$$

##### 张量积和Vectorization

##### 张量积和Tr操作

我们知道张量积和 Tr 是可以互换的，并且张量积的 Tr 其实就是原本 Tr 的乘积：

$$\text{Tr}(A \otimes B) = (\text{Tr}A)(\text{Tr}B) \quad (234)$$

#### 12.1.2 Unitary

##### Unitary变换矩阵

对于任意算符  $A$ ，作酉共轭变换

$$A' = UAU^\dagger, \quad (235)$$

则  $A'$  与  $A$  具有完全相同的本征值（包括重数）。即

$$\text{Spec}(A') = \text{Spec}(A). \quad (236)$$

换言之，unitary 变换仅仅对应于基底变换，不改变物理可观测量的谱。

证明要点. 设  $A|\psi\rangle = \lambda|\psi\rangle$ ，则

$$A'(U|\psi\rangle) = UAU^\dagger U|\psi\rangle = UA|\psi\rangle = \lambda(U|\psi\rangle), \quad (237)$$

说明  $U|\psi\rangle$  是  $A'$  对应本征值  $\lambda$  的本征态，从而本征值保持不变。□

因此，unitary 变换可视为“旋转基底”，保留体系的谱结构：

$$\text{Tr}(A') = \text{Tr}(A), \quad \det(A') = \det(A), \quad \|A'\|_1 = \|A\|_1. \quad (238)$$

### Unitary和奇异值

同样的对于一般的矩阵，Unitary 变换也不改变矩阵的奇异值。也就是说：

$$\sigma_i(UAV^\dagger) = \sigma_i(A) \quad (239)$$

## 12.1.3 Trace

### Trace的一个基本trick

对于一个量子态  $\rho$  写作：

$$\rho = \sum_{a,b} f(a,b) |a\rangle\langle b| \quad (240)$$

我们求  $\text{Tr}$  等价于形式化的把 ket 放在 bra 的右边：

$$\text{Tr}(\rho) = \sum_{a,b} f(a,b) \text{Tr}(|a\rangle\langle b|) = \sum_{a,b} f(a,b) \langle b|a\rangle \quad (241)$$

### Trace和Unitary演化

如果我们有 A,B 两个系统，然后在其上有一个  $\rho_{AB}$ 。如果在 B 系统上面进行 Unitary 演化以及 A 系统上面进行任意演化，我们发现这两个过程是可以互换的。所以特殊的对于 Partial Trace：

$$\text{Tr}_A((I_A \otimes U_B)\rho_{AB}(I_A \otimes U_B^\dagger)) = U_B \text{Tr}_A(\rho_{AB}) U_B^\dagger \quad (242)$$

### SWAP operator 和 Trace

### Trace of times

两个密度矩阵的乘积的 Trace 满足不等式：

$$|\text{Tr}(\rho\sigma)| \leq \sqrt{\text{Tr}(\rho^2) \text{Tr}(\sigma^2)} \quad (243)$$

$$\text{Tr}(\rho^2) \leq 1, \quad \text{Tr}(\sigma^2) \leq 1 \Rightarrow \text{Tr}(\rho\sigma) \leq 1 \quad (244)$$

### 单bit measure的TR

对于两个单比特态  $\rho_1, \rho_2$ ，我们有：

## 12.2 Pauli 矩阵的特殊性质

由于 Pauli 矩阵在计算之中太过于常见，我们下面列出计算之中常用的 Pauli 矩阵性质



### 12.2.1 基本定义

### 12.2.2 基本乘法性质

### 12.2.3 对易反对易关系

### 12.2.4 互相作用，pauli 矩阵作用在 pauli 基上

### 12.2.5 指数形式的 Pauli 矩阵

### 12.2.6 Pauli 矩阵的本征值与本征基

单比特的 Pauli 矩阵定义为

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

其中  $I$  是单位算符，而  $X, Y, Z$  均为厄米且酉的算符，满足  $P^2 = I$ 。因此，对于任意 Pauli 矩阵  $P \in \{X, Y, Z\}$  有

$$P^2 = I \quad \Rightarrow \quad \lambda^2 = 1,$$

即其本征值为

$$\boxed{\lambda = \pm 1}.$$

更具体地，本征值与本征态如下：

$$X: \quad X |\pm\rangle = \pm |\pm\rangle, \quad |\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle),$$

$$Y: \quad Y |\pm i\rangle = \pm |\pm i\rangle, \quad |\pm i\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm i |1\rangle),$$

$$Z: \quad Z |0\rangle = +1 |0\rangle, \quad Z |1\rangle = -1 |1\rangle.$$

**Pauli 算符期望值的取值范围** 对任意量子态  $\rho$ ，Pauli 算符  $P$  的期望值定义为

$$\langle P \rangle_\rho = \text{Tr}(\rho P).$$

由于 Pauli 矩阵的本征值仅为  $\pm 1$ ，其测量结果可以看作一个取值为  $\pm 1$  的随机变量，因此期望值必定满足

$$-1 \leq \langle P \rangle_\rho \leq 1.$$

更严格地，利用谱分解  $P = \sum_i \lambda_i |i\rangle \langle i|$ ，并注意  $\rho$  为正定且  $\text{Tr} \rho = 1$ ，可得

$$\langle P \rangle_\rho = \sum_i \lambda_i p_i, \quad p_i \geq 0, \quad \sum_i p_i = 1,$$

即为  $\pm 1$  的加权平均，从而有

$$\boxed{-1 \leq \langle P \rangle_\rho \leq 1}.$$

### 12.3 常用不等式

平方和和和的平方 对于任意实数  $a_i$ ，我们有下面不等式：

$$M \sum_{i=1}^M a_i^2 \geq \left( \sum_{i=1}^M a_i \right)^2 \quad (245)$$

## 13 Scratch book

{sec:Scra

这里会放一些写的很混沌，但懒得扔掉的东西呜呜呜呜!!!