109 量子資訊與計算

Quantum Information and Computation

Hao-Chung Cheng (鄭皓中)
haochung@ntu.edu.tw

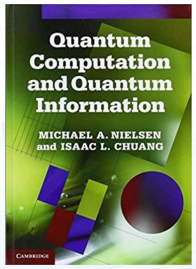
Department of Electrical Engineering
National Taiwan University

February 24, 2021

Course Information

Course Information

- **Instructor**
  - Name 鄭皓中
  - E-mail haochung@ntu.edu.tw
  - Office Hour 12:10~13:10 on Thursdays, 電二549室 (booking preferred)
- **Time** 9:10~12:10 on Wednesday (三2, 3, 4)
- **Location** 明達231室
- **Course Materials** Lecture based on slides, notes, and reference books (see the references listed in the Syllabus)



Course Information

- **Teaching Assistant**
  - Name 廖允執
  - E-mail r05222001@g.ntu.edu.tw
  - Accessibility 電二533室→ 學新513室 (by email)

- **Teaching Assistant**
  - Name 洪晟霖
  - E-mail chenglin@ntu.edu.tw
  - Accessibility 電二533室→ 學新513室 (by email)

Prerequisites

- This theory course heavily relies on Linear Algebra
  → Please check Homework 0 before committing your time to this course
- Basic knowledge of Probability and Statistics
- Basic knowledge of Computation Theory and Algorithm [Chapter 3, N&C]
- No requirement on Quantum Mechanics

Goal of this course

Theory of Quantum Information Processing

- Theory: mathematical foundation of quantum information

（注意：這是一門理論課！）

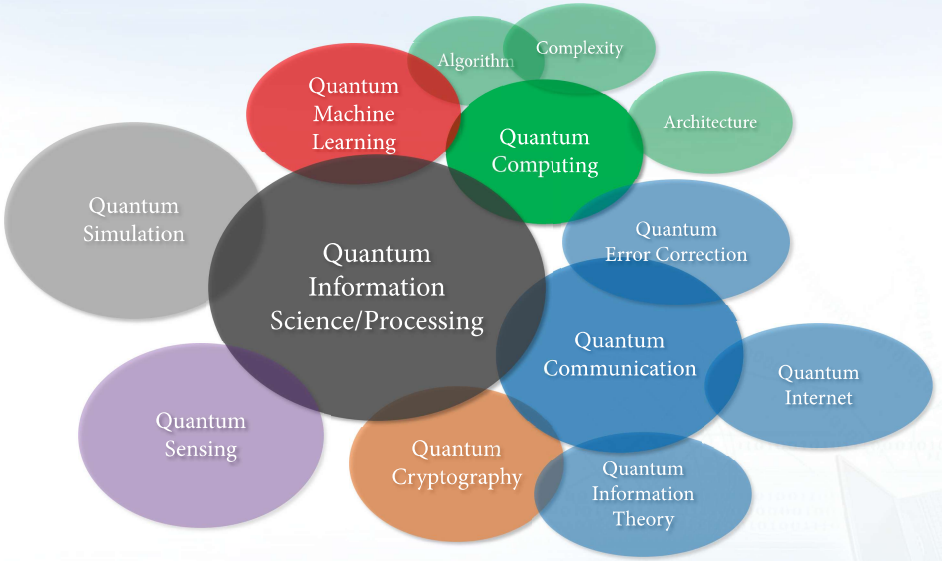
- Quantum Information Processing:
  - Computation (circuit & algorithms)
  - Information Theory
  - Communication
  - Cryptograph etc.

Grading Policy and Requirements

HW 0	HW 1	HW 2	HW 3	HW 4	Mid-term Exam	Final Project	**Total**
0 %	10 %	10 %	10 %	10 %	30 %	30 %	100%

- Homework is done individually
  - Online submission (LaTex preferred, or clearly photo-copied)
- Mid-term exam will be taken at 4/14 or 4/21
- Final Project
  - Final Report of a surveyed topic
  - Slides of the oral presentation (depends on the number of students, tentatively at 6/9 & 6/15)
    → For the students to become familiar with some of research topics

Underlined text: what you need to submit to the course website



Why Quantum Information Processing?

- See my talk: "A Glimpse of The Quantum Information Industry"

(Joint NTU GICE & GIEE Colloquium, October 26, 2020)

https://youtu.be/Rg5AhfT0o0k

Course Syllabus

https://cutt.ly/all7D0R

Course Overview

- First Part: Quantum computation [Chapters 1~6, N&C]
  - Foundations of Quantum Information Science (Postulates)
  - Closed quantum systems Isolated
  - Quantum circuits & algorithms

- Second Part: Quantum information and communication [Chapters 8~12, N&C]
  - Open quantum systems (QIP in the presence of noise)
  - Quantum operations, distance measures, entropies
  - Quantum information theory – compression & communication
  - Quantum error correction, basic quantum cryptography (if time permited)

What I hope for

- You know what really make senses by

  Quantum + [ ] ?

- You know the difference/separation between classical and quantum stuff

- You are able to follow ArXiv papers: https://arxiv.org/list/quant-ph/ and independently do quantum research

Quantum Bubble? Hype?

> "There is still a lot of value being created — it's just a case of whether there is too much hype," - Christian Weedbrook, founder of the quantum-computing firm Xanadu in Toronto, Canada.

- As a researcher/visionary, to innovate killer applications
- As a scientist, to prove no-go theorems/impossibility
- As an engineer, to find a way around the hurdles

Course Policy

- Homework, Reports should be uploaded to the NTU COOL by the deadline
  - Format: PDF (LaTeX preferred)
- No late submission!
- No plagiarism. You are encouraged to discuss homework problems with your peers, with the TAs, and with the course instructor (if time permits). However, your solutions should be based on your own understanding and written *independently* in your own words.
- Please acknowledge all sources of help on your assignments – failing to do so constitutes an academic offense.
  - Caught academic offense for the first time: final grade –10
  - Caught academic offense more than once: Failed

Sample Topics (Not limited to, will be updated)

- Fault-tolerant quantum computation
- Topological quantum computation
- Advances in quantum key distribution
- Advances in quantum communication
- Hardware for quantum computers
- Quantum machine learning
- Quantum metrologies/sensing
- Quantum simulation
- Quantum control
- Quantum finance
- Quantum computing architecture
- Advances in quantum cryptography
- Advances in quantum surface codes
- Advances in quantum information theory
- Advances in quantum state discrimination
- On quantum supremacy
- Quantum networks
- Quantum complexity theory
- Quantum circuit synthesis
- Advances in quantum algorithms
- QIP talks
- ArXiv papers: https://arxiv.org/list/quant-ph/

Presentation/Final Project Guidelines

- A professional and academic presentation (9 & 16 June, tentative)
- What to cover:
  - Necessary (but minimum) preliminaries
  - Problem statement
  - Existing results
  - Methodology and novelty
  - Implications and contributions
  - Possible future work

Why your presentation is worth time for other audience?

- The topic is chosen by 5 May (tentative), confirmed with TA via email
- Final report – a professional and academic paper (LaTeX)
  - Cover the above mentioned points
  - Due by 23 June, 2021
  - No plagiarism!

What to do after the course today?

A Taste of Quantum Information

- Reading assignment: HW0 and Chapter 1, 2 of N&C
- Public lecture by Charles Bennett (2020 Claude E. Shannon Award recipient) at QisKit

https://youtu.be/rslt-LwtDK4

Quantum Information and Computation
The Quantum Postulates

Hao-Chung Cheng (鄭皓中)
haochung@ntu.edu.tw

Department of Electrical Engineering
National Taiwan University

February 24, 2021

Outline

1. Prologue – What is information?
2. Mathematical Framework
   - Postulate 1 (State)
   - Postulate 2 (System)
   - Postulate 3 (Evolution)
   - Postulate 4 (Measurement)
3. Epilogue – Concluding Remarks

Prologue – What is information?

What is information?

- Intuitively, we "obtain information" when we acquire knowledge of an alternative that we did not know before (and consequently there is a "reduction in uncertainty").
- In classical information processing, information is conventionally represented by a classical bit (or string of bits), i.e. a Boolean variable that can take values 0 or 1. It represents the "elementary unit of information" giving the result of a single binary decision, e.g. a yes/no question.
- Then, what is bit?
- R. Landauer 1996: "Information is not a disembodied abstract entity, it is always tied to a physical representation." → "No information is without representation."
- A bit is given by any two different physical states (of some physical system) that can be perfectly distinguished (by a physical measurement apparatus).

Information is physical!

- If information is represented in physical states or degrees of freedom of some physical system, then any possible act of information processing (e.g. computation or communication) must corresponds to a physical evolution and cannot be determined by abstract thought or mathematics along.
- Any actual computer is always a physical device whose operation must obey the law of physics.


PHYSICS TODAY

HOME BROWSE INFO RESOURCES JOBS

Home > Physics Today > Volume 44, Issue 5 > 10.1063/1.881299

01 MAY 1991 • page 23

Information is Physical

There are no unavoidable energy consumption requirements per step in a computer. Related analysis has provided insights into the measurement process and the communication schannel, and has prompted speculations about the nature of physical laws.

Rolf Landauer

Thomas J. Watson Research Center, Yorktown Heights, New York

Mathematical Framework of Quantum Theory

Mathematical Framework

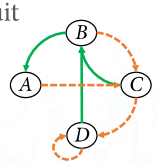
- It is the goal of quantum theory – just as of every other physical theory – to predict the results of experiments and to justify these predictions.
- Quantum theory, in its conventional formulation, is built on the theory of Hilbert spaces and operators.

Physical quantities	Mathematical characterizations
Quantum system	Hilbert space $\mathcal{H}$
Quantum state	Vectors in (or more generally density operators on) $\mathcal{H}$
Evolution of the quantum state	Unitary operators or CPTP maps
Predictions	Probabilistic statements (POVM)

Postulate 1 (State)

Deterministic Classical System and State

- A finite classical system can be described by an *alphabet*, i.e. a finite and non-empty set, whose elements may be considered to be *symbols*.
  - $\mathcal{X} = \{\boxdot, \boxdot, \boxdot, \boxdot, \boxdot, \boxdot\}$ for a dice 🎲; $\mathcal{X} = \{\spadesuit\ \clubsuit\ \heartsuit\ \diamondsuit\}$ for a card suit
  - $\mathcal{X} = \{0, 1, 2, \ldots, 9\}$ for a digit
  - A finite-state machine



- A *analog* system maybe described by a continuous set
  e.g. the thermometer 🌡 ranges [35 °C, 42 °C]
- A classical bit: finite field $\mathbb{Z}_2 = \{0, 1\}$
  → The vector space $\mathbb{Z}_2$ over field $\mathbb{Z}_2$ has dimension 1
- A column vector: $\begin{pmatrix}\text{height}\\ \text{weight}\end{pmatrix} \in \mathbb{R}^2$ for a person

Closed Quantum Systems and States

Postulate 1 (State)

- To any physical (quantum-mechanical) system, there is associated a Hilbert space, i.e. a complex inner product space.
- The (pure) physical state of the isolated system is completely described by its state vector, which is a unit vector in the Hilbert space.

- A d-level quantum system is described by a d-dimensional Hilbert space.
- The vectors are denoted with the Dirac notation, e.g. $|\psi\rangle \in \mathcal{H}$.

The Dirac Notation (1/3)

- A Hilbert space $\mathcal{H}$ is a vector space over the field $\mathbb{C}$ of complex numbers endowed with the (Euclidean) *inner product*.
  - A vector in $\mathcal{H}$ is denoted by the ket-vector $|\psi\rangle$.
  - If $\mathcal{H} \simeq \mathbb{C}^d$ (i.e. a d-dimensional complex vector space), then any $|\psi\rangle \in \mathcal{H}$ is given by a column vector of length d and complex entries.
  - Its conjugate transpose is denoted by the bra-vector: $\langle\psi| := |\psi\rangle^\dagger$, which is a row vector.
  - The inner product (or scalar product) of two vectors $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ is denoted as $\langle\psi|\phi\rangle$.
- Example: A 2-dimensional Hilbert space $\mathcal{H} = \mathbb{C}^2$ with a chosen orthonormal basis denoted $\{|0\rangle, |1\rangle\}$, i.e. the basis vectors are labelled by the bit values 0 and 1. This is called the *computational basis* or *standard basis*.

$$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}, |1\rangle = \begin{pmatrix}0\\1\end{pmatrix} \Rightarrow |\psi\rangle = a|0\rangle + b|1\rangle = \begin{pmatrix}a\\b\end{pmatrix} \in \mathbb{C}^2,\ |a|^2 + |b|^2 = 1$$

The Dirac Notation (2/3)

- Outer product: The outer product of two vectors $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ is given by $|\psi\rangle\langle\phi|$. If $\mathcal{H} = \mathbb{C}^d$, then $|\psi\rangle\langle\phi|$ is given by an $d \times d$ matrix.

- Let $\{|i\rangle\}_{i=1}^d$ be an orthonormal basis (ONB) of $\mathcal{H}$.
  - $\langle i|j\rangle = \delta_{ij}$ (orthonormality)
  - $\sum_{i=1}^d |i\rangle\langle i| = I$ (completeness relation)
  - Any operator A on $\mathcal{H}$ can be written as $A = \sum_{ij} |i\rangle\langle i|A|j\rangle\langle j| = \sum_{ij} [A]_{ij} |i\rangle\langle j|$
  - Any vector $|\psi\rangle \in \mathcal{H}$ can be expressed in this basis as: $|\psi\rangle = \sum_{i=1}^d a_i |i\rangle$, $a_i \in \mathbb{C}$.
    $\because \langle\psi|\psi\rangle = 1, \therefore \sum_{i=1}^d |a_i|^2 = 1. \rightarrow \{|a_i|^2\}_{i=1}^d$ forms a probability distribution.

Probability amplitude

The Dirac Notation (3/3)

$$|\psi\rangle = \begin{pmatrix}\psi_1\\ \vdots \\ \psi_d\end{pmatrix}, \qquad \langle\psi| = \begin{pmatrix}\overline{\psi_1} & \cdots & \overline{\psi_d}\end{pmatrix},$$

$$\langle\phi|\psi\rangle = \sum_{i=1}^d \overline{\phi_i}\psi_i, \qquad |\psi\rangle\langle\phi| = \begin{pmatrix}\psi_1\overline{\phi_1} & \cdots & \psi_1\overline{\phi_d} \\ \vdots & \ddots & \vdots \\ \psi_d\overline{\phi_1} & \cdots & \psi_d\overline{\phi_d}\end{pmatrix}.$$

An Implication of Postulate 1

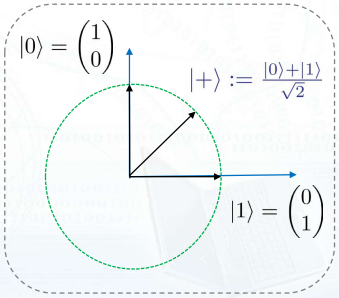
- The Superposition Principle:
  If $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, then any state which is a superposition of these states:

$$|\Psi\rangle = \frac{a|\psi\rangle + b|\phi\rangle}{\||a|\psi\rangle + b|\phi\rangle\|}, \qquad a, b \in \mathbb{C}$$

  is also a (pure) state.

- The quantum bit (qubit): $\mathcal{H} = \mathbb{C}^2$
  We say that $|\psi\rangle$ is a *superposition* of states $|0\rangle$ and $|1\rangle$.

$$|\psi\rangle = a|0\rangle + b|1\rangle,\ a, b \in \mathbb{C},\ |a|^2 + |b|^2 = 1$$



Global and Relative Phase of State Vectors

- Consider two states $|\psi\rangle$ and $|\phi\rangle$ where $|\phi\rangle = e^{i\theta}|\psi\rangle,\ \theta \in \mathbb{R}$.
  → These two states differ by the factor $e^{i\theta}$, of unit modulus, which is referred to as a global phase factor.
- These states describe *the same physical state* of a system.
  This is because there are no measurements which can be used to distinguish between such states.
- Hence, the state of a physical system is given by a *ray* in a Hilbert space, the latter being an equivalence class of unit vectors that differ by a global phase.
  → By a pure state $|\psi\rangle$, we meant $|\psi\rangle \equiv \{e^{i\theta}|\psi\rangle : \theta \in \mathbb{R}\}$.
- However, the relative phase between two states is of physical significance, i.e. $a|\psi\rangle + b|\phi\rangle$ and $a|\psi\rangle + be^{i\theta}|\phi\rangle$ do not represent the same physical state.

The Qubit Systems [Chapter 7, N&C]

- Physical implementations of qubit systems are:
  - 2-level atoms (also atoms with more levels, when only two level play a role ina particular process), and ions with two energy levels;
  - Polarisation of spin-$\frac{1}{2}$ particles;
  - Polarisation of single photons (horizontal/vertical or left/right-hand circular)
  - Ray paths in a two-path interferometer containing exactly one photon;
  - Quantum dots;
  - Modes of the electromagnetic field in a cavity resonator;
  - The two-slit experiment
  - The Stern–Gerlach experiment
  - etc.

Open Quantum Systems and States

Postulate 1* (State)

- To any physical (quantum-mechanical) system, there is associated a Hilbert space, i.e. a complex inner product space.
- The system is completely described by its *density operator*, which is positive semi-definite operator with unit trace acting on the Hilbert space.

背後有個機率分佈

- Density operator/matrix or *mixed state*: $\rho \geq 0, \mathrm{Tr}[\rho] = 1$.
- Pure state $|\psi\rangle$ is just a rank-1 operator $|\psi\rangle\langle\psi|$.

The Bloch-Sphere Representation – Visualizing Qubits

- A density operator on $\mathcal{H} = \mathbb{C}^d$ is a $d \times d$ self-adjoint matrix $\rho \in \mathcal{B}_{\mathrm{sa}}(\mathcal{H})$.
  → Characterizing the linear space $\mathcal{B}_{\mathrm{sa}}(\mathcal{H})$ with the $\langle A, B\rangle \coloneqq \mathrm{Tr}[A^\dagger B]$.

Hilbert–Schmidt inner product

- The dimension of the real vector space of $\mathcal{B}_{\mathrm{sa}}(\mathcal{H})$ is $d + 2 \cdot d(d-1)/2 = d^2$.
  → Choose a self-adjoint orthogonal basis $\{E_0 = I, E_1, \ldots, E_{d^2-1}\}$.
  → $\langle E_i, E_j\rangle = d\delta_{ij}$; → $E_1, \ldots, E_{d^2-1}$ are traceless.

$$\Rightarrow \rho = \frac{1}{d}\left(I + \vec{r}\cdot\vec{E}\right)$$

where $\vec{E} \coloneqq (E_1, \ldots, E_{d^2-1})$, and $\vec{r} \coloneqq (\mathrm{Tr}[\rho E_1], \ldots, \mathrm{Tr}[\rho E_{d^2-1}])$ is the *Bloch vector*.

Properties

- Length of the Bloch vector $\|\vec{r}\|$:

  Purity $\mathrm{Tr}[\rho^2] \in [1/d\,, 1]$: a density operator ρ is pure if and only if $\rho = \rho^2$.

$$\Rightarrow \mathrm{Tr}[\rho^2] = \tfrac{1}{d^2}\mathrm{Tr}\left[\left(I + \vec{r}\cdot\vec{E}\right)\left(I + \vec{r}\cdot\vec{E}\right)\right] = \tfrac{1}{d}\left(1 + \|\vec{r}\|^2\right) \le 1$$

$$\Rightarrow \|\vec{r}\| \le \sqrt{d-1}$$

Unit length for pure qubits

- Inner product: $\langle \rho_1, \rho_2 \rangle = \dfrac{1}{d}\left(1 + \vec{r}_{\rho_1}\cdot\vec{r}_{\rho_2}\right)$

→ The angle between the Bloch vectors of orthogonal pure states is $\cos^{-1}\left(\frac{1}{1-d}\right)$.

- Note: Not every operator with $\|\vec{r}\| \le \sqrt{d-1}$ is positive semi-definite.

Bloch Representation for a Qubit (1/4)

- For qubits, the operator basis $\{E_1, E_2, E_3\}$ are: Pauli matrices

Bit flip

$$X \equiv \sigma_x := |0\rangle\langle 1| + |1\rangle\langle 0| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow \sigma_x|0\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle$$

$$Y \equiv \sigma_y := -\mathrm{i}|0\rangle\langle 1| + \mathrm{i}|1\rangle\langle 0| = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}$$

Phase flip

$$Z \equiv \sigma_z := |0\rangle\langle 0| - |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

- Properties:

$$X^2 = Y^2 = Z^2 = -\mathrm{i}XYZ = I$$
$$XY = -YX = \mathrm{i}Z$$
$$YZ = -ZY = \mathrm{i}X$$
$$ZX = -XZ = \mathrm{i}Y$$

Bloch Representation for a Qubit (2/4)

- Cartesian coordinates to the *polar form* (pure state):

$$\vec{r} = (r_x, r_y, r_z) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$$

$$|\psi\rangle = \cos(\tfrac{\theta}{2})|0\rangle + \sin(\tfrac{\theta}{2})\mathrm{e}^{\mathrm{i}\varphi}|1\rangle,\ \theta \in [0,\pi],\ \varphi \in [0, 2\pi]$$

$$|\psi\rangle = a|0\rangle + b|1\rangle = r_1\mathrm{e}^{\mathrm{i}\phi_1}|0\rangle + r_2\mathrm{e}^{\mathrm{i}\phi_2}|1\rangle$$

$$= \mathrm{e}^{\mathrm{i}\phi_1}\left(r_1|0\rangle + r_2\mathrm{e}^{\mathrm{i}(\phi_2-\phi_1)}|1\rangle\right)$$

$$\simeq r_1|0\rangle + r_2\mathrm{e}^{\mathrm{i}(\phi_2-\phi_1)}|1\rangle$$

$$r_1^2 + r_2^2 = 1$$

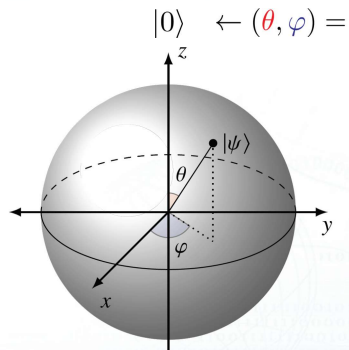
Bloch Representation for a Qubit (3/4)

$$\Rightarrow |\psi\rangle = \cos(\tfrac{\theta}{2})|0\rangle + \sin(\tfrac{\theta}{2})\mathrm{e}^{\mathrm{i}\varphi}|1\rangle,\ \theta \in [0,\pi],\ \varphi \in [0, 2\pi]$$

Classical Bit

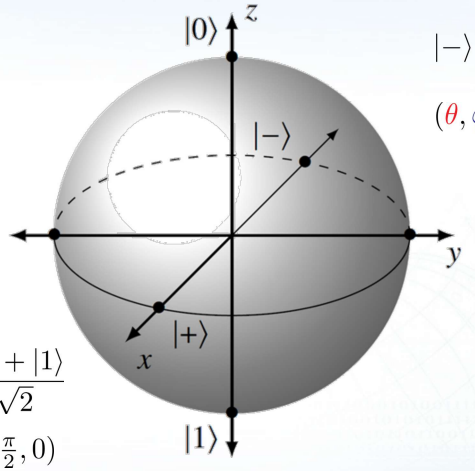
0

1



Quantum Bit

Bloch Representation for a Qubit (4/4)



$$|-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

$$(\theta, \varphi) = (\tfrac{\pi}{2}, \pi)$$

Y-Basis

$$|-\mathrm{i}\rangle = \frac{|0\rangle - \mathrm{i}|1\rangle}{\sqrt{2}}$$

$$(\theta, \varphi) = (\tfrac{\pi}{2}, \tfrac{3\pi}{2})$$

$$|\mathrm{i}\rangle = \frac{|0\rangle + \mathrm{i}|1\rangle}{\sqrt{2}}$$

$$(\theta, \varphi) = (\tfrac{\pi}{2}, \tfrac{\pi}{2})$$

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

X-Basis

$$(\theta, \varphi) = (\tfrac{\pi}{2}, 0)$$

Z-Basis or the Computation Basis

Postulate 2 (System)

Composite Classical Systems

- Composition of finite classical systems can be described by the *Cartesian product* of alphabets.
  - $\mathcal{X} \times \mathcal{X} = \{\boxdot, \boxdot, \boxdot, \boxdot, \boxdot, \boxdot\} \times \{\boxdot, \boxdot, \boxdot, \boxdot, \boxdot, \boxdot\}$ for two dices (🎲, 🎲)
  - $\mathcal{X} = \{(i, j) : i, j \in 0,1, \ldots, 9\}$ for two digits

- n classical bit: $\mathbb{Z}_2 \times \cdots \times \mathbb{Z}_2 = \mathbb{Z}_2^n$
  → The vector space $\mathbb{Z}_2^n$ over field $\mathbb{Z}_2$ has dimension n (linear in # of subsystems)

- Two column vectors: $\begin{pmatrix} \text{height} & \text{height} \\ \text{weight} & \text{weight} \end{pmatrix} \in \mathbb{R}^2 \times \mathbb{R}^2$ for two people
  → Dimension becomes $2 + 2$

Composite Quantum Systems

Postulate 2⋆ (Composition)

- For a joint system composed of two subsystems with Hilbert space $\mathcal{H}_A$ and $\mathcal{H}_B$, the Hilbert space is the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$.
- If the system i is prepared in the state ρ_i, the joint state of the total system is $\bigotimes_i \rho_i$.

Tensor Products of Spaces

- Let $\mathcal{V} \simeq \mathbb{C}^m$ and $\mathcal{W} \simeq \mathbb{C}^n$be two Hilbert spaces of dimension m and n with bases $\{|e_1\rangle, \ldots, |e_m\rangle\}$ and $\{|f_1\rangle, \ldots, |f_n\rangle\}$ respectively, then the tensor product space $\mathcal{V} \otimes \mathcal{W} \simeq \mathbb{C}^{mn}$ has dimension mn, and can be regarded as consisting of all linear combinations of $|e_i\rangle \otimes |f_j\rangle$, $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

- A natural bilinear embedding of tensor product:
  If $|\psi\rangle = \sum_i a_i |e_i\rangle \in \mathcal{V}$ and $|\phi\rangle = \sum_j b_j |f_j\rangle \in \mathcal{W}$,

$$\Rightarrow |\psi\rangle \otimes |\phi\rangle \coloneqq \sum_{ij} a_i b_j \, |e_i\rangle \otimes |f_j\rangle$$

- Note: Such mapping is *not* surjective – vectors in $\mathcal{V} \otimes \mathcal{W}$ are not all product!
  → The existence of *entangled state* in the tensor product space.

Tensor Product of Vectors

- Definition. Tensor product of vectors $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{C}^2$

$$\mathbf{a} \otimes \mathbf{b} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} := \begin{pmatrix} a_1 \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\ a_2 \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a_1 \times b_1 \\ a_1 \times b_2 \\ a_2 \times b_1 \\ a_2 \times b_2 \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ a_2 b_1 \\ a_2 b_2 \end{pmatrix} \in \mathbb{C}^{2\times 2}$$

- Properties (linearity).

$$\begin{aligned} (\mathbf{a}_1 + \mathbf{a}_2) \otimes (\mathbf{b}_1 + \mathbf{b}_2) &= \mathbf{a}_1 \otimes (\mathbf{b}_1 + \mathbf{b}_2) + \mathbf{a}_2 \otimes (\mathbf{b}_1 + \mathbf{b}_2) \\ &= \mathbf{a}_1 \otimes \mathbf{b}_1 + \mathbf{a}_1 \otimes \mathbf{b}_2 + \mathbf{a}_2 \otimes \mathbf{b}_1 + \mathbf{a}_2 \otimes \mathbf{b}_2 \end{aligned}$$

Multiple Qubits (1/3)

- Recall: a qubit is represented by a unit vector in two-dimensional linear space $\mathbb{C}^2$
- A two-qubit is represented by a unit vector in four-dimensional linear space $\mathbb{C}^{2\times 2}$
  → The computational basis:

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

- Tensor product: $|01\rangle \equiv |0\rangle \otimes |1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} := \begin{pmatrix} 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \times 0 \\ 1 \times 1 \\ 0 \times 0 \\ 0 \times 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$

Multiple Qubits (2/3)

- Product state: Given two qubits states $|a\rangle = a_1|0\rangle + a_2|1\rangle$, $|b\rangle = b_1|0\rangle + b_2|1\rangle \in \mathbb{C}^2$

$$\Rightarrow |a\rangle \otimes |b\rangle = a_1 b_1 |0\rangle \otimes |0\rangle + a_1 b_2 |0\rangle \otimes |1\rangle + a_2 b_1 |1\rangle \otimes |0\rangle + a_2 b_2 |1\rangle \otimes |1\rangle$$

$$= a_1 b_1 |00\rangle + a_1 b_2 |01\rangle + a_2 b_1 |10\rangle + a_2 b_2 |11\rangle = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ a_2 b_1 \\ a_2 b_2 \end{pmatrix}$$

- Can check $|a_1 b_1|^2 + |a_1 b_2|^2 + |a_2 b_1|^2 + |a_2 b_2|^2 = 1$
  using $|a_1|^2 + |a_2|^2 = |b_1|^2 + |b_2|^2 = 1$

Multiple Qubits (3/3)

- General 2-qubit state:

Probability amplitude

$$|\psi\rangle = a_{00}|00\rangle + a_{01}|01\rangle + a_{10}|10\rangle + a_{11}|11\rangle \in \mathbb{C}^4 \qquad \sum_{i,j\in\{0,1\}} |a_{ij}|^2 = 1$$

- Entangled state: there are states that cannot be expressed as the product form, i.e.

$$\exists|\psi\rangle \text{ s.t. } |\psi\rangle \neq |q_1\rangle \otimes |q_2\rangle, \quad \forall|q_1\rangle, |q_2\rangle \in \mathbb{C}^2$$

- The *Bell states*: $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ $|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$

$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$ $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$

Proof: $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = (a|0\rangle + b|1\rangle)(c|0\rangle + d|1\rangle) \otimes = ac|00\rangle + ad|01\rangle + bc|10\rangle + bd|11\rangle$.

Must have $ad = 0$, so either $a = 0$ or $d = 0$. Hence, either $|00\rangle$ or $|11\rangle$ has coefficient zero too, $\rightarrow \perp$

The Curse of Dimensionaliy

- For the n-fold tensor product of qubits, we write $(\mathbb{C}^2)^{\otimes n} := \otimes^n \mathbb{C}^2 = \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$ which is a space of dimension 2^n with basis $\{|i_1\rangle \otimes \cdots \otimes |i_n\rangle : i_1, \ldots i_n \in \{0,1\}\}$ labelled by 2^n n-bit strings $i_1, \ldots i_n$.
  We often write $|i_1\rangle \otimes \cdots \otimes |i_n\rangle$ simply as $|i_1 \ldots i_n\rangle$.
  This basis is called the *computational basis* of $(\mathbb{C}^2)^{\otimes n}$.

2^n $\left\{ \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} \cdots \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix} \right.$

- The dimension of the n-qubits is 2^n, which is *exponential* in the number of qubits.
  The full state description (given as the full list of amplitudes) grows *exponentially*.
  However, the description of any product state grows only linearly in n.
  The classical state space of a composite system is the *Cartesian product* of the state spaces of its constituent parts, i.e. only a linear growth of description
  → This explains some of the difficulty in simulating a quantum system on an ordinary classical computer.

Postulate 3 (Evolution)

Physical Evolution of Closed Quantum Systems

Postulate 3 (Evolution)

Any physical (finite time) evolution of a closed quantum system is represented by a *unitary* operation on the corresponding Hilbert space.

- $|\psi(t_2)\rangle = U(t_1, t_2)|\psi(t_1)\rangle$
- The time evolution of a state vector is governed by the *Schrödinger equation*: $\mathrm{i}\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle \Rightarrow U(t_1, t_2) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H(t_2 - t_1)}$

Visualizing Evolutions on Qubits

- A unitary U on $\mathbb{C}^2$ can be written in the spectral decomposition form:

$$U = \mathrm{e}^{\mathrm{i}\alpha}|\varphi\rangle\langle\varphi| + \mathrm{e}^{\mathrm{i}\beta}|\varphi^\perp\rangle\langle\varphi^\perp|$$

- Via the Bloch-sphere representation,

$$|\varphi\rangle\langle\varphi| = \tfrac{1}{2}\left(I + \vec{r}\cdot\vec{\sigma}\right) \qquad |\varphi^\perp\rangle\langle\varphi^\perp| = \tfrac{1}{2}\left(I - \vec{r}\cdot\vec{\sigma}\right)$$

- Set $\alpha = \kappa - \frac{\phi}{2}$ and $\beta = \kappa + \frac{\phi}{2}$,

$$\Rightarrow U = \mathrm{e}^{\mathrm{i}\kappa}\left(\frac{\mathrm{e}^{-\mathrm{i}\frac{\phi}{2}} + \mathrm{e}^{\mathrm{i}\frac{\phi}{2}}}{2} I + \frac{\mathrm{e}^{-\mathrm{i}\frac{\phi}{2}} - \mathrm{e}^{\mathrm{i}\frac{\phi}{2}}}{2}\vec{r}\cdot\vec{\sigma}\right)$$

$$= \mathrm{e}^{\mathrm{i}\kappa}\left(\cos\tfrac{\phi}{2}\, I - \mathrm{i}\sin\tfrac{\phi}{2}(\vec{r}\cdot\vec{\sigma})\right)$$

$$= \mathrm{e}^{\mathrm{i}\kappa}\, \mathrm{e}^{-\mathrm{i}\frac{\phi}{2}\vec{r}\cdot\vec{\sigma}}$$

Global phase can be ignored

Three parameters (r_x, r_y, r_z) are sufficient

About The Matrix Exponential

- Functional calculus for matrices: For diagonalizable $A = UDU^\dagger$,

$$e^A := I + \frac{A}{1!} + \frac{A^2}{2!} + \cdots = U\exp(D)U^\dagger$$

Holds for analytic real-valued function f

- For a special case that $A^2 = I$,

$$e^{\mathrm{i}\theta A} = I + \frac{\mathrm{i}\theta A}{1!} + \frac{-\theta^2 I}{2!} + \frac{-\mathrm{i}\theta^3 A}{3} + \cdots$$

$$= \left(I - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \cdots\right) + \mathrm{i}\left(\frac{\theta}{1!} - \frac{\theta^3}{3!} + \cdots\right) A$$

$$= \cos\theta + \mathrm{i}\sin\theta\, A$$

Euler's formula for matrices

- Note: Be careful about the *non-commutativity*: $\mathrm{e}^{A+B} \neq \mathrm{e}^A\mathrm{e}^B$ for $[A,B] \neq 0$.

Rotation About X, Y, and Z Axes

$U = \mathrm{e}^{-\mathrm{i}\frac{\phi}{2}\vec{r}\cdot\vec{\sigma}}$

- Consider the case $\vec{r} = (r_x, r_y, r_z) = (0,0,1)$,

$$U = \mathrm{e}^{-\mathrm{i}\frac{\phi}{2}}|0\rangle\langle 0| + \mathrm{e}^{\mathrm{i}\frac{\phi}{2}}|1\rangle\langle 1|$$

- For state $|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + \mathrm{e}^{\mathrm{i}\varphi}\sin\frac{\theta}{2}|1\rangle$,

$$\Rightarrow U|\psi\rangle = \mathrm{e}^{-\mathrm{i}\frac{\phi}{2}}\cos\tfrac{\theta}{2}|0\rangle\langle 0| + \mathrm{e}^{\mathrm{i}\frac{\phi}{2}+\mathrm{i}\psi}\sin\tfrac{\theta}{2}|1\rangle\langle 1|$$

Rotation along the Z axis

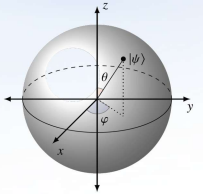
$$= \mathrm{e}^{-\mathrm{i}\frac{\phi}{2}}\left(\cos\tfrac{\theta}{2}|0\rangle\langle 0| + \mathrm{e}^{\mathrm{i}(\psi+\phi)}\sin\tfrac{\theta}{2}|1\rangle\langle 1|\right)$$

- Along Pauli X, Y, and Z axes:

$$R_x(\phi) := \mathrm{e}^{-i\frac{\phi}{2}X} = \begin{pmatrix} \cos(\frac{\phi}{2}) & -\mathrm{i}\sin(\frac{\phi}{2}) \\ -\mathrm{i}\sin(\frac{\phi}{2}) & \cos(\frac{\phi}{2}) \end{pmatrix}$$

$$R_z(\phi) := \mathrm{e}^{-i\frac{\phi}{2}Z} = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\phi} \end{pmatrix} \qquad R_y(\phi) := \mathrm{e}^{-i\frac{\phi}{2}Y} = \begin{pmatrix} \cos(\frac{\phi}{2}) & -\sin(\frac{\phi}{2}) \\ \sin(\frac{\phi}{2}) & \cos(\frac{\phi}{2}) \end{pmatrix}$$

Rotation About an Arbitrary Axis (1/2)



- Along axis $\vec{n} = (n_x, n_y, n_z) = (\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta)$

$$R_{\vec{n}}(\phi) = R_z(\varphi)R_y(\theta)R_z(\phi)R_y(-\theta)R_z(-\varphi)$$

$$= R_z(\varphi)R_y(\theta)R_z(\phi)R_y(\theta)^\dagger R_z(\varphi)^\dagger$$

$$= R_z(\varphi)R_y(\theta)\left[\cos\tfrac{\phi}{2}I - \mathrm{i}\sin\tfrac{\phi}{2}Z\right]R_y(\theta)^\dagger R_z(\varphi)^\dagger$$

$R_y(\theta)ZR_y(\theta)^\dagger = \cos\theta\, Z + \sin\theta\, X$

$$= R_z(\phi)R_y(\theta)\left[\cos\tfrac{\phi}{2}I - \mathrm{i}\sin\tfrac{\phi}{2}(\cos\theta\, Z + \sin\theta\, X)\right]R_z(\varphi)^\dagger$$

$R_z(\theta)ZR_z(\theta)^\dagger = Z$

$$= \cos\tfrac{\phi}{2}I - \mathrm{i}\sin\tfrac{\phi}{2}(\cos\theta\, Z + \sin\theta[\cos\varphi\, X + \sin\varphi\, Y])$$

$$= \cos\tfrac{\phi}{2}I - \mathrm{i}\sin\tfrac{\phi}{2}(\cos\theta\cos\varphi\, X + \sin\theta\sin\varphi\, Y + \cos\theta\, Z)$$

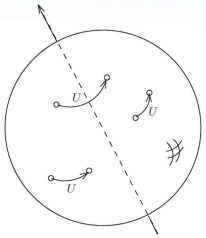
$R_z(\theta)XR_z(\theta)^\dagger = Z\cos\theta\, X + \sin\theta\, Y$

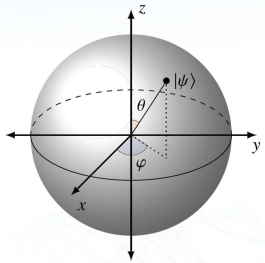
$$= \cos\tfrac{\phi}{2}I - \mathrm{i}\sin\tfrac{\phi}{2}(n_x X + n_y Y + n_z Z)$$

$$= \cos\tfrac{\phi}{2}I - \mathrm{i}\sin\tfrac{\phi}{2}(\hat{n}\cdot\vec{\sigma}) = \mathrm{e}^{-\mathrm{i}\frac{\phi}{2}\vec{n}\cdot\vec{\sigma}}$$

Rotation About an Arbitrary Axis (2/2)

- Along axis $\vec{n}$ (e.g. via the Y and Z axes are sufficient)





- Another representation [Chapters 4, N&C]

$$U = e^{i\kappa} \begin{pmatrix} e^{-i(\beta+\delta)/2}\cos\frac{\gamma}{2} & -e^{i(-\beta+\delta)/2}\sin\frac{\gamma}{2} \\ e^{i(\beta-\delta)/2}\sin\frac{\gamma}{2} & e^{i(\beta+\delta)/2}\cos\frac{\gamma}{2} \end{pmatrix} = e^{i\kappa} R_z(\beta) R_y(\gamma) R_z(\delta)$$

Physical Evolution of Open Quantum Systems

Postulate 3* (Evolution)

Any physical evolution of an open quantum system is characterized by a quantum operation, which is a completely positive and trace-preserving (CPTP) map.

- A linear mapping $\Lambda^{A\to B}: \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ is completely positive if the mapping $\Lambda^{A\to B} \otimes \mathrm{id}^{A'}$ on $\mathcal{H}_A \otimes \mathcal{H}_{A'}$ is positive for all finite dimensional extensions $\mathcal{H}_{A'}$.
  - E.g. The closed quantum evolution: $\rho \mapsto U\rho U^\dagger$.

Postulate 4 (Measurement)

Measuring a Quantum System

Postulate 4 (Measurement)

or called von Neumann measurement

- A *projective measurement* is described by a collection of projection operators $\{P_m\}$ (on the underlying Hilbert space) that satisfies the completeness condition $\sum_m P_m = I$.
- The index m refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is $|\psi\rangle$, then

$$\Pr_\psi(\text{outcome } m) = \langle\psi|P_m|\psi\rangle,$$

The Born Rule

and the system "collapses" into the *post-measurement state*

$$|\psi'\rangle = \frac{P_m|\psi\rangle}{\|P_m|\psi\rangle\|} = \frac{P_m|\psi\rangle}{\sqrt{\langle\psi|P_m|\psi\rangle}}.$$

Remarks on Physical Observables

- A projective measurement of a physical quantity (e.g. of the energy, angular momentum, etc.) carried out on a quantum system is often physically described by an Hermitian operator called *observable* O, which is embodied by the spectral decomposition $O = \sum_m a_m P_m$.
  The difference is now that we get the measurement outcome a_m not m.
- Note that the mapping $m \mapsto a_m$ is bijective. The actual choice of the naming of the distinct outcomes is of no real consequence. We can just consider $\Pr_\psi(\text{outcome } i)$, where the i-th outcome could either be the label of a projection or the i-th eigenvalue of O.
- For an example of the Pauli spin-Z observable $O = +|0\rangle\langle 0| - |1\rangle\langle 1|$, we are measuring $|\psi\rangle$ with respect to the basis $\{|0\rangle, |1\rangle\}$ and labels ± 1.

$$\Rightarrow \mathbb{E}_\psi(\text{outcome}) = \sum\nolimits_m a_m \langle\psi|P_m|\psi\rangle = \langle\psi|O|\psi\rangle.$$

Quantum Measurement Relative to a Basis

- Let $\mathcal{B} = \{|e_1\rangle, \ldots, |e_d\rangle\}$ be an orthonormal basis of $\mathcal{H}$ and write $|\psi\rangle = \sum_i a_i |e_i\rangle$.
  Now we make a quantum measurement of $|\psi\rangle$ relative to the basis $\mathcal{B}$.
  → The probability of obtaining outcome i is (Rank-1 projection)

$$\Pr_\psi(\text{outcome } i) = \langle\psi||e_i\rangle\langle e_i||\psi\rangle = |\langle e_i|\psi\rangle|^2 = |a_i|^2$$

After seeing outcome i, the state collapses to $|e_i\rangle$. If we were to apply the measurement again, we will simply see the same i with certainty

Quantum → classical

Measuring a Quantum System

Postulate 4* (Measurement)

(Positive operator-valued measure (POVM))

- A *general measurement* is described by a collection of positive semi-definite operators $\{\Pi_m\}$ (on the underlying Hilbert space) that satisfies the completeness condition $\sum_m \Pi_m = I$.
- The index m refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is ρ, then

$$\Pr_\rho(\text{outcome } m) = \mathrm{Tr}[\rho\Pi_m],$$

(The Born Rule)

and the *post-measurement state* is

$$\rho' = \frac{\sqrt{\Pi_m}\rho\sqrt{\Pi_m}}{\sqrt{\Pr(m)}}.$$

Concluding Remarks

Properties of Tensor Products (1/2)

- Definition. Tensor product of matrices $A := \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$, $B := \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix}$

$$A \otimes B = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \otimes B$$

$$:= \begin{pmatrix} a_{1,1}B & a_{1,2}B \\ a_{2,1}B & a_{2,2}B \end{pmatrix} = \begin{pmatrix} a_{1,1}\begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} & a_{1,2}\begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} \\ a_{2,1}\begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} & a_{2,2}\begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} a_{1,1}b_{1,1} & a_{1,1}b_{1,2} & a_{1,2}b_{1,1} & a_{1,2}b_{1,2} \\ a_{1,1}b_{2,1} & a_{1,1}b_{2,2} & a_{1,2}b_{2,1} & a_{1,2}b_{2,2} \\ a_{1,1}b_{3,1} & a_{1,1}b_{3,2} & a_{1,2}b_{3,1} & a_{1,2}b_{3,2} \\ a_{1,1}b_{4,1} & a_{1,1}b_{4,2} & a_{1,2}b_{4,1} & a_{1,2}b_{4,2} \end{pmatrix}$$

Properties of Tensor Products (2/2)

- Bilinearity : $(A+B)\otimes(C\otimes D) = A\otimes C + A\otimes D + B\otimes C + C\otimes D$
- Product : $(kA)\otimes B = A\otimes(kB) = k(A\otimes B)$
  $(A\otimes B)\otimes C = A\otimes(B\otimes C)$
  $A\otimes 0 = 0\otimes A = 0$
  $(A\otimes B)(C\otimes D) = (AC)\otimes(BD)$
  $A\otimes B = (I_1\otimes B)(A\otimes I_2) = (A\otimes I_2)(I_1\otimes B)$
- Inverse/Transpose : $(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}$ $\quad (A\otimes B)^T = A^T\otimes B^T$
- Multiplicativity: $\mathrm{Tr}\,[A\otimes B] = \mathrm{Tr}[A]\cdot\mathrm{Tr}[B]$
- Acting on vectors: $(A\otimes B)|\psi\rangle|\phi\rangle = (A|\psi\rangle)\otimes(B|\phi\rangle)$
  $\langle\alpha|_A\langle\beta|_B|\psi\rangle_A|\phi\rangle_B = \langle\alpha|\psi\rangle\langle\beta|\phi\rangle$
- Eigenvalues of $A\otimes B$: $\lambda_i(A)\lambda_j(B), \quad \forall(i,j)$

Concluding Remarks

- State of a quantum system:
  - Pure state (state vector or rank-1 projection) ↔ deterministic state
  - Mixed state (density operator) ↔ statistical mixture of pure states (probability distribution on the state space)
- Quantum operation:
  - Unitary transformation ↔ deterministic (and reversible) evolution
  - CPTP map (quantum channel) ↔ stochastic transformation
- Measurement:
  - Projective-valued measure (PVM) ↔ deterministic decision
  - Positive operator-valued measure (POVM) ↔ randomized decision (e.g. measure w.r.t. the X basis with 50% chances and w.r.t. the Z basis with 50%.)
- Operational interpretation of tensor product: non-interaction of subsystems.

Independence

Quantum Information and Computation
The No-Go Theorems

Hao-Chung Cheng (鄭皓中)
haochung@ntu.edu.tw

Department of Electrical Engineering
National Taiwan University

March 3, 2021

Outline

1. The No-Cloning Theorem
2. The No-Signaling Theorem
3. No-Perfect Discrimination
4. Discussions and References

What are the no-go theorems?

- In theoretical physics, *a no-go theorem* is a theorem that states that a situation (task or result) is not physically possible given some specific setup.
  - Heisenberg's uncertainty principle (1927): we cannot measure the momentum or position of a quantum particle to an arbitrary accuracy if we know the value of the other.
  - Weinberg–Witten Theorem (1980)
- Since the early days of Quantum Information Theory, no-go theorems have served as guideline in the search of a deeper understanding of quantum theory as well as for the development of applications of quantum mechanics to cryptography and computation.

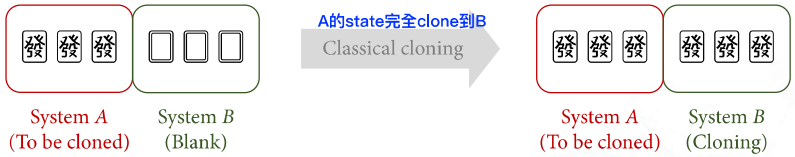
The No-Cloning Theorem

The No-Cloning Theorem (1/2)

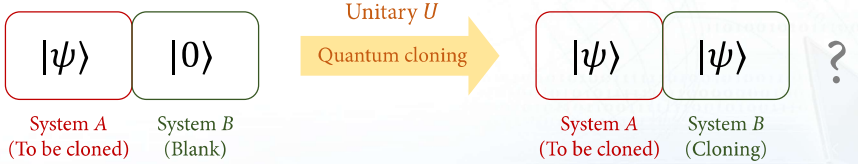
- Quantum information cannot be copied or cloned exactly.

What's quantum information?

- Classical cloning:



- Quantum cloning:



The No-Cloning Theorem (2/2)

- Setup:
  - Let $\mathcal{S}$ be any set of states that contains at lest one pair of *non-orthogonal* states
  - System A contains a state vector $|\psi\rangle \in$ to be cloned.
  - System B contains some standard blank state, e.g. $|0\rangle$
  - System M is a fixed starting state $|M_0\rangle$ as any machine/materials required

- **Theorem** [D. Dieks and W. Wootters & W. Zurek, 1982, and Yuen, 1986].
  No unitary cloning process exists that achieves exact cloning for all states in $\mathcal{S}$.

Proof of The No-Cloning Theorem (Unitary Process)

1. Let $|\psi\rangle$ and $|\phi\rangle$ be two distinct non-orthogonal states in $\mathcal{S}$.
2. The cloning process must do both the following evolutions:

對不同state都有copy功能

unit

unit

$$|\psi\rangle_A|0\rangle_B|M_0\rangle_M \overset{U}{\mapsto} |\psi\rangle_A|\psi\rangle_B|M_\psi\rangle_M$$
$$|\phi\rangle_A|0\rangle_B|M_0\rangle_M \overset{U}{\mapsto} |\phi\rangle_A|\psi\rangle_B|M_\phi\rangle_M$$

Regardless of the dimension

3. Unitary preserves inner product:

$$\langle\psi|\phi\rangle\langle 0|0\rangle\langle M_0|M_0\rangle = \langle\psi|\phi\rangle\langle\psi|\phi\rangle\langle M_\psi|M_\phi\rangle$$
$$\Rightarrow |\langle\psi|\phi\rangle| = |\langle\psi|\phi\rangle|^2\,|\langle M_\psi|M_\phi\rangle|$$

< 1

<= 1

< 1, 不可能 = 1

4. Since $|\psi\rangle \neq |\phi\rangle$, we have $0 < |\langle\psi|\phi\rangle| < 1$; $\Rightarrow 1 = |\langle\psi|\phi\rangle|\,|\langle M_\psi|M_\phi\rangle|$

Contradiction! □

Remarks (1/2)

- The no-cloning theorem remains true if further operations of adding ancillas and measurements are included.
- The proof (unitary process) only relies on Postulate 1 (State) & Postulate 3 (Unitary).
  → It holds beyond quantum mechanics:
  e.g. Generalized Probabilistic Theories (GPT), [Barrett, 2005]
  No-signalling theories, [Masanes & Acin & Gisin, 2006]
- Why considering a pair of non-orthogonal states?
  → We don't know what to be cloned (if we knew the states, just prepare them).
  → So we want a universal quantum cloning machine (at least for a pair of states)!
  → A pair of orthogonal states can be perfectly discriminated; so we can prepare them.

Remarks (2/2)

- What kind of additional information is needed to supplement the cloning?
  → No supplementary information is required for classical cloning.
  → The clone must be generated from the 'blank' system alone [Jozsa, 2002].
- It prevents us from using classical error correction techniques on quantum states. (Quantum error correction adopts other ideas; see [Shor, Steane, 1995].
- It protects Heisenberg's uncertainty relation in quantum mechanics.
- It prevents superluminal signaling (communication) via quantum entanglement.
- It forbids eavesdropper from creating copies of the transmitted quantum states
  → Vital to quantum cryptography
- However, Imperfect cloning is possible!
  → Weapons for eavesdropping attack on quantum cryptography protocols
  [Bužek & Hillery, 1996], [Bruß *et al.*, 1998]

Superluminal Signaling and Cloning (1/2)

- Superluminal signaling is possible via quantum entanglement and cloning.

1. Alice and Bob are distinctly separated and share an EPR pair:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) \quad = \frac{1}{\sqrt{2}}(|+\rangle|+\rangle + |-\rangle|-\rangle)$$

2. Alice wants to send a yes/no decision (a bit of information) to Bob at noon:
   "Yes": measure her qubit in the Z basis $\{|0\rangle, |1\rangle\}$;
   "No": measure her qubit in the X basis $\{|+\rangle, |-\rangle\}$.
3. Let Π_i denote the projection operator for Bob's outcome i:
   "Yes": $\Pr_{\text{yes}}(i) = \frac{1}{2}\langle 0|\Pi_i|0\rangle + \frac{1}{2}\langle 1|\Pi_i|1\rangle = \mathrm{Tr}\left[\Pi_i \frac{|0\rangle\langle 0|+|1\rangle\langle 1|}{2}\right] = \mathrm{Tr}\left[\Pi_i \frac{I}{2}\right]$
   "No": $\Pr_{\text{yes}}(i) = \frac{1}{2}\langle +|\Pi_i|+\rangle + \frac{1}{2}\langle -|\Pi_i|-\rangle = \mathrm{Tr}\left[\Pi_i \frac{|+\rangle\langle +|+|-\rangle\langle -|}{2}\right] = \mathrm{Tr}\left[\Pi_i \frac{I}{2}\right]$

Alice's attempted signaling is completely indistinguishable by Bob's local operations!

Superluminal Signaling and Cloning (2/2)

- Assuming cloning at Bob's side is possible:

4. After noon, Bob clones his qubits to make many copies (say, 1 million copies)
   "Yes": Bob's qubits will be all $|0\rangle$ or all $|1\rangle$;
   → all 0's or all 1's bit strings via measurement w.r.t. the Z basis
   "No": Bob's qubits will be all $|+\rangle$ or all $|-\rangle$;
   → both of $|\pm\rangle$ gives 50/50 outcome of 0 and 1 w.r.t. the Z basis
   → So Bob will obtain a uniformly random bit string of length one million
5. Now the "yes"/"no" cases are distinguishable except with probability $2/10^6$!

With (arbitrarily) high probability, Bob can instantaneously get Alice's message!

The No-Signaling Theorem

The No-Signaling Theorem

- Setup:
  - Alice and Bob share a joint state $|\phi\rangle_{AB}$.
  - Alice performs local measurement (w.r.t. some basis) on her part.
  - Bob performs local measurement (w.r.t. some basis) on her part.

Freely to choose a basis

- Theorem.
  No local actions by Alice can change the output probability distribution of any local quantum process by Bob.
  Namely, Alice cannot convey any information to Bob by performing any local operations.

The quantum formalism prevents us from harnessing the nonlocality for instantaneous communication.

Proof of The No-Signaling Theorem (1/2)

1. (Alice didn't perform measurement)

Bob performs a local measurement w.r.t. basis $\{|b\rangle\}$ labeled by outcomes $\{b\}$.

$$|\phi\rangle_{AB} = \sum_{a,b} c_{ab}|a\rangle|b\rangle$$

$\{|a\rangle\}$ is an ONB on system A

Probability of Bob seeing outcome b:

Bob

$$\begin{aligned}\Pr(b) &= \langle\phi|(I_A \otimes |b\rangle\langle b|)|\phi\rangle \\ &= \langle\phi|(I_A \otimes |b\rangle\langle b|)|\textstyle\sum_{a',b'} c_{a'b'}|a'\rangle|b'\rangle \\ &= \langle\phi|\textstyle\sum_{a',b'} c_{a'b'}(I_A \otimes |b\rangle\langle b|)|a'\rangle|b'\rangle \\ &= \langle\phi|\textstyle\sum_{a'} c_{a'b}|a'\rangle|b\rangle \\ &= \textstyle\sum_a |c_{ab}|^2\end{aligned}$$

Proof of The No-Signaling Theorem (1/2)

2. (Alice first performs measurement w.r.t. basis $\{|a\rangle\}$)

If Alice gets an outcome a,

Alice

$$\Pr(a) = \langle\phi|(|a\rangle\langle a| \otimes I_B)|\phi\rangle = \textstyle\sum_b |c_{ab}|^2$$

$$|\phi'\rangle_{AB} = \frac{(|a\rangle\langle a| \otimes I_B)|\phi\rangle_{AB}}{\Pr(a)}$$

Bob then measures his part w.r.t $\{|b\rangle\}$, and gets outcome b with probability:

$$\begin{aligned}\Pr(b|a) = \langle\phi'|(I_A \otimes |b\rangle\langle b|)|\phi'\rangle &= \tfrac{1}{\Pr(a)}\langle\phi|(I_A \otimes |b\rangle\langle b|)(|a\rangle\langle a| \otimes I_B)|\phi\rangle \\ &= \tfrac{1}{\Pr(a)}\langle\phi|((|a\rangle\langle a| \otimes |b\rangle\langle b|)|\phi\rangle \\ \Rightarrow \Pr(a,b) = \Pr(b|a)\Pr(a) &= ((|a\rangle\langle a| \otimes |b\rangle\langle b|)|\phi\rangle \\ \Rightarrow \Pr(b) = \textstyle\sum_a \Pr(a,b) &= \textstyle\sum_a |c_{ab}|^2\end{aligned}$$

□

Remarks

- The marginal probability ~~distribution~~ for b is the same as the case of Alice no having done anything on her side; it is *experimentally indistinguishable* to Bob.
- Bob couldn't tell the difference no matter which local operations were chosen by Alice.
  → holds for the case of local unitaries/ancillas or general measurements by Alice.
- The choices of measurement cannot affect the probability of Bob's measurement result. However, the value a that Alice obtains *can* affect Bob's probabilities (if he knew a).
- The order of local measurements $I_A \otimes \Pi_b$ & $\Pi_a \otimes I_B$ doesn't matter ∵ they commute.
- No-signaling: $\sum_a \Pr(a,b|x,y) = \sum_a \langle\phi|\Pi_a^x \otimes \Pi_b^y|\phi\rangle = \langle\phi|I_A \otimes \Pi_b^y|\phi\rangle = \Pr(b|y)$
  (x and y denote indices of local measurements chose by Alice and Bob respectively.)

$$\textstyle\sum_a \Pr(a,b|x,y) = \Pr(b|x,y) = \Pr(b|y)$$

Remark on The Communication Complexity

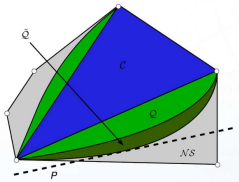
- Study of the non-local effects dates back to the 1935 'EPR' paper by A. Einstein, B. Podolsky, and N. Rosen.
- Later in 1960's J. Bell introduced the *Bell inequalities*, providing a simple and experimental accessible way of demonstrating the non-local effects.
- In early 1990's, it was realized that if Alice and Bob shared entangled states and were allowed classical communication too, although entanglement by itself cannot provide classical communication (by no-signaling), it can greatly (in some cases exponentially) reduce the amount of classical communication needed to achieve some distributed tasks, involving inputs from Alice and Bob.
  → resulting a new research area, called *quantum communication complexity*.

[Buhrman & Cleve & Massar & de Wolf, 2010]

Correlations beyond Quantum

- Why no-signaling? [Brandao & Harrow, 2017]
  - Foundational: minimum assumption for a plausible theory
  - Operational: yields the well-defined "partial trace" or "miginalization"
  - Computational: yields efficient linear program
- Why Quantum Mechanics?
  - Which properties of QM are genuine quantum?
  - Theory beyond QM – how to generalize concepts like entanglement or quantum correlations?
- No-Signaling Theories or Generalized Probabilistic Theories (GPT)

[Navascués *et al.*, 2015] [Barrett, 2005]



No-Perfect Discrimination

Distinguishing Non-Orthogonal States

- Suppose the unknown quantum system is promised to be one of the two non-orthogonal states $|\psi_0\rangle$ (null hypothesis H_0) and $|\psi_1\rangle$ (alternative hypothesis H_1). We wish to determine which one it is, i.e. the value of subscript i.

$|\psi_0\rangle$ — 0

$|\psi_1\rangle$ — 1

Using the same measurement!

- We have seen that this is impossible to do *with certainty* (why?) Can we still obtain *some* information about i, and then how much?

e.g. random guessing

- Problem: What is the optimal success probability and the optimal measurement?

$$\mathsf{P}_\mathrm{s} = \Pr\{H_0\}\Pr\{H_0|H_0\} + \Pr\{H_1\}\Pr\{H_1|H_1\}$$

Depends on measurements

Holevo–Helström Theorem

- Pure states: Given two equally likely states $|\psi_0\rangle$ and $|\psi_1\rangle$ with $|\langle\psi_0|\psi_1\rangle| = \cos\theta$,

$$\Rightarrow \mathsf{P}_\mathrm{s}^\star = \tfrac{1}{2}(1+\sin\theta) = \tfrac{1}{2}\left(1+\sqrt{1-|\langle\psi_0|\psi_1\rangle|^2}\right)$$

- Remark:
  - Perfect discrimination if and only if $|\psi_0\rangle \perp |\psi_1\rangle$.
  - Since $|\langle\psi_0|\psi_1\rangle|^n \to 0$, discrimination is *reliable* in the asymptotic limit.

- Mixed states: Given two states ρ_0 and ρ_1 with prior probability p_0 and p_1,

$$\Rightarrow \mathsf{P}_\mathrm{s}^\star = \tfrac{1}{2}\left(1+\|p_0\rho_0 - p_1\rho_1\|_1\right)$$

$$\Pi_0^\star = \{p_0\rho_0 - p_1\rho_1 > 0\}$$

|M| = sqrt(M^*M)

Trace norm $\|M\|_1 \coloneqq \mathrm{Tr}[|M|]$

Proof – Pure States (1/2)

1. Perform measurement $\{\Pi_0, \Pi_1\}$ and then optimize it (assuming equiprobable):

$$\begin{aligned}
\mathsf{P}_\mathrm{s} &= \tfrac{1}{2}\Pr\{\text{outcome } 0 \,|\, |\psi_0\rangle\text{was sent}\} + \tfrac{1}{2}\Pr\{\text{outcome } 1 \,|\, |\psi_1\rangle\text{was sent}\} \\
&= \tfrac{1}{2}\langle\phi_0|\Pi_0|\phi_0\rangle + \tfrac{1}{2}\langle\phi_1|\Pi_1|\phi_1\rangle \\
&= \tfrac{1}{2}\mathrm{Tr}\left[\Pi_0|\phi_0\rangle\langle\phi_0|\right] + \tfrac{1}{2}\mathrm{Tr}\left[\Pi_1|\phi_1\rangle\langle\phi_1|\right] \\
&= \tfrac{1}{2}\mathrm{Tr}\left[\Pi_0\left(|\phi_0\rangle\langle\phi_0| - |\phi_1\rangle\langle\phi_1|\right)\right]
\end{aligned}$$

Complete relation $\Pi_1 = I - \Pi_0$

2. $\Delta := |\phi_0\rangle\langle\phi_0| - |\phi_1\rangle\langle\phi_1|$

(1) At most two non-zero eigenvalues
(2) $\mathrm{Tr}[\,\Delta\,] = 0$

$\Rightarrow \Delta = \lambda|p\rangle\langle p| - \lambda|m\rangle\langle m|$

Proof – Pure States (2/2)

3. Consider the 2-dimensional subspace spanned by $|\psi_0\rangle$ and $|\psi_1\rangle$
→ Write states $|\psi_0\rangle, |\psi_1\rangle$ in terms of the basis $\{|\psi_0\rangle, |\psi_0^\perp\rangle\}$:

$$|\psi_0\rangle = \begin{pmatrix}1\\0\end{pmatrix} \qquad |\psi_1\rangle = c_0|\psi_0\rangle + c_1|\psi_0^\perp\rangle = \begin{pmatrix}c_0\\c_1\end{pmatrix}$$

$|c_0|^2 + |c_1|^2 = 1$

$$\Delta = \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} - \begin{pmatrix}c_0\\c_1\end{pmatrix}\begin{pmatrix}\overline{c_0} & \overline{c_1}\end{pmatrix} = \begin{pmatrix}|c_1|^2 & -c_0\overline{c_1}\\ -c_1\overline{c_0} & -|c_1|^2\end{pmatrix}$$

$$\Rightarrow \lambda = \pm|c_1| = \pm\sin\theta, \quad \cos\theta = |\langle\psi_0|\psi_1\rangle|$$

4. The optimal value is achieved at $\Pi_0 = |p\rangle\langle p|$:

$$\mathsf{P}_\mathrm{s} = \tfrac{1}{2} + \tfrac{\lambda}{2}\mathrm{Tr}\left[\Pi_0\left(|p\rangle\langle p| - |m\rangle\langle m|\right)\right]$$

$$\Rightarrow \mathsf{P}_\mathrm{s}^\star = \tfrac{1}{2}(1+\sin\theta)$$

□

Proof – Mixed States

1. Achievability (Sufficiency).

$$\begin{aligned}
\mathsf{P}_\mathrm{s} &= p_0\mathrm{Tr}\left[\rho_0\Pi_0\right] + p_1\mathrm{Tr}\left[\rho_1\Pi_1\right] \\
&= \tfrac{1}{2}\mathrm{Tr}\left[(p_0\rho_0 + p_1\rho_1)(\Pi_0 + \Pi_1)\right] + \tfrac{1}{2}\mathrm{Tr}\left[(p_0\rho_0 - p_1\rho_1)(\Pi_0 - \Pi_1)\right] \\
&= \tfrac{1}{2} + \tfrac{1}{2}\mathrm{Tr}\left[(p_0\rho_0 - p_1\rho_1)(\Pi_0 - \Pi_1)\right]
\end{aligned}$$

Let $A := p_0\rho_0 - p_1\rho_1$ and choose $\Pi_0 = \{p_0\rho_0 \succ p_1\rho_1\}$.

$$\Rightarrow P_\mathrm{s} = \tfrac{1}{2} + \tfrac{1}{2}\mathrm{Tr}\left[A\{A > 0\} - A\{A \le 0\}\right] = \tfrac{1}{2} + \tfrac{1}{2}\mathrm{Tr}\left[|A|\right]$$

2. Optimality (Necessity).
Hölder's inequality: $|\langle A, B\rangle| \le \|A\|_p\|B\|_q, \quad \frac{1}{p} + \frac{1}{q} = 1$

$$\Rightarrow \mathrm{Tr}\left[A(\Pi_0 - \Pi_1)\right] \le \|A\|_1 \underbrace{\|\Pi_0 - \Pi_1\|_\infty}_{\le 1} \le \|A\|_1$$

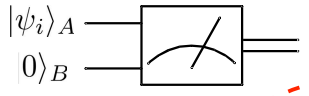
□ 不會考

Remarks (1/2)

- One can apply a unitary U before measurement → resulting another measurement

$$\Pr(i) = (\langle\phi|U^\dagger)\Pi_i(U|\phi\rangle) = \langle\phi|U^\dagger\Pi_i U|\phi\rangle$$

- One can adjoint ancilla states and apply measurement on the joint systems



$$\begin{aligned}
\Pr(i) &= \langle\phi|_A \otimes \langle 0|_B P_{AB,i}|\phi\rangle_A \otimes |0\rangle_B \\
&= \langle\phi|_A \underbrace{I_A \otimes \langle 0|_B P_{AB,i} I_A \otimes |0\rangle_B}_{=:\,\Pi_i} |\phi\rangle_A
\end{aligned}$$

POVM: $0 \le \Pi_i \le I, \Sigma_i \Pi_i = I$

- The optimal measurement depends on the states $|\psi_i\rangle$, but we don't know i.
→ The Maximum *a posteriori* decoding (MAP) rule

Remarks (2/2)

- Ambiguous state discrimination: gives one of the i's, could be erroneous.
- Unambiguous state discrimination: gives one of the i's and an *inconclusive* answer.
  - If measurement outcome i occurs, then the state was *certainly* $|\psi_i\rangle$
  - If measurement outcome 'fail' occurs, we have lost all information about the given state
- Projection measurements (PVMs) are sufficient for the binary pure/mixed state case.
- But general measurements (POVMs) might outperform PVMs in
  - Multiple ambiguous state discrimination
  - Unambiguous state discrimination

  → This motivates the study of POVMs (just as the classical scenario):

PVM ↔ Deterministic decisions
POVM ↔ Randomized decisions

Discussions & References

Other No-Go Theorems

- No-signaling: $\sum_a \Pr(a,b|x,y) = \Pr(b|\cancel{x},y)$
- No-perfect-discrimination: $\forall|\psi_0\rangle,|\psi_1\rangle,\ \mathsf{P}_{\mathrm{s}}^{\star} = \frac{1}{2}+\frac{1}{2}\sqrt{1-|\langle\psi_0|\psi_1\rangle|^2}$
- No-cloning: $\forall|\psi_0\rangle \not\perp |\psi_1\rangle,\ \nexists U \text{ s.t. } |\psi_i\rangle \overset{U}{\mapsto} |\psi_i\rangle|\psi_i\rangle$
- No-deleting: [Zurek and Pati & Braunstein, 2000] $\forall|\psi_0\rangle \not\perp |\psi_1\rangle,\ \nexists U \text{ s.t. } |\psi_i\rangle|\psi_i\rangle \overset{U}{\mapsto} |\psi_i\rangle|0\rangle$
- No-partial-erasure: [Pati & Sanders, 2006] $\forall|\psi_0\rangle \not\perp |\psi_1\rangle,\ \nexists U \text{ s.t. } |\psi_i(\theta,\phi)\rangle \mapsto |\psi_i(\theta)\rangle$
- No-flipping: [Bužek & Hillery & Werner, 1999] $\nexists$ universal-NOT operation s.t. $|\psi\rangle \mapsto |\psi^{\perp}\rangle$
- No-broadcasting: [Barnum *et al.*, 1996] $\forall[\rho_0,\rho_1]\neq 0, \nexists \mathrm{CPTP}\ \Lambda\ , \tilde{\rho}_i = \Lambda(\rho_i\otimes\tau)$ s.t. $\mathrm{Tr}_A[\tilde{\rho}_i] = \rho_i$ and $\mathrm{Tr}_B[\tilde{\rho}_i] = \rho_i$
- No-programming: [Nielsen & Chuang, 1997] $\forall U_0\neq U_1,\ \nexists|\mathcal{P}_{U_0}\rangle \not\perp |\mathcal{P}_{U_1}\rangle \text{ s.t. } |\psi\rangle||\mathcal{P}_{U_i}\rangle \mapsto U_i|\psi\rangle\otimes|\mathcal{P}'_{U_i}\rangle$
- No-hiding: [Braunstein & Pati, 2007]

If the quantum information is missing then it must move to somewhere else and it cannot be hidden in the correlations between a pair of systems.

What is information?

- The classical information is represented by *distinguishable* states of a physical system.

  → In classical physics, it is axiomatic that any two different states are *perfectly distinguishable* by a experimental measurement.
- In quantum physics, two (or more) states cannot be perfectly distinguished unless they are *mutually orthogonal*.
- In quantum information theory, information is encoded in the states of a quantum system.
  → For closed systems, these are given by the state vectors of the underlying Hilbert space.
- In principle, we can *perfectly prepare* any desired pure state (but of course there are difficulties in experimental implementations).
- But if we receive such a state (of *unknown* identity), we cannot identify it *with certainty*; yet we are still receiving a kind of definite signal or message, albeit *unreadable*.
- We use the term *quantum information* to refer to what we acquire when we receive a quantum state.

Concluding Remarks

- What types of operations one *can* or *cannot* perform on quantum information?
  (E.g. one can swap two unknown states, and can teleportate an unknown state.)
- To (exactly) clone or to delete, one needs to know the state, and one cannot know it unless one measures that state.
- An (single) unknown quantum state cannot be measured without being disturbed.
  → A single (qubit) measurement outcome only gives one bit of classical information.
  → A qubit requires infinite bits of classical information to describe it.
  → Multiple copies of the unknown quantum states are required.
- *If* determination of a quantum state by a single measurement was possible
  → then perfect cloning would be possible
  → then superluminal signaling would be possible
- Go ahead to devise a no-go theorem (e.g. for the final project)!

References (1/4)

- D. Diecks, "Communication by EPR devices," *Physics Letters A*, 92(6): 271–272, 1982.
- W. K. Wootters and W. H. Zurek, "A Single Quantum Cannot be Cloned," *Nature*, 299: 802–803, 1982.
- H. P. Yuen, "Amplification of quantum states and noiseless photon amplifiers," *Physics Letters A*, 113(8), 13:405–407, 1986.
- D. Bruss, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, J. A. Smolin, "Optimal Universal and State-Dependent Quantum Cloning," *Physical Review A*, 57(2368), 1998.
- V. Buzek and M. Hillery, "Quantum Copying: Beyond the No-Cloning Theorem," *Physical Review A*, 54(1844), 1996.
- R. Jozsa, "A stronger no-cloning theorem," arXiv:quant-ph/0204153.
- V. Scarani, S. Iblisdir, N. Gisin, and A. Acín, "Quantum cloning," *Reviews of Modern Physics*, 77(1225), 2005.

References (2/4)

- W. H. Zurek, "Schrödinger's sheep," *Nature*, 404:130–131, 2000.
- A. K. Pati, S. L. Braunstein "Impossibility of deleting an unknown quantum state," *Nature*, 303: 164–165, 2000.
- M. Horodecki, R. Horodecki, A. S. De, and U. Sen, "No-deleting and no-cloning principles as consequences of conservation of quantum information," arXiv:quant-ph/0306044.
- A. K. Pati, "General impossible operations in quantum information," *Physical Review A*, 66, 062319, 2002.
- V. Bužek, M. Hillery, and R. F. Werner, "Optimal manipulations with qubits: Universal-NOT gate," *Physical Review A*, 60, R2626, 1999.
- A. K. Pati and B. C. Sanders, "No partial erasure of quantum information," *Physics Letters A*, 359:31–36, 2006.
- S. L. Braunstein and A. K. Pati, "Quantum Information Cannot Be Completely Hidden in Correlations: Implications for the Black-Hole Information Paradox," *Physical Review Letters*, 98, 080502, 2007.

References (3/4)

- Howard Barnum, Carlton M. Caves, Christopher A. Fuchs, Richard Jozsa, and Benjamin Schumacher, "Noncommuting Mixed States Cannot Be Broadcast," *Physical Review Letters*, 76, 2818, 1996.
- M. A. Nielsen and I. L. Chuang, "Programmable Quantum Gate Arrays," *Physical Review Letters*, 79(2), 1997.
- A. M. Kubicki, C. Palazuelos, D. Pérez-García, "Resource Quantification for the No-Programming Theorem," *Physical Review Letters*, 122, 080505, 2019.
- H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, "Nonlocality and communication complexity," *Reviews of Modern Physics*, 82(665), 2010.
- LI. Masanes, A. Acín, and N. Gisin, "General properties of nonsignaling theories," *Physical Review A*, 73(012112), 2006.
- J. Barrett, "Information processing in generalized probabilistic theories," arXiv:quant-ph/0508211.

References (4/4)

- F. G. S. L. Brandao and A. W. Harrow, “Quantum de Finetti Theorems under Local Measurements with Applications", *Communications in Mathematical Physics*, 353(2), 469-506, 2017
- Navascués, Yelena Guryanova, Matty J. Hoban, Antonio Acín, " Almost quantum correlations," *Nature Communications* 6, 6288, 2015.
- H. Barnum, J. Barrett, M. Leifer, and A. Wilce, "Generalized No-Broadcasting Theorem," *Physical Review Letters*, 99(240501), 2007.
- A. S. Holevo, *Statistical Structure of Quantum Theory*, Springer, 2001.
- A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, Springer, 2011.
- A. S. Holevo, *Quantum Systems, Channels, Information*, de Gruyter, 2013.

Quantum Information and Computation
The Basic Quantum Protocols

Hao-Chung Cheng (鄭皓中)

haochung@ntu.edu.tw

Department of Electrical Engineering

National Taiwan University

March 10, 2021

Outline

Motivations

Prologue – A Secure Communication (1/2)

1. Bob randomly sends a green or red paper to Alice.
2. Alice sends '0' by returning the paper with the *same* color;
   sends '1' by sending the paper with the *opposite* color.
3. Bob check the received paper to see whether the color was the same as before.

Can Eavesdropper learn anything from wiretapping Alice's message?

- '0': same color
- '1': opposite color

50%

50%

Prologue – A Secure Communication (2/2)

- In 1926 Vernam proposed the first provably secure cryptographic protocol, known as the *one-time pad* , or *Vernam cipher.*
- The key is represented by a *random string* of bits, which is used to lock and unlock the confidential message.
- The message itself is another string of bits. *Binary addition* is used for ciphering.

$0\oplus 0=1\oplus 1=0;\ 0\oplus 1=1\oplus 0=1$

Ciphertext: b⊕k=11101010

Source text: b=01101100 Deciphered text: b⊕k⊕k=01101100

Key: k=10000110

Key: k=10000110

Remarks on The One-Time Pad

- Classical communication is used and assumed to be *noiseless*.
- If the key is a random string then the transmitted ciphertext is a random string too.
  → No one intercepting the ciphertext has access to any meaningful information except the length of the message (any message of this length is equally probable).
- Note that the key shouldn't be reused; it has be random; that's why "one-time".
- The shared key forms an ensemble of pairs of classical bits described by a joint probability $p_{00} = p_{11} = 1/2$ and $p_{01} = p_{10} = 0$.
  → *Classically maximally correlated state*, *shared randomness*, or denoted as [cc].
- Encoded bit string (nor the key) contains no information of the original messages (Shannon, 1949). → It is the correlation between the source & key that does!
- Can we replace any of the *source*, *channel*, and *key* by its quantum counterpart?
- The protocol is secure only if the key distribution is secure and hidden from others.

Non-Local Resources

Communication

- What is communication? (We consider noiseless communication for now.)
  → Simply put, the state in system A ends up in system B.

- **Noiseless cbit channel**: $0 \mapsto 0; 1 \mapsto 1$.
  Measure the state w.r.t. to $\{|0\rangle, |1\rangle\}$ and send the post-measurement state.
  We call this ability "1 cbit", or represent it as $[c \to c]$.

- **Noiseless qbit channel**: $|x\rangle_A \mapsto |x\rangle_B$ for all basis states
  This implies to $|\psi\rangle_A \mapsto |\psi\rangle_B$ for all $|\psi\rangle_A \in \mathcal{H}$.
  We call this ability "1 qubit", or $[q \to q]$.

A qubit channel sends a "system".

Non-Local Resources

- **Quantum correlation (Entanglement)**:
  "1 ebit", or $[qq]$ is an EPR pair $\frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B)$ shared between the systems A & B.
- Non-local resources: If two spatially separated parties share it or if one party uses it to communicate to another; e.g. 1 cbit, 1 qubit, 1 ebit.
- Resource inequality: `this` ≥ `that`
  The resources enumerated in "`this`" can be used to do any task that can be performed using "`that`" .
  ∃At least one protocol that uses the resource on the left to give resource on the right.
- For example, 1 qubit ≥ 1 cbit, or $[q \to q] \geq [c \to c]$:
  - Alice has a classical bit $b \in \{0,1\}$ and creates a qubit state vector $|b\rangle_A$.
  - She uses a qubit channel to transfer this to Bob.
  - Bob measures the received $|b\rangle_B$ w.r.t. $\{|0\rangle, |1\rangle\}$ to readout b.

Can we send more cbits via 1 qubit?

Basis Resources Inequalities (1/2)

- A qubit can deliver a cbit: 1 qubit ≥ 1 cbit
- A d-dimensional state can *never* be used to perfectly send more than d messages.
  It is even impossible to communicate at an asymptotic rate higher than the trivial one cbit per qubit sent (Holevo bound, 1973).

$$1 \text{ qubit} \gneqq N \text{ cbits } \forall N > 1$$

- Finite cbits cannot perfectly represent a qubit.
  (Otherwise would violate the no-cloning or the no-perfect discrimination).

$$N \text{ cbits} \gneqq 1 \text{ qubit } \forall N \geq 1$$

Basis Resources Inequalities (2/2)

- Entanglement along cannot be used for communication (by no-signaling).

$$N \text{ ebits} \gneqq 1 \text{ cbit } \forall N \geq 1 \qquad N \text{ ebits} \gneqq 1 \text{ qubit } \forall N \geq 1$$

- Communication (bandwidth) & entanglement are assumed to be perfect but not free.

$$1 \text{ cbit} \gneqq N \text{ cbits } \forall N > 1 \qquad M \text{ cbits} + 1 \text{ qubit} \gneqq N \text{ qubits } \forall N > 1, M \geq 1$$

$$1 \text{ ebit} \gneqq N \text{ ebits } \forall N > 1$$

- Classical communication cannot be used to generate entanglement. [§12.5, N&C]
  (The rigorous proof would be given later in this semester.)

$$N \text{ cbit} \gneqq 1 \text{ ebit } \forall N \geq 1$$

Entanglement Distribution

Entanglement Distribution

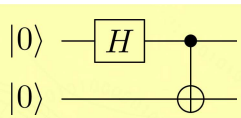
1 qubit ≥ 1 ebit

1. Alice locally prepares an EPR pair on her systems A and A'.
   - Hadamard gate H creates *superposition* on A (it is self-inverse, hence unitary).

$$|b\rangle \xrightarrow{H} \frac{1}{\sqrt{2}}\left(|0\rangle + (-1)^b|1\rangle\right), \quad b \in \{0,1\} \qquad H := \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$$

   - Controlled-NOT gate entangles systems A and A'.

$$|c\rangle \mapsto |c\rangle, \quad |t\rangle \mapsto |c \oplus t\rangle$$



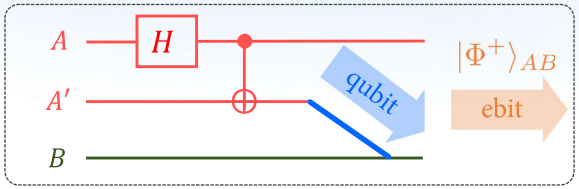
$$|00\rangle \overset{H\otimes I}{\mapsto} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\otimes|0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) \overset{CNOT}{\mapsto} \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

2. Alice sends system A' to Bob:
   (via 1 qubit channel)

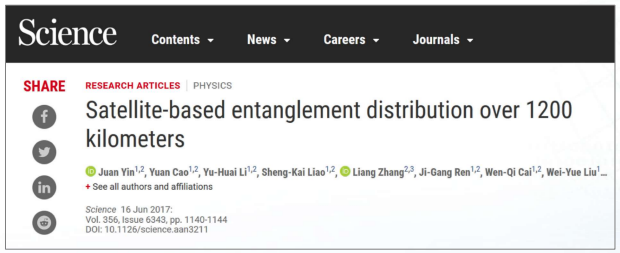
$$\frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_{A'} + |1\rangle_A|1\rangle_{A'}) \mapsto \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B)$$

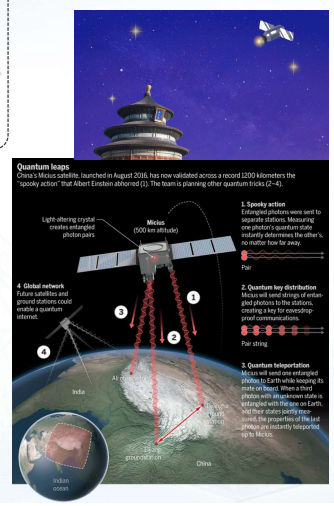
Remarks on Entanglement Distribution

- Protocol:



- Can a qubit be used to generate more ebits? No!
- Experimental implementation





Quantum Dense Coding

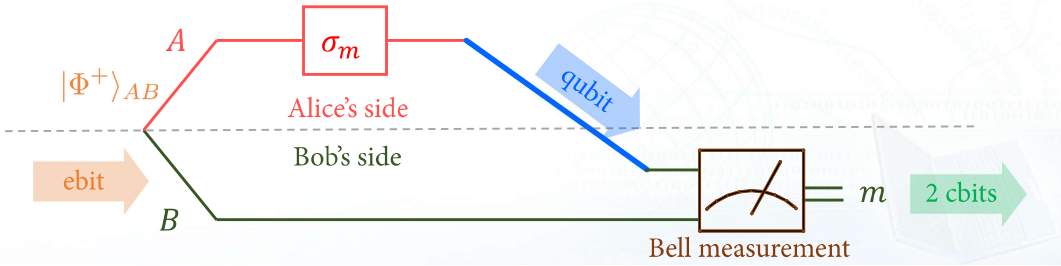
Quantum Dense Coding

1 qubit + 1 ebit ≥ 2 cbits

- Recall that 1 qubit can perfectly communicate *at most* 1 cbit.
- In 1992, Bennett and Wiesner proposed a *dense coding protocol* to perfectly communicate 2 cbits with 1 qubit and 1 ebit; later experimentally verified by Zeilinger in 1995/1996.
- Recall the one-time pad: 1 cbit + [cc] ≥ 1 private cbit.
- In the quantum dense coding scenario:
  - The secrete key (1 maximally correlated classical state) was replaced with 1 maximally entangled state (1 ebit).
  - The classical communication (1 cbit) was replaced with quantum communication (1 qubit).
- Encode cbits into qubits (with the assistance of ebits).
- Question: Is the quantum dense coding secure?

Protocol of Dense Coding

1. Alice shares a Bell state (EPR pair) $|\Phi^+\rangle_{AB}$ (say, prepared by a third party Charlie).
2. If Alice's message is $m \in \{0,1,2,3\}$ (or $ij \in \{00,01,10,11\}$), she applies to her qubit:
   $\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\sigma_1 = X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $\sigma_2 = Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ $\sigma_3 = -iY = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$
3. Alice sends her part to Bob via 1 qubit channel.
4. Bob applies the *Bell measurement* $\{|\Phi^\pm\rangle, |\Psi^\pm\rangle\}$ on both qubits to obtain 2 cbits.



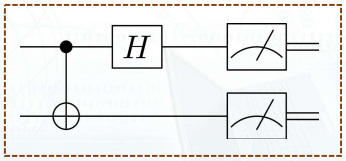
The Bell Measurement

- The four Bell states (the EPR states or the Bell basis) comprise an orthonormal basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$ of two qubits.

$$|\Phi^+\rangle = \tfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \quad |\Psi^+\rangle = \tfrac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$
$$|\Phi^-\rangle = \tfrac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \quad |\Psi^-\rangle = \tfrac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

- They characterize two classical bits: $P_{00} = |\Phi^+\rangle\langle\Phi^+|$ $P_{01} = |\Psi^+\rangle\langle\Psi^+|$
  $P_{10} = |\Phi^-\rangle\langle\Phi^-|$ $P_{11} = |\Psi^-\rangle\langle\Psi^-|$
- $U = \text{C-NOT}(H \otimes I)$ changes the computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ to the Bell basis.

  → To measure w.r.t. the Bell basis, apply U^{-1} and then measure w.r.t. the computational basis.



Rotate the Bell basis back to the computational basis measurement = Bell measurement

Proof of Dense Coding

If Alice wants to send	Alice applies
00	I
01	X
10	Z
11	$XZ = -iY$
ij	X^jZ^i

$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $\sigma_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ $\sigma_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

ij	After Alice action and communication	Bob's end	→ C-NOT gate	→ H gate
00	$(I \otimes I)\|\Phi^+\rangle$	$= \frac{1}{\sqrt{2}}(\|00\rangle + \|11\rangle)$	$= \frac{1}{\sqrt{2}}(\|00\rangle + \|10\rangle)$	$= \|00\rangle$
01	$(X \otimes I)\|\Phi^+\rangle$	$= \frac{1}{\sqrt{2}}(\|10\rangle + \|01\rangle)$	$= \frac{1}{\sqrt{2}}(\|11\rangle + \|01\rangle)$	$= \|01\rangle$
10	$(Z \otimes I)\|\Phi^+\rangle$	$= \frac{1}{\sqrt{2}}(\|00\rangle - \|11\rangle)$	$= \frac{1}{\sqrt{2}}(\|00\rangle - \|10\rangle)$	$= \|10\rangle$
11	$(XZ \otimes I)\|\Phi^+\rangle$	$= \frac{1}{\sqrt{2}}(\|10\rangle - \|01\rangle)$	$= \frac{1}{\sqrt{2}}(\|11\rangle - \|01\rangle)$	$= \|11\rangle$

Bell states

□

Remarks on Dense Coding

- We have seen 1 qubit + 1 ebit ≥ 2 cbits, but why not 1 qubit + 1 qubit ≥ 2 cbits?

  The shared entanglement is *independent* of the messages that will be set!
  (Hence the ebit is treated as a *resource* in practice.)

- One might think that 1 qubit can carry 2 cbits.

  However, if Eve intercept the transmitted 1 qubit, she will learn nothing from it!
  → The dense coding protocol is *secure* (just like the one-time pad).
  → The information of 2 cbits is not hidden in the 1 qubit.
  → It is the *correlation* between the 2 systems (the four Bell states) that contain 2 cbits.

- It belongs to the research of *"entanglement-assisted classical communication"*.

- Question: Can we consume less 1 qubit & 1 ebit (asymptotic average cost)?

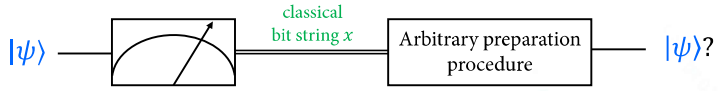
→ No! The dense coding protocol is *optimal*.

Quantum Teleportation

Quantum Teleportation

2 cbits + 1 ebit ≥ 1 qubit

- Suppose Alice wants to communicate an unknown qubit state $|\psi\rangle$ to Bob (i.e. a quantum message), but is only able to send a classical bit strings.
  Can she do it? (Encoding qubits into cbits)



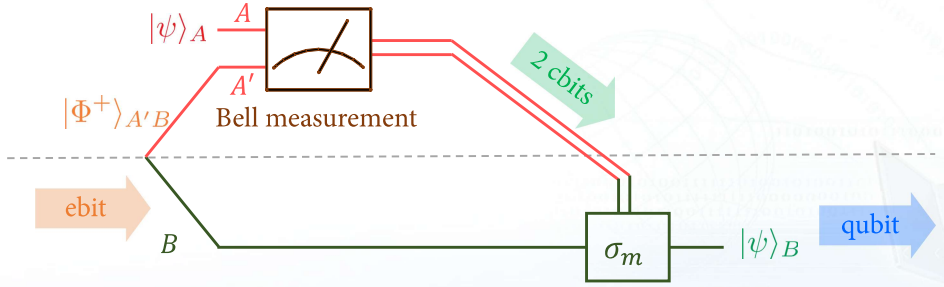
- This is not possible because of the resource inequality N cbits $\ngeq$ 1 qubit $\forall N \geq 1$

- In 1993, Bennett *et al.* invented the *quantum teleportation* that transmits 1 qubit (i.e. an arbitrary qubit state) to Bob with 2 cbits and 1 ebit.
  The protocol was later experimentally verified in 1998.


nature
Explore our content ▾ Journal information ▾
nature > letters > article
Published: 09 August 2017
Ground-to-satellite quantum teleportation
Ji-Gang Ren, Ping Xu, [...] Jian-Wei Pan
Nature 549, 70–73(2017) | Cite this article

Protocol of Teleportation

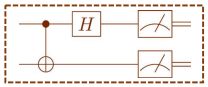
1. Alice shares a Bell state (EPR pair) $|\Phi^+\rangle_{A'B}$ (say, prepared by a third party Charlie).
2. Alice applies the Bell measurement on her state $|\psi\rangle_A$ and system A'.
3. Alice sends the 2-cbit outcome $m \in \{0,1,2,3\}$ to Bob.
4. Bob applies σ_m on his system B according to the received m, getting $|\psi\rangle_B$.



Proof of Teleportation (1/2)

$$|\psi\rangle_A \otimes |\Phi^+\rangle_{A'B} = (a|0\rangle_A + b|1\rangle_A) \otimes \frac{1}{\sqrt{2}} (|00\rangle_{A'B} + |11\rangle_{A'B})$$

$$= \frac{1}{\sqrt{2}} (a|000\rangle_{AA'B} + a|011\rangle_{AA'B} + b|100\rangle_{AA'B} + b|111\rangle_{AA'B})$$

Bell measurement 

$$\overset{CNOT}{\mapsto} \frac{1}{\sqrt{2}} (a|000\rangle_{AA'B} + a|011\rangle_{AA'B} + b|110\rangle_{AA'B} + b|101\rangle_{AA'B})$$

$$= \frac{1}{\sqrt{2}} [a|0\rangle_A (|00\rangle_{A'B} + |11\rangle_{A'B}) + b|1\rangle_A (|10\rangle_{A'B} + |01\rangle_{A'B})]$$

H on $|0\rangle_A$ → $\frac{|0\rangle_A + |1\rangle_A}{\sqrt{2}}$; H on $|1\rangle_A$ → $\frac{|0\rangle_A - |1\rangle_A}{\sqrt{2}}$

Proof of Teleportation (2/2)

$$\frac{1}{\sqrt{2}} \left[a \frac{|0\rangle_A + |1\rangle_A}{\sqrt{2}} (|00\rangle_{A'B} + |11\rangle_{A'B}) + b \frac{|0\rangle_A - |1\rangle_A}{\sqrt{2}} (|10\rangle_{A'B} + |01\rangle_{A'B}) \right]$$

$$= \frac{1}{2} [|00\rangle (a|0\rangle + b|1\rangle) + |01\rangle (a|1\rangle + b|0\rangle) + |10\rangle (a|0\rangle - b|1\rangle) + |11\rangle (a|1\rangle - b|0\rangle)]$$

$$= \frac{1}{2} [|00\rangle|\psi\rangle + |01\rangle(X|\psi\rangle) + |10\rangle(Z|\psi\rangle) + |11\rangle(-Y|\psi\rangle)]$$

- Alice performs measurements on her 2-qubits to get '00', '01', '10', '11' equally likely

Measurement outcome	Post-measurement state
00	$\|00\rangle\|\psi\rangle$
01	$\|01\rangle(X\|\psi\rangle)$
10	$\|10\rangle(Z\|\psi\rangle)$
11	$\|11\rangle(XZ\|\psi\rangle)$

⇒ Once Alice received her bits, tell Bob how to correct the state

$$|ij\rangle(X^j Z^i |\psi\rangle) \overset{I \otimes Z^i X^j}{\mapsto} |ij\rangle \otimes |\psi\rangle$$

□

Remarks on Teleportation

- Each Bell measurement outcome occurs with 25%. → Alice learn *nothing* of $|\psi\rangle$.
- The qubit information is teleported over *any given distance* in a way that is *secure*. (though it does not teleport any physical object).
  → Eves learns nothing from the intercepted 2 cbits (so does Bob!).
- The system embodying $|\psi\rangle$ is not transferred from Alice to Bob.
  There is no any physical objects that was teleported!
  Only the "*information*" of the state's identity is transferred.
- Quantum teleportation does not violate the *no-cloning* or *no-signaling* theorems.
- It belongs to the research of *"entanglement-assisted quantum communication"*.
- Question: Can we consume less than 2 cbits & 1 ebit (asymptotic average cost)?
  → No! The quantum teleportation protocol is *optimal*.

The Intuition Behind the Scenes

$$|\Phi_{00}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \quad |\Phi_{01}\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}$$
$$|\Phi_{10}\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}} \quad |\Psi_{10}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$$

- Dense coding:
  Local operations can transform one of the Bell states to others in the Bell basis.

Choose $\{\sigma_m\}_m$ as Pauli matrices

$$(\sigma_m \otimes I)|\Phi_0\rangle = |\Phi_m\rangle$$
$$\Rightarrow \tfrac{1}{2}\langle \sigma_m, \sigma_{m'} \rangle = \langle \Phi_m | \Phi_{m'} \rangle$$

- Teleportation:

$$|\psi\rangle_A \otimes |\Phi^+\rangle_{A'B} = (a|0\rangle_A + b|1\rangle_A) \otimes \frac{1}{\sqrt{2}} (|00\rangle_{A'B} + |11\rangle_{A'B})$$
$$= \frac{1}{2} \left[|\Phi^+\rangle \otimes |\psi\rangle + |\Phi^-\rangle \otimes (X|\psi\rangle) + |\Psi^+\rangle \otimes (Z|\psi\rangle) + |\Psi^-\rangle \otimes \overset{XZ}{(XY|\psi\rangle)} \right]$$
$$= \textstyle\sum_m \frac{1}{2} |\Phi_m\rangle \otimes (\sigma_m |\psi\rangle)$$

The choice of the transform $|\psi\rangle \mapsto \sigma_m|\psi\rangle$ depends only on the measurement, not on the identity of $|\psi\rangle$

Optimality of The Protocols

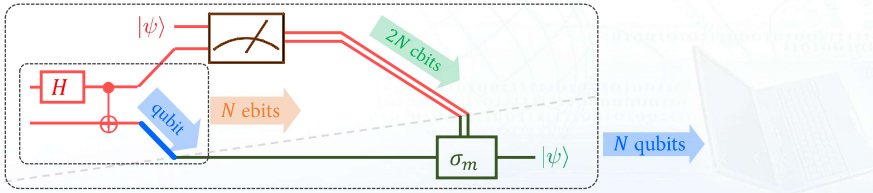
Optimality of Entanglement Distribution

- 1 qubit cannot generate more ebits: $1 \text{ qubit} \gneq N \text{ ebits } \forall N > 1$
- Assuming it was possible, use the generated N ebits for teleportation:

Assumed: $1 \text{ qubit} \geq N \text{ ebits } \exists N > 1$
N Teleportations: $2N \text{ cbits} + N \text{ ebits} \geq N \text{ qubits}$
$\Rightarrow 2N \text{ cbits} + 1 \text{ qubit} \geq N \text{ qubits}, N > 1$

The math is valid if the composition is

Contradicts with: $M \text{ cbits} + 1 \text{ qubit} \gneq N \text{ qubit } \forall N > 1, M \geq 1$ →←



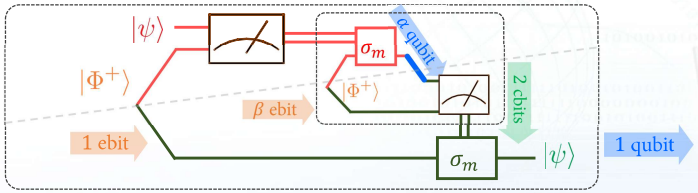
Optimality of Dense Coding (1/2)

- One cannot use less qubits and cbits (ebits): $\alpha \text{ qubit} + \beta \text{ ebit} \gneq 2 \text{ cbits } \forall \alpha < 1 \text{ or } \forall \beta < 1$

1. Assuming $\alpha < 1$, use the generated 2 cbits to supply the teleportation:

Assumed: $\alpha \text{ qubit} + \beta \text{ ebit} \geq 2 \text{ cbits } \exists \alpha < 1$
Teleportation: $2 \text{ cbits} + 1 \text{ ebit} \geq 1 \text{ qubit}$
$\Rightarrow \alpha \text{ qubit} + (1 + \beta) \text{ ebits} \geq 1 \text{ qubit}, \alpha < 1$

Contradicts with: $1 \text{ qubit} \gneq N \text{ qubits } \forall N > 1$ & $N \text{ ebits} \gneq 1 \text{ qubit } \forall N \geq 1$ →←



Optimality of Dense Coding (2/2)

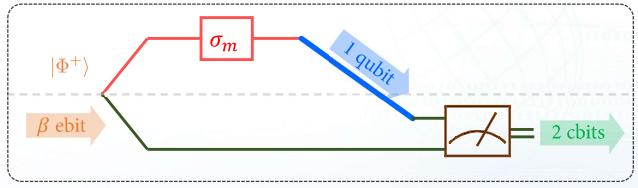
- One cannot use less qubits and cbits (ebits): $\alpha \text{ qubit} + \beta \text{ ebit} \gneq 2 \text{ cbits } \forall \alpha < 1 \text{ or } \forall \beta < 1$

1. Assuming $\alpha = 1$ and $\beta < 1$, relate it to the entanglement distribution:

Assumed: $1 \text{ qubit} + \beta \text{ ebit} \geq 2 \text{ cbits } \exists \beta < 1$
Entanglement distribution: $\beta \text{ qubit} \geq \beta \text{ ebit}$
$\Rightarrow (1 + \beta) \text{ qubit} \geq 2 \text{ cbits}, \beta < 1$

Contradicts with: $1 \text{ qubit} \gneq N \text{ cbits } \forall N > 1$ →←

Bob essentially gets 2 qubits



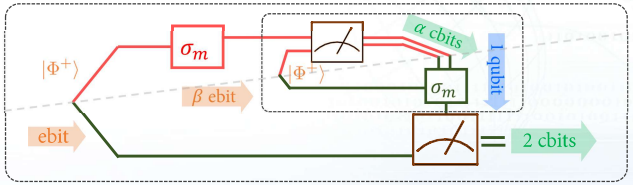
Optimality of Teleportation (1/2)

- One cannot use less qubits and cbits (cbits and ebits): α cbits $+ \beta$ ebit $\gtrneqq 1$ qubit $\forall \alpha < 2$ or $\forall \beta < 1$

1. Assuming $\alpha < 2$, use the generated 1 qubit to supply the dense coding:

Assumed: α cbits $+ \beta$ ebit ≥ 1 qubit $\exists \alpha < 2$
Dense coding: 1 qubit $+ 1$ ebit ≥ 2 cbits
$\Rightarrow \alpha$ cbits $+ (1 + \beta)$ ebit ≥ 2 cbits, $\alpha < 2$

Contradicts with: 1 cbit $\gtrneqq N$ cbits $\forall N > 1$ & N ebits $\gtrneqq 1$ cbit $\forall N \geq 1$ →←



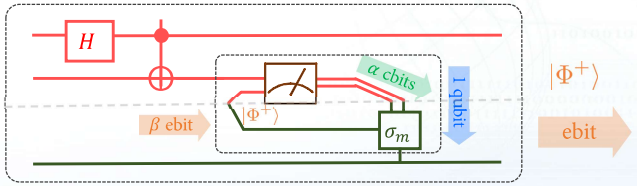
Optimality of Teleportation (2/2)

- One cannot use less qubits and cbits (cbits and ebits): α cbits $+ \beta$ ebit $\gtrneqq 1$ qubit $\forall \alpha < 2$ or $\forall \beta < 1$

2. Assuming $\beta < 1$, use the generated 1 qubit to supply entanglement distribution:

Assumed: α cbits $+ \beta$ ebit ≥ 1 qubit $\exists \beta < 1$
Entanglement distribution: 1 qubit ≥ 1 ebit
$\Rightarrow \alpha$ cbits $+ \beta$ ebit ≥ 1 ebit, $\beta < 1$

Contradicts with: 1 cbit $\gtrneqq N$ ebits $\forall N \geq 1$ & N ebits $\overset{1}{\gtrneqq} 1$ ebit $\overset{N}{\forall} N > 1$ →←



Quantum Key Distribution

Key Distribution

- The one-time pad protocol is *secure* only if the key distribution is secure and hidden from others, but how to do it in a *secure* way?
- Some public key crypto systems (such as RSA) relies on the computational hardness of the integer factorization.
- *Quantum key distribution* (QKD) provides a method for Alice and Bob to generate a shared secret key over public classical and quantum channels without the need to meet or to use a trusted intermediary party.
  Moreover, it is *provably secure* against eavesdropping.
  - BB84 (C. Bennett and G. Brassard 1984) uses four qubit non-orthogonal states;
  - B92 (C. Bennett 1992) uses only two non-orthogonal qubit states;
  - E91 (A. Ekert 1991) uses an entangled pair of qubits and the Bell theorem.
  - Etc. [Gisin et al., 2002] & [Pirandola, 2020]

Mutually Unbiased Bases

- *Mutually unbiased bases* (MUB): $\mathcal{B}_0 = \{|\psi_{00}\rangle, |\psi_{10}\rangle\}$ and $\mathcal{B}_1 = \{|\psi_{01}\rangle, |\psi_{11}\rangle\}$.

$$
\begin{aligned}
|\psi_{00}\rangle &= |0\rangle \\
|\psi_{10}\rangle &= |1\rangle \\
|\psi_{01}\rangle &= |+\rangle = \tfrac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\
|\psi_{11}\rangle &= |-\rangle = \tfrac{1}{\sqrt{2}}(|0\rangle - |1\rangle)
\end{aligned}
$$

- $\mathcal{B}_0 = \{|\psi_{00}\rangle, |\psi_{10}\rangle\}$ is the computational basis (or the Pauli Z eigenbasis);
  $\mathcal{B}_1 = \{|\psi_{01}\rangle, |\psi_{11}\rangle\}$ is the conjugate basis (or the Pauli X eigenbasis).

- These bases are called mutually unbiased if any state of one basis is measured in the other basis, the outcomes are always equally likely.

Protocol of BB84 (1/3)

1. Alice generates two uniformly random binary strings:
   $\boldsymbol{x} = x_1 x_2 \cdots x_m$ (x_i represents the bit value of key she's trying to send);
   $\boldsymbol{y} = y_1 y_2 \cdots y_m$ (y_i is her choice of the basis for that bit),
   and she prepares the m qubits in the states $|\psi_{x_1 y_1}\rangle |\psi_{x_2 y_2}\rangle \cdots |\psi_{x_m y_m}\rangle$ sending to Bob.

   Using such a random choice of MUB for encoding each bit value is sometimes called *conjugate coding*.

   Alice 沒有告訴Bob他用哪個basis，所以Bob得到的result是他猜的

2. Bob receives the m qubits but they may no longer be in the states $|\psi_{x_i y_i}\rangle$ due to the noise of the quantum channel or eavesdropping, but let's assume they are perfect.

   Bob chooses a uniformly random bit string $\boldsymbol{y'} = y'_1 y'_2 \cdots y'_m$ and measures the i^{th} received qubit in basis $\mathcal{B}_{y'_i}$ to get a result x'_i; let $\boldsymbol{x'} = x'_1 x'_2 \cdots x'_m$ be Bob's outcomes.

   If $y'_i = y_i$, then $x'_i = x_i$. Otherwise, x'_i is completely uncorrelated with x_i.

Protocol of BB84 (2/3)

3. Alice and Bob publicly reveal and compare their choice of bases, i.e. $\boldsymbol{y}$ and $\boldsymbol{y'}$ (but they do NOT reveal the strings $\boldsymbol{x}$ and $\boldsymbol{x'}$).

   They discard all bits x_i and x'_i for which $y'_i \neq y_i$ leaving shorter strings of expected length $m/2$. Call these strings $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{x}}'$.

   Under our assumptions of no noise and no eavesdropping in the quantum channel, we would have the perfectly correlated bits $\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{x}}'$ (as the key).

- The 'quantum' part has done. In reality there always be noise and eavesdropping in the transmission. To address these issues, the BB84 protocol concludes with the following Steps 4 & 5 from techniques in classical cryptography.
- The key distribution (Steps 1~3) can be done easily without quantum tricks.

  → What's matter is the security of the protocol (i.e. Eavesdropper isn't be recognized).

Protocol of BB84 (3/3)

4. (**Information reconciliation**) Alice and Bob publicly compare a random sample of their strings (say half of them) to estimate the *bit error rate* (BER), the proportion of bits in $\tilde{\boldsymbol{x}}'$ that are not equal to those in $\tilde{\boldsymbol{x}}$, and discard all the announced bits.

   Assume the remaining bits have the same proportion of errors as those checked. They can correct these remaining errors (albeit at unknown positions) to obtain two strings that agree in a high percentage of positions with high probability (at the expense of sacrificing some more bits) if the BER is not too large. Otherwise, abort

5. (**Privacy amplification**) From the estimated BER, Alice and Bob can estimate the maximum amount of information that an eavesdropper is likely to have obtained about the remaining bits.
   They can replace their strings by even shorter strings about which the eavesdropper has no knowledge of (with high probability).

An Example

Alice's bit string x	1	0	1	1	0	1	0	1
Alice's basis string y	1	0	0	1	0	1	1	0
X: |+>,|–> Z: |0>, |1>	X	Z	Z	X	Z	X	X	Z
Qubit states	$\|-\rangle$	$\|0\rangle$	$\|1\rangle$	$\|-\rangle$	$\|0\rangle$	$\|-\rangle$	$\|+\rangle$	$\|1\rangle$
Bob's basis string y'	1	1	0	0	1	0	1	0
	X	X	Z	Z	X	Z	X	Z
Bob's resulting states	$\|-\rangle$	$\|+\rangle$	$\|1\rangle$	$\|1\rangle$	$\|-\rangle$	$\|0\rangle$	$\|+\rangle$	$\|1\rangle$
Bob's resulting bits x'	1	0	1	1	1	0	0	1
Right basis?	Y	N	Y	N	N	N	Y	Y
Key string $\tilde{x} = \tilde{x'}$	1		1				0	1

About Eavesdropper's Attacks (1/2)

- Goal of the Eavesdropper: To learn the key without being recognized.
  Note that if classical systems are measured, then any introduced disturbance can be undone by Eve in theory; thus her presence cannot be detected.
- **The Intercept-resend attack**: Eve can intercept each transmitted qubit separately, measure it in some chosen basis to acquire some information, and then send on the post-measurement state to Bob.
- **General coherent attack**: Eve can introduce an auxiliary (possibly very large) probe quantum system E of her own and unitarily interact E with many of the passing qubits. She measure E to acquire information, which now can be joint information about many qubits. Her measurement here can even be postponed until she overhears Alice & Bob's public discussions in Steps 2~5, and did what she likes.

About Eavesdropper's Attack (2/2)

- Recall that standard classical strategy for eavesdropping on classical bits (reading them and retaining a copy, and then sending them on perfectly intact) is not available in the quantum protocol because of the no-cloning theorem and the use of non-orthogonal states in the set of encoding states.
- In Step 5, the BER provides an upper bound on the amount of information that Eve can have gained because non-orthogonal states cannot be perfectly distinguished (i.e. information disturbance trade-off).
- Hence, the privacy amplification techniques from classical cryptography can be shown to provide *information-theoretic security* against any possible eavesdropping strategy obeying quantum mechanics.

[Gisin et al., 2002], [Mayers, 2001], [Shore-Preskill, 2000], [Pirandola, 2020]

An Intercept-Resend Attack

- Assume that the quantum channel is noiseless but Eve intercepts each passing qubit and measures it in the so-called *Breidbart basis*:

$$|\alpha_0\rangle = \cos\tfrac{\pi}{8}|0\rangle + \sin\tfrac{\pi}{8}|1\rangle$$
$$|\alpha_1\rangle = -\sin\tfrac{\pi}{8}|0\rangle + \cos\tfrac{\pi}{8}|1\rangle$$

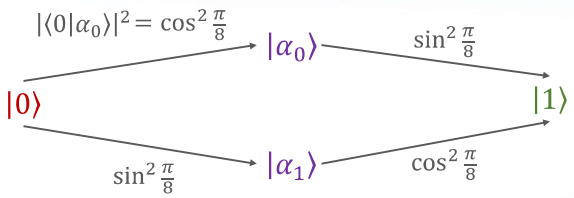
- Of course Eve can choose either $\mathcal{B}_0$ or $\mathcal{B}_1$ as before but the Breidbart basis lies "midway" between the above two bases.
  → The eavesdropping will result in a disturbance amounting to BER = 25%.
  Eve will learn each bit of $\tilde{x}$ with probability $\cos^2\frac{\pi}{8} \approx 0.85$.
- If spotting any difference in Step 4, they just abort the protocol since they have noticed Eve's presence. Then they run the protocol again.

→ The probability of not noticing Eve will thus be exponentially small.

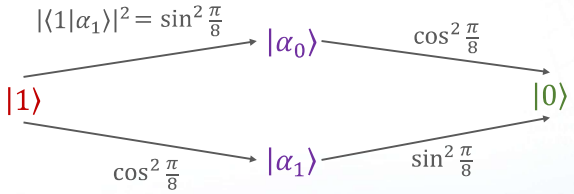
Analysis for The BER (1/2)

$$|\alpha_0\rangle = \cos\tfrac{\pi}{8}|0\rangle + \sin\tfrac{\pi}{8}|1\rangle$$
$$|\alpha_1\rangle = -\sin\tfrac{\pi}{8}|0\rangle + \cos\tfrac{\pi}{8}|1\rangle$$

Alice sent — Eve's outcome — Bob's outcome



$$\Pr(\text{error}) = 2\cos^2\tfrac{\pi}{8}\sin^2\tfrac{\pi}{8} = 1/4$$
$$\Pr(\text{Eve learns}) = |\langle 0|\alpha_0\rangle|^2 = \cos^2\tfrac{\pi}{8}$$

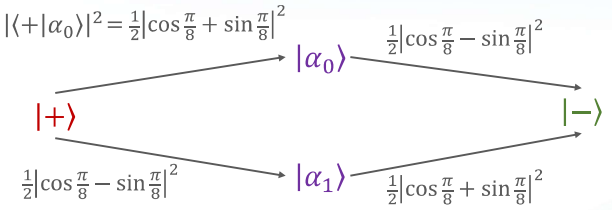


$$\Pr(\text{error}) = 2\cos^2\tfrac{\pi}{8}\sin^2\tfrac{\pi}{8} = 1/4$$
$$\Pr(\text{Eve learns}) = |\langle 1|\alpha_1\rangle|^2 = \cos^2\tfrac{\pi}{8}$$

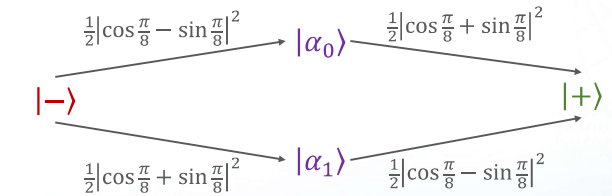
Analysis for The BER (2/2)

$$|\alpha_0\rangle = \cos\tfrac{\pi}{8}|0\rangle + \sin\tfrac{\pi}{8}|1\rangle$$
$$|\alpha_1\rangle = -\sin\tfrac{\pi}{8}|0\rangle + \cos\tfrac{\pi}{8}|1\rangle$$

Alice sent — Eve's outcome — Bob's outcome



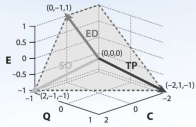
$$\Pr(\text{error}) = \tfrac{1}{2}|\cos\tfrac{\pi}{8} + \sin\tfrac{\pi}{8}|^2|\cos\tfrac{\pi}{8} - \sin\tfrac{\pi}{8}|^2 = 1/4$$
$$\Pr(\text{Eve learns}) = |\langle +|\alpha_0\rangle|^2 = \tfrac{1}{2}|\cos\tfrac{\pi}{8} + \sin\tfrac{\pi}{8}|^2 = \cos^2\tfrac{\pi}{8}$$



$$\Pr(\text{error}) = \tfrac{1}{2}|\cos\tfrac{\pi}{8} + \sin\tfrac{\pi}{8}|^2|\cos\tfrac{\pi}{8} - \sin\tfrac{\pi}{8}|^2 = 1/4$$
$$\Pr(\text{Eve learns}) = |\langle -|\alpha_1\rangle|^2 = \tfrac{1}{2}|\cos\tfrac{\pi}{8} + \sin\tfrac{\pi}{8}|^2 = \cos^2\tfrac{\pi}{8}$$

Concluding Remarks & References

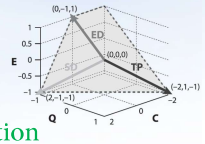
Resource Inequalities



- 1 qubit $\geq$ 1 cbit
- 1 qubit $\gneq N$ cbits $\forall N > 1$
- 1 qubit $\geq$ 1 ebit
- N cbits $\gneq$ 1 qubit $\forall N \geq 1$
- N ebits $\gneq$ 1 cbit $\forall N \geq 1$
- 1 qubit $\gneq N$ ebits $\forall N > 1$
- N ebits $\gneq$ 1 qubit $\forall N \geq 1$
- 1 cbit $\gneq N$ cbits $\forall N > 1$
- M cbits + 1 qubit $\gneq N$ qubit $\forall N > 1, M \geq 1$
- 1 ebit $\gneq N$ ebits $\forall N > 1$
- N cbit $\gneq$ 1 ebit $\forall N \geq 1$
- 1 qubit + 1 ebit $\geq$ 2 cbits
- α qubit + β ebit $\gneq$ 2 cbits $\forall \alpha < 1$ or $\forall \beta < 1$
- 2 cbits + 1 ebit $\geq$ 1 qubit
- α cbits + β ebit $\gneq$ 1 qubit $\forall \alpha < 2$ or $\forall \beta < 1$

etc…

Concluding Remarks



- One-time pad: public classical channel + [cc] = private classical communication
- Dense coding: public quantum channel + [qq] = private classical communication
- Teleportation: public classical channel + [qq] = private quantum communication
- Beyond qubits: Provided a d-dimensional maximally entangled state $|\Phi_d\rangle \coloneqq \frac{1}{\sqrt{d}}\sum_i |ii\rangle$.
  - A total of d^2 messages ($2 \log d$ bits) can be sent via a d-dimensional quantum system and $|\Phi_d\rangle$.
  - As for teleportation, $2 \log d$ bits + $\log d$ ebits $\geq \log d$ ebits.

logd qbits + logd ebits >= 2logd cbits

- If entanglement is free, then 1 qubit is worth exactly two cbits.

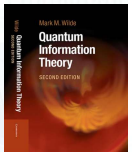
  Indeed, 1 qubit + 1 ebit = 2 cobits (*coherent communication* by Harrow, 2004).
- The basic quantum protocols are building blocks for large information-processing systems.
- *Entanglement distillation protocol* [Werner, 2001] & [Plenio-Virmani, 2007], [Horodeck et al., 2009]
- Extensions for Quantum Shannon Theory and capacity region for qubit/ebit/cbit.
  [Chapter 8, Wilde, 2017]

References (1/4)

- A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problems of Information Transmission*, 9:177-183, 1973.
- C. H. Bennett, S. J. Wiesner, "Communication via one- and two-particle operators on Einstein–Podolsky–Rosen states," *Physical Review Letters*, 69:2881-2884, 1992.
- K. Mattle, H. Weinfurther, P. G. Kwiat, and A. Zeilinger, "Dense coding in experimental quantum communication," *Physical Review Letters*, 76:4656-4659, 1996.
- C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels," *Physical Review Letters*, 70:1895-1899, 1993.
- D. Boschi, S. Branca, F. De Martini, L. Hardy, and S. Popescu, "Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein–Podolsky–Rosen channels," *Physical Review Letters*, 80:1121-1125, 1998.

References (2/4)

- R. F. Werner, "All teleportation and dense coding schemes," *Journal of Physics A*, 34:7081, 2001.
- M. B. Plenio and S. Virmani, "An introduction to entanglement measures. *Quantum Information and Computation*," 7(1), 2007.
- R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Review of Modern Physics*, 81, 865, 2009.
- C. H. Bennett and G. Brassard. "Quantum cryptography: Public key distribution and coin tossing". In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, 175:8, 1984.
- D. Mayers, "Unconditional security in quantum cryptography," *Journal of the ACM*, 48:351, 2001.
- P. W. Shor and J. Preskill, "Simple proof of security of the BB84 quantum key distribution protocol," *Physical Review Letters*, 85:441–444, 2000.
- M. Wilde, Quantum information theory. Cambridge Press, 2017.



References (3/4)

- C. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing", *Theoretical Computer Science. Theoretical Aspects of Quantum Cryptography* – celebrating 30 years of BB84, 560, Part 1: 7–11, 2014.
- V. Scarani, "The security of practical quantum key distribution," *Review of Modern Physics*, 81(3): 1301–1350, 2009.
- C. H. Bennett, "Quantum cryptography using any two nonorthogonal states," *Physical Review Letters*, 68, 3121, 1992.
- A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Physical Review Letters*, 67, 661, 1991.
- N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," *Review of Modern Physics*, 74: 145-195, 2002.
- S. Pirandola *et al.*, "Advances in quantum cryptography," *Advances in Optics and Photonics*, 12(4):1012–1236, 2020.
- A. Harrow, "Coherent Communication of Classical Messages," Physical Review Letters, 92, 097902, 2004.

References (4/4)

- C. H. Bennett, P. Hayden, D. Leung, P. W. Shor, A. Winter, "Remote preparation of quantum states," *IEEE Transactions on Information Theory*, 51:56–74, 2005.
- I. Devetak, A. Harrow, A. Winter. "A resource framework for quantum Shannon theory," *IEEE Transactions on Information Theory*, 54(10), 4587–4618, 2008.
- A. Harrow, P. Hayden, D. Leung, "Superdense coding of quantum states," *Physical Review Letters*, 92:187901, 2004.
- I. Devetak, J. Yard, "Exact cost of redistributing multipartite quantum states," *Physical Review Letters* 100:230501, 2008.
- Terhal B.M. (2016) Quantum Dense Coding. In: Kao MY. (eds) Encyclopedia of Algorithms. Springer, New York, NY.
- Anshu A., Devabathini V.K., Jain R., Mukhopadhyay P. (2016) Teleportation of Quantum States. In: Kao MY. (eds) Encyclopedia of Algorithms. Springer, New York, NY.
- Renner R. (2016) Quantum Key Distribution. In: Kao MY. (eds) Encyclopedia of Algorithms. Springer, New York, NY.

Quantum Information and Computation
The Quantum Circuit Model

Hao-Chung Cheng (鄭皓中)

haochung@ntu.edu.tw

Department of Electrical Engineering

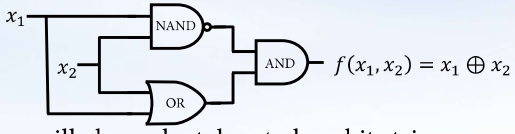
National Taiwan University

March 17, 2021

Outline

Prologue

What is Computation?



- *Computational task:* The input to a computation will always be taken to be a bit string. The input size is the number of bits in the bit string.
  → e.g. "given an n-bit string A (for any n), is A prime?"
- *Circuit model:* For each n, the computation with inputs of size n begins with the input string $x = b_1 \dots b_n$ extended with an extra 0's. These latter bits provide "extra working space (called *ancilla*)" that may by needed in computation.
- A Boolean circuit C_n is a collection of logic gates (e.g. AND, OR, NOT) and wires (registers) that performs a mapping from Boolean inputs to Boolean outputs.
- A circuit *computes* some Boolean function $f: \{0,1\}^n \to \{0,1\}^m$ if the output gets the correct value $f(x)$ for every input $x \in \mathbb{Z}_2^n := \{0,1\}^n$.
- Classical digital computers (under the Boolean circuit model) uses binary digit (*bits*, 0s or 1s) to store, transfer, manipulate data by executing algorithms.
- An *algorithm* is a well-defined procedure, with finite description, for solving an information-processing tasks. → It's a *hardware-independent* method.

Randomized Classical Computation

- It is useful to extend the model of classical (deterministic) computation to incorporate classical probabilistic choices. → *Randomized circuits*.
- For input bits $b_1 \dots b_n$, we extend the starting string $b_1 \dots b_n 0 \dots 0$ to $b_1 \dots b_n r_1 \dots r_k 0 \dots 0$, where $r_1 \dots r_k$ is a sequence of random bits (e.g. coin flips).
- The output is now a sample from a probability distribution over all possible output strings depending on the random choice of $r_1 \dots r_k$.
- A randomized circuit (or algorithm) computes a function f if it successfully outputs the correct answer $f(x)$ with probability at least $2/3$ for *every* inputs x (probability taken over the random bits $r_1 \dots r_k$).
- The randomized algorithm may give wrong answers (with prob. $\leq 1/3$), but we can independently run the algorithm n times and output the majority among the output answers. The probability that this majority is wrong is at most $1/3^n \leq \varepsilon$.

→ We output the correct answer with probability $\geq 1 - \varepsilon$.

Ancillias: 0 ... 0

The worst case scenario

With arbitrary high precision

Quantum Computation Model

- For inputs of size n, the starting string $b_1 \dots b_n 0 \dots 0$ is replaced by a sequence of qubits in the corresponding computational basis state $|b_1\rangle \dots |b_n\rangle |0\rangle \dots |0\rangle$.
- A *quantum circuit* C_n (quantum gate array) replaces the AND, OR, and NOT gates by *elementary quantum gates*, which are unitary operation on (at most 3) qubits.
- The output of the computation is the result of performing a quantum measurement (in the computational basis) on a specified subset of the qubits.
- A *quantum computation* or *quantum algorithm* is defined by a family of quantum circuits $(C_1, C_2, \dots)$.
- A *quantum computer* is a device that leverages quantum mechanics representation of information (qubits) to perform computation.
- In architecture design, it is referred to as the *quantum processing unit* (QPU).

Elementary Quantum Gates

- Pauli gates $X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $Y := \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}$ $Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
- Hadamard gates

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

- Rotation gates

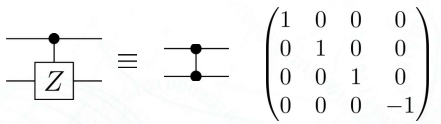
$$R_x(\phi) := \mathrm{e}^{-\mathrm{i}\frac{\phi}{2}X} = \begin{pmatrix} \cos(\frac{\phi}{2}) & -\mathrm{i}\sin(\frac{\phi}{2}) \\ -\mathrm{i}\sin(\frac{\phi}{2}) & \cos(\frac{\phi}{2}) \end{pmatrix}$$

$$R_y(\phi) := \mathrm{e}^{-\mathrm{i}\frac{\phi}{2}Y} = \begin{pmatrix} \cos(\frac{\phi}{2}) & -\sin(\frac{\phi}{2}) \\ \sin(\frac{\phi}{2}) & \cos(\frac{\phi}{2}) \end{pmatrix}$$

Phase gate

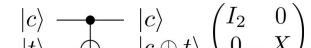
$$R_z(\phi) := \mathrm{e}^{-\mathrm{i}\frac{\phi}{2}Z} = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\phi} \end{pmatrix} \quad S := R_z(\tfrac{\pi}{2}) \quad T := R_z(\tfrac{\pi}{4})$$

- Controlled-Z gates



$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

- CNOT gate

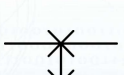


$$\begin{pmatrix} I_2 & 0 \\ 0 & X \end{pmatrix}$$

- Toffoli (CCNOT) gate

$$\begin{pmatrix} I_3 & 0 & 0 \\ 0 & I_3 & 0 \\ 0 & 0 & X \end{pmatrix}$$

- Swap gate $|\psi\rangle|\phi\rangle \mapsto |\phi\rangle|\psi\rangle$



$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Universal Quantum Gate Sets [§4.5, N&C]

- Definition. A set of quantum gates is *universal* if for any integer $n \geq 1$, any n-qubit unitary operator can be approximated to arbitrary accuracy (in terms of the *operator norm*) by using only gates from that set.
- A set composed of any 2-qubit entangled gate with all 1-qubit gates is universal.
  → Any $d \times d$ unitary matrix U can be written as $U = V_1 \cdots V_k$, $k \leq d(d-1)/2$.
  → Such gate sets implement any n-qubit unitary exactly but its size is infinite.
- (1) All 1-qubit gates and the CNOT gate are universal.
- (2) The set $\{H, T\}$ is universal for 1-qubit gates; hence $\{\text{CNOT}, H, T\}$ is a universal set.
- (3) The set $\{H, \text{CCNOT}\}$ is universal for all unitaries with real entries.
- The Solovay–Kitaev theorem: We can approximate any 1-qubit gates up to error ε using a number of gates from the universal set that is only polylog$(1/\varepsilon)$.

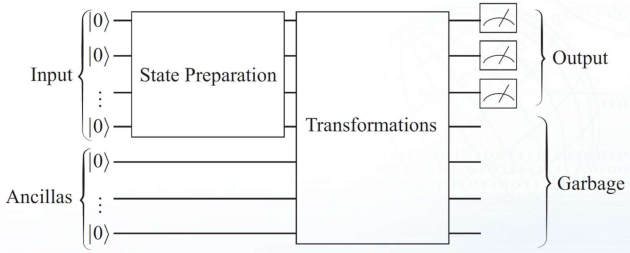
V_k's are 2×2 unitaries

Classical vs. Quantum Computation

- Any quantum operation in a closed quantum system is reversible, i.e. $\forall$ state vector $|\psi\rangle$ and unitary U, $\exists$ a quantum operation $U^{-1} = U^\dagger$ such that $U^{-1}U|\psi\rangle = |\psi\rangle$.
- In classical computation, certain gates (e.g. AND and OR) are not reversible.
- Any irreversible classical computation can be efficiently simulated by a reversible classical computation:
  Given a (probably irreversible) Boolean function $f: \mathbb{Z}_2^n \to \mathbb{Z}_2^m$, construct a reversible form: $\tilde{f}: \mathbb{Z}_2^{n+m} \to \mathbb{Z}_2^{n+m}$ by $\tilde{f}(x, y) \coloneqq \big(x, y \oplus f(x)\big), \forall x \in \mathbb{Z}_2^n, y \in \mathbb{Z}_2^m$.
  → The function $\tilde{f}$ is self-inverse: $\tilde{f}\big(\tilde{f}(x, y)\big) = (x, y)$ since $b \oplus b = 0 \cdots 0 \; \forall b \in \mathbb{Z}_2^*$.
- Any Boolean function f that is efficiently computable by a classical circuit is also efficiently computable by a quantum circuit using $|x, y\rangle \mapsto |x, y \oplus f(x)\rangle$.
- Toffoli gate is universal for classical computation, → replace it by its quantum part.

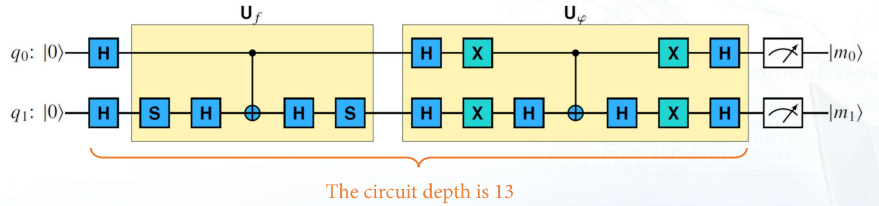
Circuit Implementation of Quantum Algorithms

- Elementary quantum gates can be composed into bigger unitary operations by taking *tensor products* (if gates are applied in parallel to different parts of the register), and ordinary *matrix products* (if gates are applied sequentially).
- A prescription commonly observed in quantum algorithms is:
  - efficiently encode information into a small number of qubits;
  - cleverly build up *entanglement* and *interference* during the algorithm;
  - design a final measurement that yields desired outcomes with high probability.



Computational Resources

- We are interested in how various kinds of *computational resources* (principally *time* – number of steps, or *space* – amount of memory needed) grow as a function of input size n.
  - *Time complexity.* The time complexity of a unitary transformation U is related to the number of gates of the smallest circuit (i.e. the *size* of the circuit) that implements U.
  - *Query complexity.* Query complexity is the number of times an algorithm needs to query a given black-box function (often called an *oracle*) to solve a problem.
  - *Depth.* The depth of a circuit is the maximum of the depths of its wires.



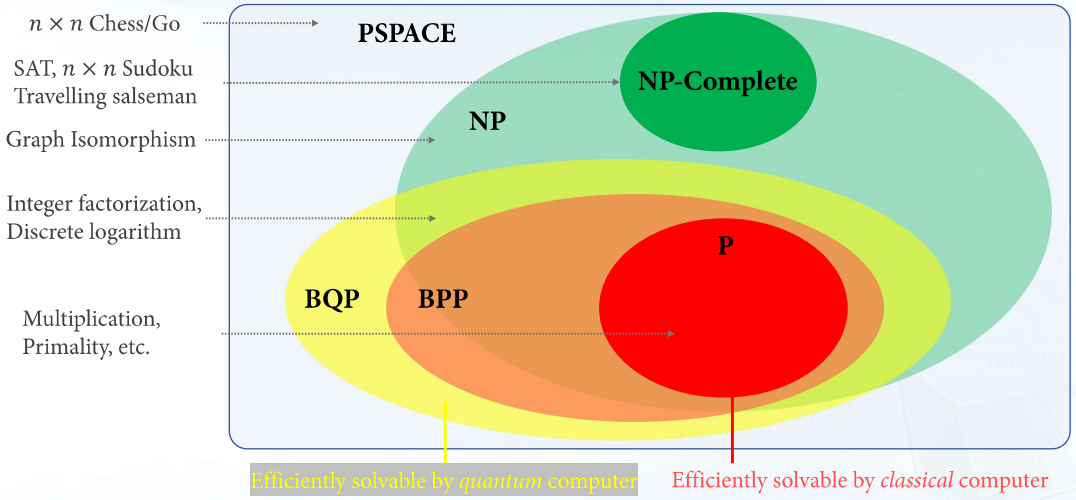
The Big-O Notation for Asymptotic Analysis

- $O(f(n))$ denotes the set of functions g for which $\exists c > 0, N \in \mathbb{N}$ so that $g(n) \leq cf(n)$.
  $\rightarrow T(n) = O(f(n))$ means T grows no faster than f.
  $\rightarrow T(n) = O(\text{poly}(n))$ means $T = O(n^k)$ for some constant k.
  $\rightarrow T(n) \neq O(\text{poly}(n))$, we say it is *superpolynomial*; e.g. $e^n, 2^{\sqrt{n}}, n^{\log n}$.
- $\Omega(f(n))$ denotes the set of functions g for which $\exists c > 0, N \in \mathbb{N}$ so that $g(n) \geq cf(n)$.
  $\rightarrow T(n) = O(f(n))$ means T grows at least as fast as f.
- $\Theta(f(n))$ denotes the set of functions g that are both in $O(f(n))$ and in $\Omega(f(n))$.
- $g(n) = o(f(n))$ means $\limsup_{n\to\infty} \frac{f(n)}{g(n)} = 0$.
- Poly-time computations are regarded as "tractable" in practice; otherwise "intractable". The term *efficient algorithm* is also synonymous with poly-time algorithm.

Complexity Classes

- How many steps (in the worst case) does the algorithm require for any input of size n? → We analyse the algorithm especially when n is large, i.e. *asymptotic analysis*.
- Most problems of interest can be reformulated as *decision problems* (where the answer is "yes/no") in a very natural way without losing their intrinsic complexity.
- **P** (polynomial time): problems that can be solved in polynomial time by a classical algorithm and it gives the correct answer with certainty.
- **BPP** (bounded-error probabilistic poly time): problems that there exists a poly-time randomized algorithm that gives the correct answer with prob. at least $2/3$.
- **NP** (non-deterministic poly time): problems whenever the answer is "yes", there is a short proof that the verifier can deterministically verify it by in poly time.
- **BQP** (bounded-error quantum poly time): quantum analogue of **BPP**.

Relations – A Glimpse of The [Complexity Zoo]



The Oracle Model

The Oracle Model (Black Box Promise Problems)

- A query model involves a *black-box function* called an *oracle* that computes some function $f: \{0,1\}^n \to \{0,1\}^m$.
  → We can query the black box by giving it inputs and this is the only access we have to the function and its values (we don't know its interior workings).
- An algorithm is trying to learn some properties of f by evaluating f on a set of inputs and analyzing the outputs but not examining how f is implemented internally.
- There is often an a priori *promise* on f, i.e. *a priori* restriction on the possible form of f.
- The number of queries required by the algorithm is called the *query complexity* of the algorithm.

Examples of The Oracle Model

The Search Problem

- ***Given***: A black box for a Boolean function $f: \{0,1\}^n \to \{0,1\}$.
- ***Promise***: There is a unique x such that $f(x) = 1$.
- ***Problem***: Find this special x.

Periodicity

- ***Given***: A black box for a function $f: \mathbb{Z}_n \to \mathbb{Z}_n$.
- ***Promise***: f is periodic i.e. there is a least p such that $f(x+p) = f(x)$ for all x.
- ***Problem***: Find the period p.

The Quantum Oracle

- The *quantum oracle* for any Boolean function $f: \{0,1\}^n \to \{0,1\}^m$ is given by the quantum gate U_f:

$$U_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle, \forall x \in \{0,1\}^n, y \in \{0,1\}^m$$

→ Again, applying the oracle twice gives the original states. Hence it is indeed a *unitary* operation.
→ If restricted to computational basis, U_f acts just like the *reversible* version of any Boolean function.

Quantum Parallelism

- We set the input register to an equal superposition of all 2^n possible n-bit strings:

$$U_f : \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle|0\rangle \mapsto \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle|f(x)\rangle$$

- In *one* run of the protocol, we obtain a final state with depends on all of the function values. On the other hand, we would need *exponentially* many queries to have full access to the function f.
  → An important feature is that we have created a superposition of *exponentially many* terms with only a *linear* number of the Hadamard gates:

$$H^{\otimes n}|0\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}}(|0\rangle + |1\rangle) \otimes \cdots \otimes (|0\rangle + |1\rangle) = \frac{1}{\sqrt{2^n}} \sum_{x_1,\ldots,x_n=0}^{1} |x_1\rangle \cdots |x_n\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle$$

Why Using The Query Model?

- The query complexity provides a *lower bound* on the number of steps (gates) required to solve a given problem.
  → It is useful for ruling out fast quantum algorithms.
- We can give both a quantum computer and a classical computer access to the same oracle and see which performs better.
  → The query model can be used to prove fast quantum algorithm *relative* to the oracle.
- If we find an *efficient* algorithm for a problem in query complexity,
  → then if we are given an *explicit* circuit realizing the black-box transformation, we will have an *efficient* algorithm for an explicit computational task.
  (E.g. Shor's algorithm for integer factorization.)

The Deutsch–Jozsa Algorithm

The Deutsch–Jozsa Algorithm (1/2)

The "Balanced versus Constant" Problem

– ***Given***: A black box for a Boolean function $f: \{0,1\}^n \to \{0,1\}$.

– ***Promise***: The function f is *either* a (i) constant function, or (ii) a balanced function, i.e. $f(x) = 0$ or 1 for exactly half of the inputs.

– ***Problem***: Determine (*with certainty*) whether f is balanced or constant.

- Ex:

x	f_0	f_1	f_x	$f_{\bar{x}}$
0	0	1	0	1
1	0	1	1	0

Constant Balanced

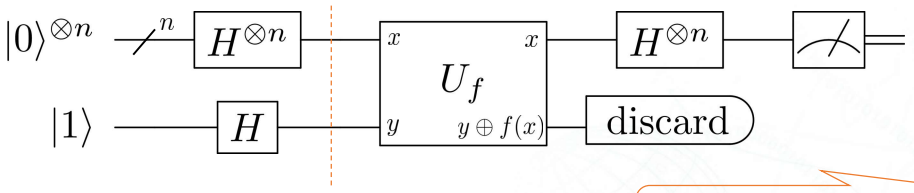
D. Deutsch, R. Jozsa, "Rapid Solution of Problems by Quantum Computation," *Proc. Royal Society A*, 439:553-558, 1992.

The Deutsch–Jozsa Algorithm (2/2)

- Classical solutions:
  There are 2^n inputs in total. So in the worst case, we have to try more than half of them, i.e. at least $2^n/2 + 1$ queries, to determine the function *for sure*.
  → It's a sufficient and necessary condition even for *randomized* classical algorithms.
- A quantum solution (by David Deutsch and Richard Jozsa in 1992):
  Just *one* query to the oracle suffices (with $O(n)$ extra processing steps).
  (The quantum oracle is $U_f: |x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$.)

Quantum Advantage (in terms of the query complexity)!

Circuit Diagram of the DJ Algorithm



$\left(\frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle\right)\otimes|-\rangle$

See if the measurement outcome is all zeros or not

Proof of the DJ Algorithm (1/2)

1. Create equal superposition for all n-bit strings in the first n-qubit register and set the last register to be the state $|-\rangle$.

$$|0\rangle^{\otimes n}\otimes|1\rangle \overset{H^{\otimes n+1}}{\mapsto} \left(\frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle\right)\otimes|-\rangle$$

2. Apply the quantum oracle $U_f\colon |x\rangle|y\rangle = |x\rangle|y\oplus f(x)\rangle$ on the above state.

For each x:
$$U_f : |x\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) \mapsto |x\rangle\left(\frac{|f(x)\rangle-|1\oplus f(x)\rangle}{\sqrt{2}}\right)$$
$$= \begin{cases} |x\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) & \text{if } f(x)=0 \\ -|x\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) & \text{if } f(x)=1 \end{cases}$$
$$= (-1)^{f(x)}|x\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$$

Phase kickback

Proof of the DJ Algorithm (2/2)

The resulting state (for superposition): $=: |f\rangle$

$$U_f : \left(\frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle\right)|-\rangle \mapsto \left(\frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}(-1)^{f(x)}|x\rangle\right)|-\rangle$$

3. If f is constant: $|f\rangle = \pm\frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle \overset{H^{\otimes n}}{\mapsto} \pm|0\rangle^{\otimes n}$

will get all zero outcomes *certainly*

If f is balanced: the sum in $|f\rangle$ contains an equal number of plus and minus terms

$\rightarrow \langle f|\left(\frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle\right) = 0$

will get non-zero strings *certainly*

$\rightarrow \langle f|H^{\otimes n}|0\rangle^{\otimes n} = 0 \rightarrow$ $H^{\otimes n}|f\rangle$ is orthogonal to $|0\rangle^{\otimes n}$ $\rightarrow H^{\otimes n}|f\rangle = \sum_{x\neq 0\ldots 0} a_x|x\rangle$

Remarks on the DJ Algorithm

- The problem is solved with *one* query (*exponential speedup*) to the quantum oracle, $(2n+1)$ Hadamard gates, one X gate, and n single-qubit measurements.
- The *exactly-zero* scenario in computation is an unrealistic idealization and for *realistic* computation we should always accept some (suitably small) level of error.
  → There exists a classical (randomized) algorithm that solves the problem with probability $> 1-\epsilon$, with only a *constant* number of queries, $O(1/\log\epsilon)$.
  → Query K times independently, output all 0/1 with probability $2/2^K < \varepsilon$.
- The *Bernstein–Vazirani algorithm*: *polynomial speedup* over bounded-error randomized algorithms.
- The *Simon's algorithm*: exponential speedup over bounded-error randomized algo.
- An totally algebraic approach: $H^{\otimes n}|x\rangle = \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}(-1)^{x\cdot y}|y\rangle$.

$x\cdot y$ means bitwise inner product modulo 2

Intuition behind the DJ Algorithm (1/2)

Interference & Entanglement

- Key ingredients: *Quantum Parallelism* & the *phase kickback trick*
  → The information we need is encoded into the relative phase.

- In the computational basis, the CNOT gate appears to do nothing to the control bit.
  → In fact, it affects the control quit just as much as it does the target qubit!

CNOT: $|1\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) \mapsto |1\rangle\left(X\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)\right) = -|1\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$

$|-\rangle$ is an eigenstate of the X gate with eigenvalue -1

CNOT: $|0\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) \mapsto |0\rangle\left(I\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)\right) = |0\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$

$|-\rangle$ is an eigenstate of the identity gate with eigenvalue 1

⇒ *Phase kickback* via CNOT: $|b\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) \mapsto (-1)^b|b\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right), \forall b \in \{0,1\}$

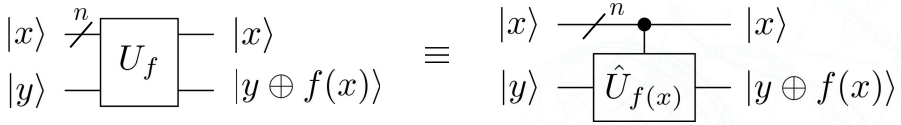
CNOT: $(\alpha_0|0\rangle + \alpha_1|1\rangle)\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) \mapsto (\alpha_0|0\rangle - \alpha_1|1\rangle)\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$

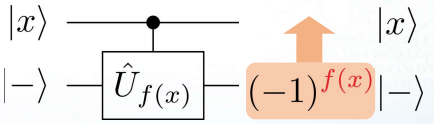
Relative phase between $|0\rangle/|1\rangle$

Intuition behind the DJ Algorithm (2/2)

- Replace the Controlled-NOT gate to C-$\hat{U}_{f(x)}$, where $\hat{U}_{f(x)}: |b\rangle \mapsto |b \oplus f(x)\rangle$.
  → $|-\rangle$ is the eigenstate of target register $\hat{U}_{f(x)}$;
  Its eigenvalue $(-1)^{f(x)}$ can be *kicked back* in front of the control register.

Phase kickback via U_f: $|x\rangle|-\rangle \mapsto (-1)^{f(x)}|x\rangle|-\rangle, \forall x \in \{0,1\}^n$





Concluding Remarks (1/2)

- *Deterministic* classical computation → a deterministic classical theory of physics.
  Probabilistic classical computation → adding the possibility of a randomness.
  Quantum computation → working in a quantum mechanical framework.
- In a fully quantum algorithm, we would not measure the state immediately after the first step. This way the quantum probability amplitudes will have a chance to interfere.
  → *Principle of deferred measurement*: move measurements to the end of the circuit.
- We do not know if NP ⊂ BQP yet (except for the oracle problems), but we aim to find classically hard but quantumly easy tasks.
- The advantage (if any) of quantum algorithms as compared to classical algorithms is based on the use of superpositions and entanglement for specific problems.
  - Searching for *global properties* of a function (e.g. balanced or not, period): we don't compute functional values and then compare them, but rather investigate the correlations between the states.
  - Amplitude amplification (see next lecture).

Concluding Remarks (2/2)

- The Deutsch–Jozsa Algorithm shows a quantum speedup in the oracle problems.
- But how to construct the quantum oracle?
  → The black box's interior workings are inaccessible to us. It's unlike standard computational task with a bit string as input and no hidden ingredients.
  - For query complexity: don't care too much since we assume the black-box is given.
  - For implementation: (i) *quantum circuit synthesis*
    (ii) translate the classical circuit (if any) into the quantum one
- In principle
  - The implementation of the oracle must be as *fast* and *efficient* as possible.
  - We want to call the oracle the *fewest* number of times as possible to minimize the complexity of the algorithm.
- Caveat: if the access to the classical oracle is given by a RAM, we need *QRAM*.

[Giovannetti, Lloyd, Maccone, 2008]

Appendix

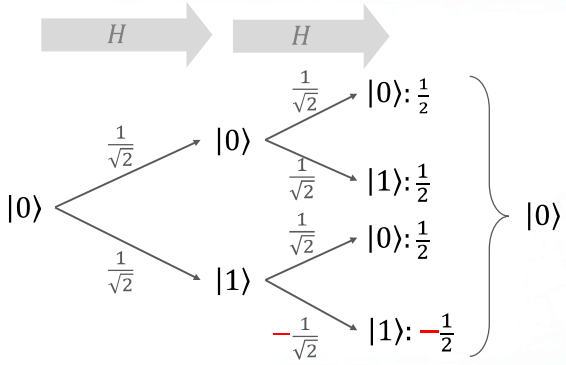
– Entanglement is necessary – Quantum Interference – **BQP ⊂ PSPACE**

Entanglement is necessary for advantage in quantum computation

- What is the *sufficient condition* for advantage in quantum computation?
  → We don't even know if there is (except for the oracle problems)!
- Consider circuits $\{\mathcal{C}_n\}_n$ that solves a problem A in **BQP** using 1 & 2 qubit gates. Further suppose for any input to the circuit $\mathcal{C}_n$ (for any n), the quantum state at every stage of the process is *enentangled*. → The problem A is also in **BPP**.
- Reasoning:
  1. Suppose the state at some stage is given as $|\alpha_1\rangle|\alpha_2\rangle\cdots|\alpha_n\rangle$ and in the next step applies a 2-qubit gate U to qubits 1 & 2: $(U|\alpha_1\rangle|\alpha_2\rangle)|\alpha_3\rangle\cdots|\alpha_n\rangle$.
  2. By a 4×4 matrix multiplication, $U|\alpha_1\rangle|\alpha_2\rangle$ takes the form $a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$. Then factorize it into $|\beta_1\rangle|\beta_2\rangle$ in just constant (in n) time.
  3. Single qubit measurement: the probabilities of the outcome for each qubit is easily simulated by using the squared amplitude of each qubit state.

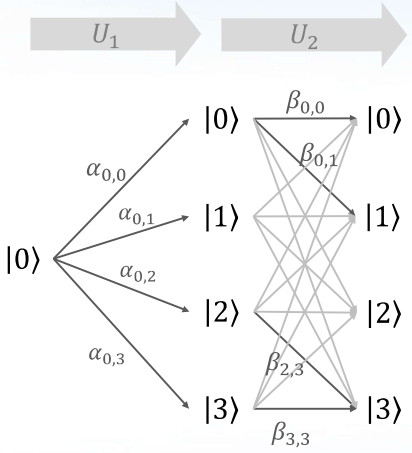
Quantum Interference (1/2)

- Using *Feynman's sum-over-path* (*path-integral*) description for $|0\rangle \xrightarrow{H} |+\rangle \xrightarrow{H} |0\rangle$:



1. Each path has an amplitude given by the product of numbers along the path.
2. Each final basis state has an amplitude given by the sum over all paths from the start to it.
3. The probability for the transition from initial $|0\rangle$ to a final basis is the modulus square of sum over all paths to it.

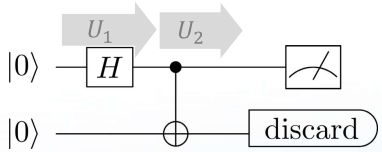
Quantum Interference (2/2)



Pr(final outcome is 3) $= \left|\sum_j \alpha_{0,j}\,\beta_{j,3}\right|^2$

Why BQP ⊂ PSPACE? [§4.5.5, N&C], [§5, Preskill]

- **PSPACE**: problems that can be solved with an algorithm that uses a polynomial amount of space or memory (and can thus generally run for exponential time).
- At first glance, it seems we need exp-space to store an n-qubit state $\sum_{x\in\{0,1\}^n} \alpha_x|x\rangle$.
- Intuition: WLOF, suppose the decision outcome given by measuring the first qubit.
  1. Don't need to do the full matrix multiplication all the ways on all qubits.
     → Calculating the probabilities using *Feyman's sum-over-path method*.
  2. This takes poly-space since there're only poly-gates (**BQP**) using $\{H, T, \text{CNOT}\}$ and individual terms in the sum can be erased after being added to the running total.
- Ex:



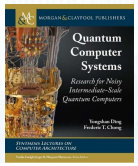
$$\Pr(\text{first qubit is } 0) = |\langle 00|U_2U_1|00\rangle|^2 + |\langle 01|U_2U_1|00\rangle|^2$$

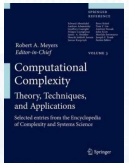
Final Remarks

- It's good to have $\{H, T, \text{CNOT}\}$ being the universal set and the *Solovay–Kitaev Thm.*
- → But approximating arbitrary unitary gates is generally *inefficient*. [§4.5.4, N&C]
- Most Boolean functions need exponentially many gates (either classical or quantum).
- Classical probabilistic algorithms can be easily simulated by quantum algorithms (via probabilistic transition diagrams such as Feynman's method); is the converse true?
- If there is no entanglement, or a sufficiently small amount of entanglement, then there are also efficient classical algorithms for simulating quantum systems.
  → Circuits using only CNOT, H, X, Z, T (known as the *Clifford group*) can be *efficiently* simulated on a classical computer (*Gottesman–Knill theorem*). What's beyond this set?
- Fast, nearly-optimal circuit synthesis for the set of single-qubit {Clifford, T} circuits.

References (1/2)

- Frederic T. Chong and Yongshan Ding. *Quantum Computer Systems: Research for Noisy Intermediate-Scale Quantum Computers*. Morgan & Claypool Publishers, 2020.
- R. A. Meyers (Editor). *Computational Complexity: Theory, Techniques, and Applications*. Springer, 2012.
- Complexity Zoo: https://complexityzoo.net/Complexity_Zoo
- John Preskill's lecture note: http://theory.caltech.edu/~preskill/ph219/chap5_13.pdf
- Ryan O'Donnell's lecture note: https://www.cs.cmu.edu/~odonnell/quantum15/lecture23.pdf
- Ronald de Wolf's lecture note: https://homepages.cwi.nl/~rdewolf/qcnotes.pdf
- Andrew Childe's lecture note: https://www.cs.umd.edu/~amchilds/qa/qa.pdf





References (2/2)

- V. Giovannetti, S. Lloyd, L. Maccone, "Quantum random access memory," *Physical Review Letters*, 100(160501), 2008.
- D. K. Par, F. Petruccione, and J.-K. K. Rhee, "Circuit-Based Quantum Random Access Memory for Classical Data," *Scientific Reports*, 3949, 2019.
- D. Gottesman, "The Heisenberg representation of quantum computers." arXiv:quant-ph/9807006, 1998.
- S. Aaronson and D. Gottesman, "Improved Simulation of Stabilizer Circuits," *Physical Review A*, 70(052328), 2004.
- V. Kliuchnikov, D. Maslov, and M. Mosca, "Fast and efficient exact synthesis of single qubit unitaries generated by Clifford and T gates," *Quantum Information and Computation*, 13(7–8):607, 2013.
- B. Giles and P. Selinger, "Remarks on Matsumoto and Amano's normal form for single-qubit Clifford+T operators," arXiv:1312.6584, 2013.

Quantum Information and Computation
Amplitude Amplification Algorithm

Hao-Chung Cheng (鄭皓中)
haochung@ntu.edu.tw

Department of Electrical Engineering
National Taiwan University

March 24, 2021

Outline

Motivation

The Phone Book Problem

- The phone book problem involves finding the correct name in a telephone book for a given telephone number (e.g. 0955556666) out of total N names.
- The distribution of the telephone numbers is supposed to be *random*. The phone book is stored in the oracle.

unstructured

Query:
"Does 灶門炭治郎 have the number 0955556666?"

The oracle answers in this case with "No".

Index	Name	Phone number
1	灶門炭治郎	0952074925
2	艾倫耶格爾	0987878787
⋮	⋮	⋮
l	奇犴揍敵客	0955556666
$l+1$	流川楓	0900000000
⋮	⋮	⋮

The Search Problem

- We are searching in a large database with $N = 2^n$ size to find a target item.
- Assume that the database is *unsorted* but given each item we can easily check whether or not it's the one we seek → each access to the database is called a *query*.
- Goal: we should locate the item with some constant level of probability (half say) independent of the size N.

The Unstructured Search Problem

- ***Given***: A black box for a Boolean function $f: \{0,1\}^n \to \{0,1\}$.
- ***Promise***: There is a unique x such that $f(x) = 1$.
- ***Problem***: Find this special x.

The Quantum Search Algorithm

- Classical solution: $\Theta(N) = \Theta(2^n)$ queries are sufficient and necessary.
  → If we examine an item and find it bad,
  we gain no further information about the location of the good item.
- Quantum solution (by Lov Grover 1996):
  → $O(\sqrt{N})$ queries are sufficient to locate the good item *with high probability.*
  → *Quadratic speedup* (in terms of the query complexity).
  → It is shown to optimal (by Bennett *et al.* 1997), i.e. $\Omega(\sqrt{N})$ queries are needed.
- Proof strategies for *amplitude amplification*:
  - Geometric method via rotation operations.
  - Manipulating probability amplitudes.
  - Algebraic approach (see HW2).

Applications and **NP** Problems

- Basically any classical algorithm that has some search-component can be improved using Grover's algorithm as a subroutine, e.g. finding shortest paths, scheduling, etc.
- The satisfiability problem SAT:
  Given a Boolean formula $\phi(x_1, \dots, x_n)$ with n variables and single bit output, we want to know if there is an assignment $x_1 \cdots x_n = b_1 \cdots b_n$ with $\phi(b_1, \dots, b_n) = 1$.
  A brute force evaluation of all 2^n possible assignments takes exponential time.
  → SAT has been proved to be in **NP**-complete, but we don't know if it is in **P**.
- *Unstructured search* can be thought of as a model for solving problems in **NP** by brute force search – if brute-force search is basically the best thing we can do classically to solve some **NP**-hard problems, then that computation can be sped up quadratically using Grover's algorithm, e.g. quantum counting, amplitude estimation.

[Durr–Høyer, 1996], [Brassard et al., 2002], [Aaronson–Rall, 2020]

Amplitude Amplification

Amplitude Amplification

- Thought Experiment:
  Suppose we have three qubits and the computational basis vectors for them each correspond to a possible solution to some problem.

$$a_0|000\rangle + a_1|001\rangle + a_2|010\rangle + a_3|011\rangle + a_4|100\rangle + a_5|101\rangle + a_6|110\rangle + a_7|111\rangle$$

  → To devise an algorithm that chooses the best solution among them.

- *Amplitude amplification* is a process that we manipulate the state so that the basis vector which represents the best solution has the coefficient a_j with the largest probability.

- Ultimately, this is the goal of every quantum algorithm: have the results of the final qubit measurements correspond with high probability to the best solution.

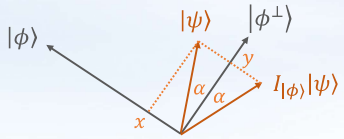
Inspection

- Intuitively, the ket vector corresponding to the good item occurs with only an *exponentially small amplitude* in the total superposition.

- If one can create a *phase difference* for the marked ket vector, then the *amplitude amplification algorithm* can strengthen the amplitude of the marked ket.

NEEDLE IN A HAYSTACK

Reflection Operator



- Let $|\phi\rangle$ be any unit ket vector.
  Define the "*reflection* in the subspace that is orthogonal to $|\phi\rangle$": $I_{|\phi\rangle} \coloneqq I - 2|\phi\rangle\langle\phi|$

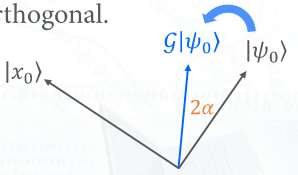
- Properties:
  a) Let $|\phi^\perp\rangle$ be any chosen unit vector orthogonal to $|\phi\rangle$. Then any ket vector $|\psi\rangle$ may be uniquely expressed as $|\psi\rangle = x|\phi\rangle + y|\phi^\perp\rangle$ and then $I_{|\phi\rangle}|\psi\rangle = -x|\phi\rangle + y|\phi^\perp\rangle$.
  b) It holds that $-I_{|\phi\rangle}|\psi\rangle = x|\phi\rangle - y|\phi^\perp\rangle = I_{|\phi^\perp\rangle}$.
  c) For any unitary operator, we have $UI_{|\phi\rangle}U^\dagger = I_{U|\phi\rangle}$.
  d) For any $|\alpha\rangle, |\beta\rangle$, let $\mathcal{P}$ be the (real) plane spanned by $|\alpha\rangle$ and $|\beta\rangle$.
  Then $I_{|\alpha\rangle}$ and $I_{|\beta\rangle}$ preserve the plane $\mathcal{P}$, i.e. $\forall|\psi\rangle \in \mathcal{P}$, we have $I_{|\alpha\rangle}|\psi\rangle, I_{|\beta\rangle}|\psi\rangle \in \mathcal{P}$.
  ($|\psi\rangle$ is only modified by a multiple of $|\alpha\rangle$, $\because I_{|\alpha\rangle}|\psi\rangle = |\psi\rangle - 2\langle\alpha|\psi\rangle|\alpha\rangle$.)
  e) For any $x \in \{0,1\}^n$, we have $I_{|x_0\rangle}|x\rangle = \begin{cases} |x\rangle, & x \neq x_0 \\ -|x_0\rangle, & x = x_0 \end{cases}$.

Grover's Search Algorithm

- Define the *Grover diffusion operator* $\mathcal{G} \coloneqq -H_n I_{|0\rangle} H_n I_{|x_0\rangle}$ ($H_n \coloneqq H^{\otimes n}$)
- Claims:
  a) In the plane $\mathcal{P}(x_0)$ spanned by (the unknown target) $|x_0\rangle$ and $|\psi_0\rangle \coloneqq H_n|0\rangle$, $\mathcal{G}$ is rotation through angle 2α where $\sin\alpha = 1/\sqrt{N}$.
  b) In the subspace orthogonal to $\mathcal{P}(x_0)$, $\mathcal{G} = -I$.

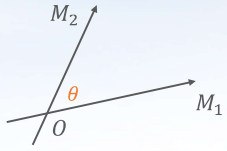
All amplitudes of $|\psi_0\rangle$ and matrix elements of $\mathcal{G}$ are *real numbers*

- Iteration:
  - For large N, $\langle x_0|\psi_0\rangle = 1/\sqrt{N}$; states $|x_0\rangle$ and $|\psi_0\rangle$ are nearly orthogonal.
    $\rightarrow 2\alpha \approx 2\sin\alpha = 2/\sqrt{N}$.
  - Repeatedly applying $\mathcal{G}$ to the starting state $|\psi_0\rangle$.
    $\Rightarrow$ About $\frac{\pi/2}{2\alpha} \approx \frac{\pi\sqrt{N}}{4}$ iterations will be needed to reach $|x_0\rangle$.
  - More precisely, # of iteration needed is the integer near $\frac{\cos^{-1}\frac{1}{\sqrt{N}}}{2\sin^{-1}\frac{1}{\sqrt{N}}}$.

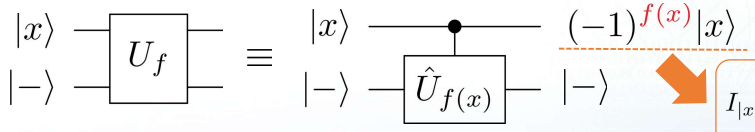


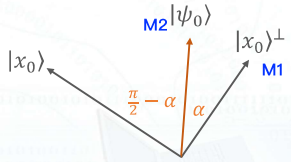
← Independent of $|x_0\rangle$!

The Grover Iteration Operator



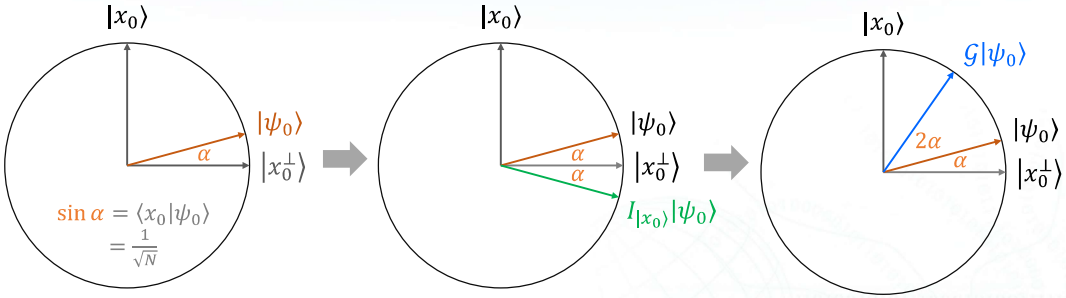
- $\mathcal{G} := -H_n I_{|0\rangle} H_n I_{|x_0\rangle} = -I_{H_n|0\rangle} I_{|x_0\rangle} = -I_{|\psi_0\rangle} I_{|x_0\rangle} = I_{|\psi_0\rangle}{}^{\perp} I_{|x_0\rangle}$
- Lemma: Let M_1 and M_2 be two mirror lines in the Euclidean plane $\mathbb{R}^2$ intersecting at a point O and let θ be the angle in the plane from M_1 to M_2 . Then the operation of reflection in M_1 followed by reflection in M_2 is (anticlockwise) rotation by angle 2θ about the point O.
- $\rightarrow \mathcal{G}$ is rotation by angle 2α, where α is the angle between mirror lines $|x_0\rangle^{\perp}$ & $|\psi_0\rangle$, i.e. $\langle x_0|\psi_0\rangle = \cos(\frac{\pi}{2}-\alpha) = \sin\alpha$.
- How to implement $I_{|x_0\rangle}$? → *Phase kickback*!





$$I_{|x_0\rangle}|x\rangle = \begin{cases} |x\rangle & \text{if } x \neq x_0 \\ -|x_0\rangle & \text{if } x = x_0 \end{cases}$$

Geometric Illustration



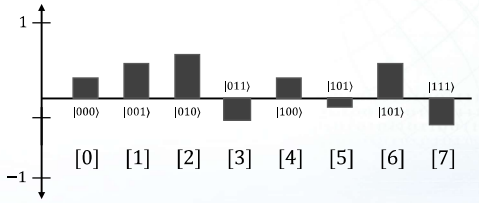
Amplitudes and Circuit

Perspective from Amplitudes

- Previously, we have seen how to reach the target state $|x_0\rangle$ from $|\psi_0\rangle$ via rotation. How are the *probability amplitudes* evolving through the iteration?
- Fix a basis (say the computational basis). We can express the probability amplitude of any state vector as a *vector* or a *discrete-time signal*.

$$|\psi\rangle = \sum_x a_x|x\rangle = A = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \end{pmatrix}$$

- For real-valued probability amplitudes, a 3-qubit state vector is like:

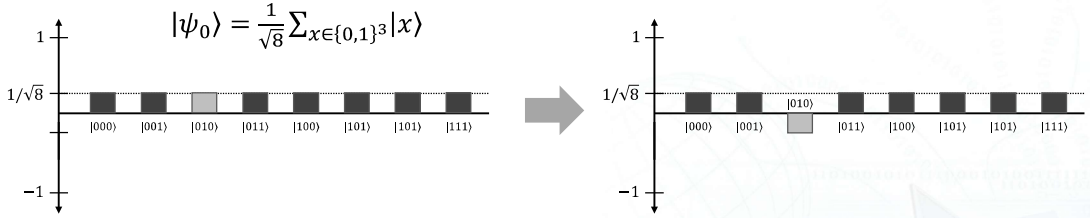


A Two-Step Circuit Subroutine (1/3)

$\mathcal{G} = -I_{|\psi_0\rangle} I_{|x_0\rangle}$

- Goal: Transform the superposition such that $|x_0\rangle$ will be measured with high prob.

1. *Flipping the sign* → creating the phase difference: $I_{|x_0\rangle}|x\rangle = \begin{cases} |x\rangle & \text{if } x \neq x_0 \\ -|x_0\rangle & \text{if } x = x_0 \end{cases}$

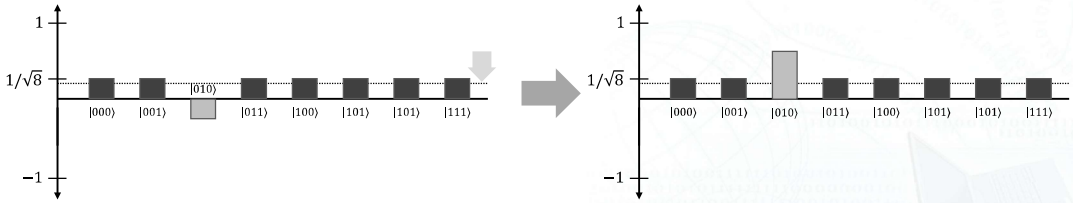


A Two-Step Circuit Subroutine (2/3)

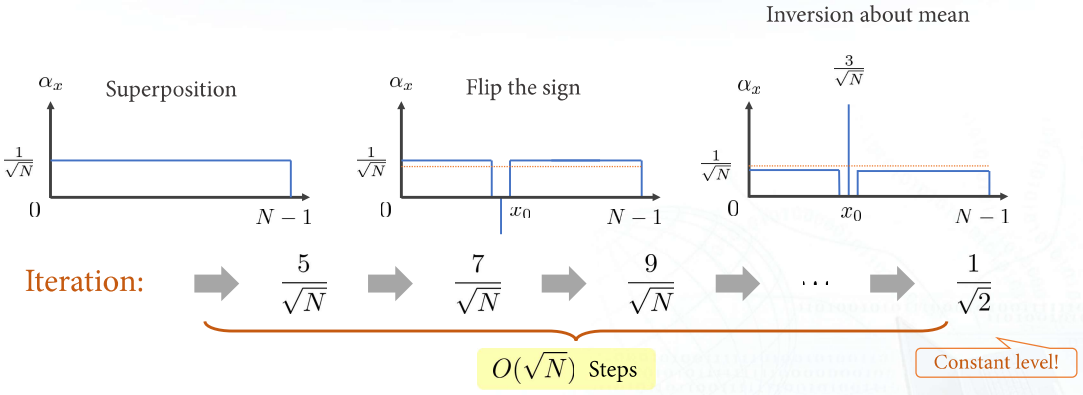
$\mathcal{G} = -I_{|\psi_0\rangle} I_{|x_0\rangle}$

2. *Inversion about mean*:
Let $\{a_x\}$ be a collection of numbers and $\bar{a}$ be the mean. Then the numbers $\{2\bar{a} - a_x\}$ are the inversion about mean $\bar{a}$.

$$-I_{|\psi_0\rangle} = 2|\psi_0\rangle\langle\psi_0| - I, \ |\psi_0\rangle := \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle$$



A Two-Step Circuit Subroutine (3/3)



Why Inverse about the Mean?

- The *inversion about mean* operator $|\psi_0\rangle := \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle = H^{\otimes n}|0\rangle^{\otimes n}$

$$I_{|\psi_0\rangle} := 2|\psi_0\rangle\langle\psi_0| - I \ = H^{\otimes n}\left(2(|0\rangle\langle 0|^{\otimes n} - I\right) H^{\otimes n}$$

- Check for an n-qubit state $\phi\rangle = \sum_x a_x|x\rangle$:

$$\begin{aligned} \Rightarrow -I_{|\psi_0\rangle}|\phi\rangle &= (2|\psi_0\rangle\langle\psi_0| - I)\,|\phi\rangle \\ &= 2|\psi_0\rangle\langle\psi_0|\phi\rangle - |\phi\rangle \\ &= 2|\psi_0\rangle\langle\psi_0|\phi\rangle - \textstyle\sum_x a_x|x\rangle \\ &= 2\bar{a}\textstyle\sum_x|x\rangle - \sum_x a_x|x\rangle \\ &= \textstyle\sum_x (2\bar{a} - a_x)\,|x\rangle \end{aligned}$$

$$\begin{aligned} |\psi_0\rangle\langle\psi_0|\phi\rangle &= \frac{1}{\sqrt{N}}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}\frac{1}{\sqrt{N}}\begin{pmatrix}1 & \cdots & 1\end{pmatrix}\begin{pmatrix}a_0\\ a_1\\ a_2\\ \vdots \\ a_{2^n-1}\end{pmatrix} \\ &= \frac{1}{N}\left(\sum_x a_x\right)\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix} \\ &= \bar{a}\sum_x|x\rangle \end{aligned}$$

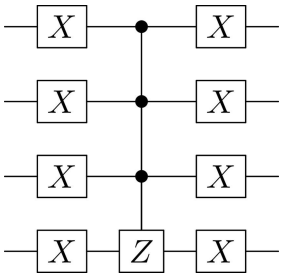
Mean of the amplitudes

How to Implement Inverse about the Mean? (1/2)

- Circuit model for $2|0\rangle\langle 0|^{\otimes n} - I = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{pmatrix}$

Multi-controlled Z

$=$



$$X^{\otimes n}(\mathrm{MCZ})X^{\otimes n}|x_1\rangle \dots |x_n\rangle$$

$$= X^{\otimes n}(\mathrm{MCZ})|\bar{x}_1\rangle \dots |\bar{x}_n\rangle$$

$Z|x\rangle = (-1)^x|x\rangle$

$$= X^{\otimes n}(-1)^{\text{if all } \bar{x}_i=1}|\bar{x}_1\rangle \dots |\bar{x}_n\rangle$$

$$= (-1)^{\text{if all } \bar{x}_i=1}|x_1\rangle \dots |x_n\rangle$$

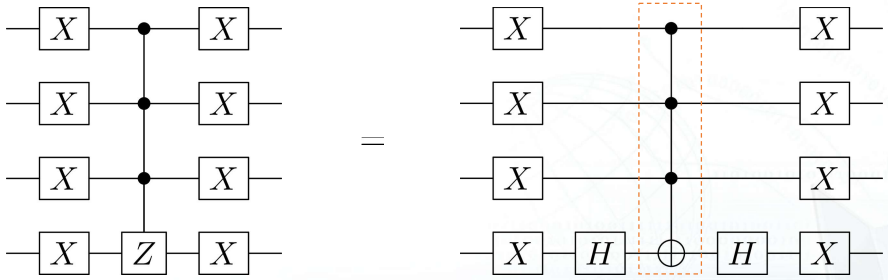
$$= (-1)^{\text{if all } x_i=0}|x_1\rangle \dots |x_n\rangle$$

Note that the global phase doesn't matter

How to Implement Inverse about the Mean? (1/2)

- How to implement $2|0\rangle\langle 0|^{\otimes n} - I = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{pmatrix}$?

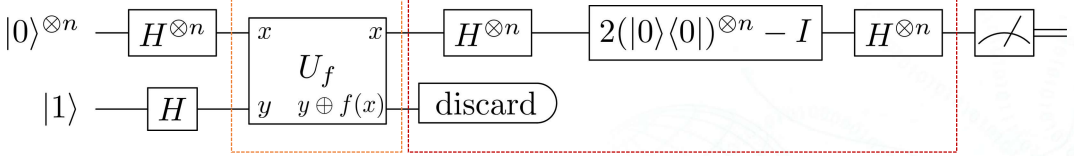
Multi-controlled *Toffoli* gate



[Schuch-Siewert, 2003], [Shende - Markov, 2009]

Circuit Diagram

Repeat $\frac{\pi}{4}\sqrt{2^n}$ times



$I_{|x_0\rangle}|x\rangle = (-1)^{f(x)}|x\rangle$

$$I_{|\psi_0\rangle} := 2|\psi_0\rangle\langle\psi_0| - I$$

$$= H^{\otimes n}\left(2(|0\rangle\langle 0|^{\otimes n} - I\right)H^{\otimes n}$$

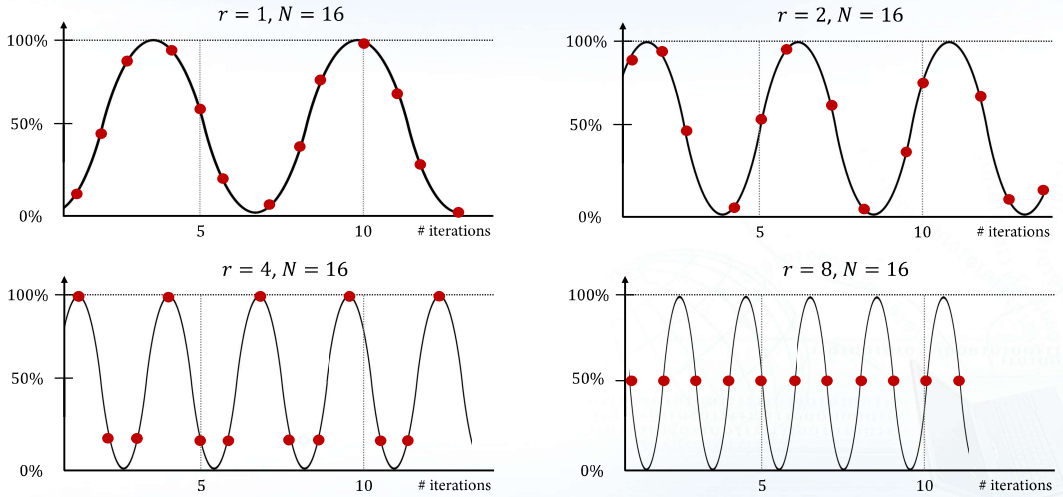
Concluding Remarks

After Thought

- Searching with multiple, say r , items would require $O\left(\sqrt{N/r}\right)$ queries.
- Whether there is quantum advantage for searching in which kinds of structure in database is still largely open and a topic of current research.
- There is currently no way to quickly input large amounts of data before the loading operations exhaust the physical qubit coherence times.
- In Grover's search algorithm, we do not have to pre-load all the data as long as the oracle can correctly identify the object we need.
- How to construct the oracle? → It could actually be very complicated. [Figgatt et al., 2017]

  → Grover's algorithm does not really improve searching in the phone book problem.
- Grover's algorithm works best as a subroutine for information that is already represented within the states of the qubits or implicitly usable.

Success Probability vs # Iterations



Appendix

Optimality of Quantum Search Algorithms [§6.6, N&C]

- Motivation: Previously, we've seen using $O(\sqrt{N})$ queries is *sufficient* to locate the target item with high probability. Can we use fewer queries, say $O(\log N)$?

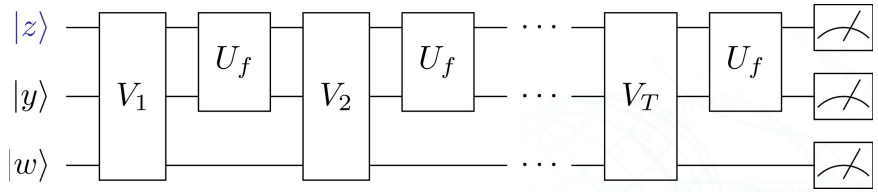
  → If it was so, any problem whose solution can be verified efficiently can be solved efficiently. Namely, **NP ⊆ BQP**! Unfortunately, it is not true.

- Theorem (Optimality). Any quantum algorithm that achieves the search problem for a unique target item in an unstructured database of size N using T queries with failure probability $\varepsilon < 1/2$ must satisfy $T \geq \left(\frac{1}{2} - \varepsilon\right)\sqrt{N} = \Omega(N)$.
- This was originally proved by Bennett *et al.* in 1997 before Grover's quantum search algorithm was found. [Bennett *et al.*, 1997]
- The lower bound can be improved to show that Grover's algorithm is *exactly optimal* and its performance cannot be improved by even one query.

[Zalka, 1999], [Høyer–Špalek, 2005]

General Quantum Query Algorithms

- Consider an arbitrary quantum algorithm which completes the above mentioned task. Such an algorithm operates on three registers: an n-qubit input register, a 1-qubit output register, and an ancilla register of (arbitrary) m qubits for workspace. A computational basis state can be written as $|z\rangle|y\rangle|w\rangle$, $z \in \mathbb{Z}_2^n$, $y \in \mathbb{Z}_2$, $w \in \mathbb{Z}_2^m$.



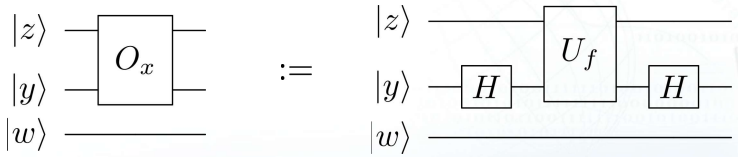
- The operation starts in the state $|0\rangle|0\rangle|0\rangle$, with a sequence of arbitrary but fixed unitary operators $V_1, V_2, \ldots$, which do not depend on f, interleaved with oracle U_f (operating on the input & output registers).

Proof Idea (1/6)

$f(x) = 1$; otherwise 0

- For any possible target $x \in \mathbb{Z}_2^n$, we define $O_x := (I \otimes H)U_f(I \otimes H)$, WLOG.
  Counting queries to O_x is the same as counting queries to U_f.
  Introduce such a new oracle will simplify our analysis since it diagonalizes U_f; it can be verified that:

$$\begin{cases} O_x|z\rangle|y\rangle|w\rangle = (-1)^{y\delta_{xz}}|z\rangle|y\rangle|w\rangle \\ (O_x - I)\,|z\rangle|y\rangle|w\rangle = -2y\delta_{xz}|z\rangle|y\rangle|w\rangle \end{cases}$$



Proof Idea (2/6)

- Assume the quantum algorithm makes T queries. We define the state of the system after $t \geq T$ queries, given that $f(x) = 1$ as
  - $|\psi_{x,t}\rangle := O_x\, V_t\, O_x\, V_{t-1} \cdots O_x V_1\, |0\rangle$;
  - and $|\phi_t\rangle := V_t V_{t-1} \cdots V_1|0\rangle$ be the result of not applying the oracle queries.
- *Goal*: to bound $\||\psi_{x,T}\rangle - |\phi_T\rangle\|$. Deviation after T steps
- *Intuition*: if this norm is too small, then all the states $\{|\psi_{x,T}\rangle\}_{x\in\mathbb{Z}_2^n}$ are roughly the same (no amplification); it is not possible to correctly identify x with high prob.
- *Approach*:
  1. Assuming the algorithm works with error probability $< \varepsilon$, the norm cannot be two small. → to derive a lower bound to the norm in terms of ε.
  2. To derive an universal upper bound to the norm in terms of T for all possible O_x and V_t's.

Proof Idea (3/6)

Euclidean norm (2-norm)

1. (Lower bound) Invoke a lemma:
   Let $|\psi_1\rangle$ and $|\psi_2\rangle$ be such that $\|\psi_1 - \psi_2\| \leq \delta$. Then the best error probability of distinguishing them is $P_e^\star(\psi_1, \psi_2) \geq \frac{1}{2}(1 - \delta)$.
   (States are *closer* → *harder* to distinguish them → *higher* error probability.)
   - Proof of the lemma: Let p_1 and p_2 be the prob. distributions of the measurement (which can be chosen arbitrarily) for $|\psi_1\rangle$ and $|\psi_2\rangle$, and let $a_x := |\langle x|\psi_1\rangle|$, $b_x := |\langle x|\psi_2\rangle|$.

     Using the Cauchy–Schwarz inequality we get

$$\begin{aligned} \sum_x |p_1(x) - p_2(x)| &= \sum_x |a_x^2 - b_x^2| = \sum_x |a_x - b_x|\,(a_x + b_x) \\ &\leq \sqrt{\sum_x (a_x - b_x)^2}\sqrt{\sum_x (a_x + b_x)^2} \\ &\leq \sqrt{\sum_x |\langle x|\psi_1\rangle - \langle x|\psi_2\rangle|^2}\left(\sqrt{\sum_x a_x^2} + \sqrt{\sum_x b_x^2}\right) \\ &= 2\,\||\psi_1\rangle - |\psi_2\rangle\| \end{aligned}$$

Proof Idea (4/6)

1. (Lower bound) Invoke a lemma:
   Let $|\psi_1\rangle$ and $|\psi_2\rangle$ be such that $\|\psi_1 - \psi_2\| \le \delta$. Then the best error probability of distinguishing them is $P_e^\star(\psi_1, \psi_2) \ge \frac{1}{2}(1-\delta)$.
   (States are closer → hard to distinguish them → higher error probability.)

   Euclidean norm (2-norm)

   - Proof of the lemma:

     On the other hand, the classical Holevo bound for distinguishing them is

$$\begin{aligned} P_e^\star(p_1, p_2) &= \frac{1}{2}\left(1 - \frac{1}{2}\sum_x |p_1(x) - p_2(x)|\right) \\ &\ge \frac{1}{2}\left(1 - \|\psi_1 - \psi_2\|\right) \ge \frac{1}{2}(1-\delta) \end{aligned}$$

⇒ Using the lemma, we get: $P_e^\star(\psi_1, \psi_2) < \varepsilon$ implies $\left\| |\psi_{x,T}\rangle - |\phi_T\rangle \right\| > 1 - 2\varepsilon$.

Proof Idea (5/6)

2. (Upper bound)

   Using iterative sum to get an upper bound.

   Define $|D_{x,t}\rangle = O_x|\phi_t\rangle - |\phi_t\rangle$

   Then,

$$\begin{aligned} |\psi_{x,1}\rangle &= |\phi_1\rangle + |D_{x,1}\rangle \\ |\psi_{x,2}\rangle &= |\phi_2\rangle + |D_{x,2}\rangle + O_x V_2 |D_{x,1}\rangle \\ &\vdots \\ |\psi_{x,T}\rangle &= |\phi_T\rangle + |D_{x,T}\rangle + O_x V_T |D_{x,T-1}\rangle + \cdots + O_x V_T \cdots O_x V_2 |D_{x,1}\rangle \end{aligned}$$

   Triangle inequality:

$$\begin{aligned} \||\psi_{x,T}\rangle - |\phi_T\rangle\| &= \| |D_{x,T}\rangle + O_x V_T |D_{x,T-1}\rangle + \cdots + O_x V_T \cdots O_x V_2 |D_{x,1}\rangle | \\ &\le \sum_{t=1}^T \| |D_{x,t}\rangle \| = \sum_{t=1}^T \|(O_x - I)|\phi_t\rangle\| \end{aligned}$$

Proof Idea (6/6)

2. (Upper bound)

   Cauchy–Schwarz inequality:

$$\| |\psi_{x,T}\rangle - |\phi_T\rangle \|^2 \le \left(\sum_{t=1}^T \|(O_x - I)|\phi_t\rangle\| \right)^2 \le T \sum_{t=1}^T \|(O_x - I)|\phi_t\rangle\|^2$$

   Taking average over all $x \in \{0,1\}^n$:

$$\begin{aligned} \frac{1}{N}\sum_{x\in\{0,1\}^n} \| |\psi_{x,T}\rangle - |\phi_T\rangle \|^2 &\le \frac{T}{N} \sum_{t=1}^T \sum_{x\in\{0,1\}^n} \|(O_x - I)|\phi_t\rangle\|^2 \\ &\le \frac{T^2}{N} \max_{|\phi\rangle} \sum_{x\in\{0,1\}^n} \|(O_x - I)|\phi\rangle\|^2 \\ &\le \frac{4T^2}{N} \max_{|\phi\rangle} \sum_{x\in\{0,1\}^n} \sum_w |\langle\phi|x\rangle|1\rangle|w\rangle|^2 = \frac{4T^2}{N} \end{aligned}$$

   $(O_x - I)\,|z\rangle|y\rangle|w\rangle = -2y\delta_{xz}|z\rangle|y\rangle|w\rangle$

3. Combining with the lower bound: $(1-2\varepsilon)^2 \le \frac{4T^2}{N}$

□

Final Remarks

- General lower bounds on quantum query complexity is actually not easy.
  It is about not what a quantum computer *can* do, but what it *cannot* do.
  In other words, lower bounds (to the query complexity) give us a guide to what we can hope to achieve with quantum computers.
  The term *optimality* allows us to understand the limit of the power of quantum algorithms in certain problems.
- Standard approaches were developed
  - The Polynomial methods [Beals *et al.*, 2001]
  - The quantum adversary methods [Ambainis, 2002], [Ambainis, 2006], etc.
- For some survey papers, we refer to [Bernstein–Vazirani, 1997], [Høyer–Špalek, 2005], [Burhman–de Wolf, 2002], and again lecture notes by Andrew Childes, Ronald de Wolf, and Ryan O'Donnell.

References (1/3)

- L. K. Grover, "A fast quantum mechanical algorithm for database search, in *Proceedings of 28th Annual ACM Symposium on the Theory of Computing*, 1996.
- G. Brassard, P. Høyer, A. Tapp, "Quantum algorithm for the collision problem, arXiv:quant-ph/9705002.
- G. Brassard, P. Høyer, A. Tapp, "Quantum counting," in *Proceedings of ICALP'98, Lecture Notes in Computer Science*, 1443:820–831, 1998. arXiv:quant-ph/9805082.
- C. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, "Strengths and weaknesses of quantum computing," *SIAM Journal of Computing*, 26(5):1510–1523, 1997.
- C. Zalka, "Grover's quantum searching algorithm is optimal," *Physical Review A*, 60(4):2746–2751, 1999.
- P. Høyer and R. Špalek, "Lower bounds on quantum query complexity," in *Bulletin of the European Association for Theoretical Computer Science*, 87:78–103, 2005.
- R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf, "Quantum lower bounds by polynomials," *Journal of the ACM*, 48(4):778–797, 2001.

References (2/3)

- A. Ambainis, "Quantum lower bounds by quantum arguments," *Journal of Computer and System Sciences*, 64(4):750–767, 2002.
- A. Ambainis. "Polynomial degree vs. quantum query complexity," *Journal of Computer and System Sciences*, 72(2):220–238, 2006.
- P. Høyer, T. Lee, and R. Šalek, "Negative weights make adversaries stronger," in *Proc. 39th ACM Symposium on Theory of Computing*, pp. 526–535, 2007.
- H. Buhrman and R. de Wolf. "Complexity measures and decision tree complexity: a survey," *Theoretical Computer Science*, 288:21–43, 2002.
- B. W. Reichardt, "Span programs and quantum query complexity: The general adversary bound is nearly tight for every Boolean function," in *Proc. 50th IEEE Symposium on Foundations of Computer Science*, pp. 544–551, 2009.
- E. Bernstein, U. Vazirani, "Quantum complexity theory," *SIAM Journal on Computing*, 26:141–1473, 1997.
- http://twistedoakstudios.com/blog/Post2644_grovers-quantum-search-algorithm

References (3/3)

- C. Figgatt, D. Maslov, K. A. Landsman, N. M. Linke, S. Debnath & C. Monroe , "Complete 3-Qubit Grover search on a programmable quantum computer," *Nature Communications*, 8:1918, 2017.
- E. Arikan, "An information-theoretic analysis of Grover's algorithm." *Quantum Communication and Information Technologies,* Springer, Dordrecht, 339–347, 2003.
- I. L. Chuang, N. Gershenfeld, and M. Kubinec, "Experimental Implementation of Fast Quantum Searching," *Physical Review Letters*, 80(15), 1998.
- N. Schuch, J. Siewert, "Programmable networks for quantum algorithm," *Physical Review Letters*, 91, 027902, 2003. & also see V. V. Shende, I. L. Markov, "On the CNOT-cost of TOFFOLI gates," *Quantum Information and Computation*, 9(5-6):461-486, 2009.
- C. Durr, P. Høyer, "A Quantum Algorithm for Finding the Minimum," arXiv:quant-ph/9607014
- G. Brassard, P. Hoyer, M. Mosca, A. Tapp, "Quantum Amplitude Amplification and Estimation," *AMS Contemporary Mathematics*, 305:53-74, 2002.
- Scott Aaronson, Patrick Rall, "Quantum Approximate Counting, Simplified," in *Symposium on Simplicity in Algorithms*, 24-32, 2020.

Quantum Information and Computation
Phase Estimation Algorithm

Hao-Chung Cheng (鄭皓中)

haochung@ntu.edu.tw

Department of Electrical Engineering

National Taiwan University

March 31, 2021

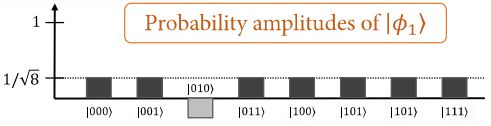
Outline

Motivation

Motivation (1/2)

- Suppose a quantum system is given by one of the following three states. We know how to determine them (with certain probability) by using Grover's idea.

$$|\phi_0\rangle = \tfrac{1}{\sqrt{8}}(|0\rangle + |1\rangle + |2\rangle + |3\rangle + |4\rangle + |5\rangle + |6\rangle + |7\rangle)$$
$$|\phi_1\rangle = \tfrac{1}{\sqrt{8}}(|0\rangle + |1\rangle - |2\rangle + |3\rangle + |4\rangle + |5\rangle + |6\rangle + |7\rangle)$$
$$|\phi_2\rangle = \tfrac{1}{\sqrt{8}}(|0\rangle + |1\rangle + |2\rangle + |3\rangle + |4\rangle + |5\rangle - |6\rangle + |7\rangle)$$

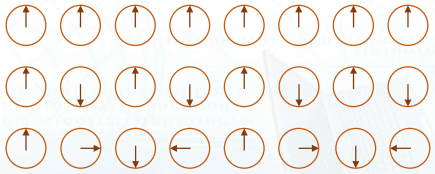


- Now we are instead given by one of the following three states:

Each probability amplitude has the same magnitude

Phases of each probability amplitude a_j

$$|\phi_0\rangle = \tfrac{1}{\sqrt{8}}(|0\rangle + |1\rangle + |2\rangle + |3\rangle + |4\rangle + |5\rangle + |6\rangle + |7\rangle)$$
$$|\phi_1\rangle = \tfrac{1}{\sqrt{8}}(|0\rangle - |1\rangle + |2\rangle - |3\rangle + |4\rangle - |5\rangle + |6\rangle - |7\rangle)$$
$$|\phi_2\rangle = \tfrac{1}{\sqrt{8}}(|0\rangle + \mathrm{i}|1\rangle - |2\rangle - \mathrm{i}|3\rangle + |4\rangle + \mathrm{i}|5\rangle - |6\rangle - \mathrm{i}|7\rangle)$$



Motivation (2/2)

$|\phi\rangle = \sum_{y=0}^{7} A_y|y\rangle$

- Let us view the probability amplitudes of them as discrete-time signals:

$$|\phi_0\rangle = \tfrac{1}{\sqrt{8}}(|0\rangle + |1\rangle + |2\rangle + |3\rangle + |4\rangle + |5\rangle + |6\rangle + |7\rangle) = \tfrac{1}{\sqrt{2^3}}\begin{pmatrix}1 & 1 & 1 & 1 & 1 & 1 & 1 & 1\end{pmatrix}^\dagger$$
$$= \tfrac{1}{\sqrt{2^3}}\sum_{y=0}^{2^3-1} \mathrm{e}^{2\pi\mathrm{i}\frac{0}{2^3}y}|y\rangle \qquad \Rightarrow A_y = \tfrac{1}{\sqrt{2^3}}\mathrm{e}^{2\pi\mathrm{i}\frac{0}{2^3}y}$$
$$|\phi_1\rangle = \tfrac{1}{\sqrt{8}}(|0\rangle - |1\rangle + |2\rangle - |3\rangle + |4\rangle - |5\rangle + |6\rangle - |7\rangle) = \tfrac{1}{\sqrt{2^3}}\begin{pmatrix}1 & -1 & 1 & -1 & 1 & -1 & 1 & -1\end{pmatrix}^\dagger$$
$$= \tfrac{1}{\sqrt{2^3}}\sum_{y=0}^{2^3-1} \mathrm{e}^{2\pi\mathrm{i}\frac{4}{2^3}y}|y\rangle \qquad \Rightarrow A_y = \tfrac{1}{\sqrt{2^3}}\mathrm{e}^{2\pi\mathrm{i}\frac{4}{2^3}y}$$
$$|\phi_2\rangle = \tfrac{1}{\sqrt{8}}(|0\rangle + \mathrm{i}|1\rangle - |2\rangle - \mathrm{i}|3\rangle + |4\rangle + \mathrm{i}|5\rangle - |6\rangle - \mathrm{i}|7\rangle) = \tfrac{1}{\sqrt{2^3}}\begin{pmatrix}1 & \mathrm{i} & -1 & -\mathrm{i} & 1 & \mathrm{i} & -1 & -\mathrm{i}\end{pmatrix}^\dagger$$
$$= \tfrac{1}{\sqrt{2^3}}\sum_{y=0}^{2^3-1} \mathrm{e}^{2\pi\mathrm{i}\frac{2}{2^3}y}|y\rangle \qquad \Rightarrow A_y = \tfrac{1}{\sqrt{2^3}}\mathrm{e}^{2\pi\mathrm{i}\frac{2}{2^3}y}$$

- The prob. amplitudes of the above states are the *Fourier series* with some phase ϕ. If we can perform an inverse Fourier transform on them, we can extract ϕ.

Quantum Fourier Transform (1/4)

- Let $|0\rangle, |1\rangle, \ldots, |N-1\rangle$ be the computational basis in $\mathbb{C}^N$.

$\mathbb{Z}_N$: integers modulo N

- The *quantum Fourier transform* (QFT) over $\mathbb{Z}_N$, denoted by Q_N, is the transform:

$$Q_N : |x\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{y\in\mathbb{Z}_N} \mathrm{e}^{2\pi \mathrm{i} \frac{x}{N} y} |y\rangle =: \frac{1}{\sqrt{N}} \sum_{y\in\mathbb{Z}_N} \omega^{xy} |y\rangle$$

Primitive N-th root of unity $\omega := \mathrm{e}^{2\pi \mathrm{i}/N}$

$$Q_N^{-1} : |x\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{y\in\mathbb{Z}_N} \omega^{-xy} |y\rangle$$

Inverse quantum Fourier transform

- For arbitrary state $|\phi\rangle = \sum_{x=0}^{N-1} a_x |x\rangle$, we get $Q_N|\phi\rangle = \frac{1}{\sqrt{N}} \sum_{x\in\mathbb{Z}_N} \sum_{y\in\mathbb{Z}_N} a_x \omega^{xy} |y\rangle$.

Quantum Fourier Transform (2/4)

- The jk-th matrix entry is $[Q_N]_{jk} := \langle j|Q_N|k\rangle = \omega^{jk},\ j,k \in \{0,1,\ldots,N-1\}$

$$Q_N = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \cdots & \omega^{2(N-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \cdots & \omega^{3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \omega^{3(N-1)} & \cdots & \omega^{(N-1)(N-1)} \end{pmatrix}$$

$$\overset{\omega^N=1}{=} \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \cdots & \omega^{N-2} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \cdots & \omega^{N-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{N-2} & \omega^{N-3} & \cdots & \omega \end{pmatrix}$$

Examples:

$$Q_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = H$$

$$Q_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \mathrm{e}^{2\pi \mathrm{i}/3} & \mathrm{e}^{-2\pi \mathrm{i}/3} \\ 1 & \mathrm{e}^{-2\pi \mathrm{i}/3} & \mathrm{e}^{2\pi \mathrm{i}/3} \end{pmatrix}$$

$$Q_4 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \end{pmatrix}$$

Quantum Fourier Transform (3/4)

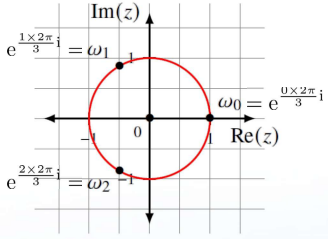
- Geometric series: $1 + \alpha + \alpha^2 + \cdots + \alpha^{N-1} = \begin{cases} N, & \alpha = 1 \\ \frac{1-\alpha^N}{1-\alpha}, & \alpha \neq 1 \end{cases}$

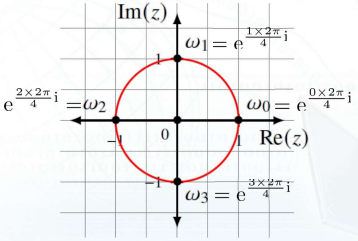
- Setting $\alpha = \omega^K$ for $\omega := \mathrm{e}^{\frac{2\pi \mathrm{i}}{N}}$, we have $\alpha = 1$ if and only if K is a multiple of N

$$\Rightarrow 1 + \omega^K + \omega^{2K} + \cdots + \omega^{(N-1)K} = \begin{cases} N, & \text{if } K \text{ is a multiple of } N \\ \frac{1-\omega^{NK}}{1-\omega} = 0, & \text{otherwise} \end{cases}$$

$K \equiv 0 \bmod N$

$K = 1$:





Quantum Fourier Transform (4/4)

- The jk-th element of $Q_N^\dagger Q_N$ is the $1/N$ times the sum of the j-th row of $Q_N^\dagger$ lined up against the k-th column of Q_N.
  → the geometric series with $\alpha = \omega^{k-j}$ divided by N, → δ_{jk}, → QFT is *unitary*

$$\langle j|Q_N^\dagger Q_N|k\rangle = \frac{1}{\sqrt{N}} \begin{pmatrix} & \vdots & \\ 1 & \omega^{-j} & \omega^{-2j} & \cdots & \omega^{-(N-1)j} \\ & \vdots & \end{pmatrix} \frac{1}{\sqrt{N}} \begin{pmatrix} & 1 & \\ \cdots & \omega^k & \cdots \\ & \omega^{2k} & \\ & \vdots & \\ & \omega^{(N-1)k} & \end{pmatrix}$$

$$= \frac{1}{N} \sum_{y\in\mathbb{Z}_N} (\omega^{jy})^* \omega^{ky} = \frac{1}{N} \sum_{y\in\mathbb{Z}_N} \omega^{(k-j)y} = \begin{cases} 1 & (k-j) = 0 \bmod N \equiv j = k \\ 0 & (k-j) \neq 0 \bmod N \equiv j \neq k \end{cases}$$

- The inverse quantum Fourier transform is just to apply $Q_N^{-1} = Q_N^\dagger$.
  → For any $Q_N|x\rangle$, we get $Q_N^\dagger Q_N|x\rangle = Q_N^{-1} Q_N|x\rangle = |x\rangle$.

Phase Parameter of Quantum Fourier State

Finding the Phase of a quantum Fourier state (with Promise)

- ***Input***: The state $\frac{1}{\sqrt{N}}\sum_{y\in\mathbb{Z}_N} e^{2\pi i\phi y}|y\rangle$ for some $\phi \in [0,1)$.
- ***Promise***: There is an integer $x \in \mathbb{Z}_N$ such that $\phi = x/N$.
- ***Problem***: Obtain a good estimate of the phase parameter ϕ.

- Such state can be written as $\frac{1}{\sqrt{N}}\sum_{y\in\mathbb{Z}_N} e^{2\pi i\phi y}|y\rangle = \frac{1}{\sqrt{N}}\sum_{y\in\mathbb{Z}_N} e^{2\pi i\frac{x}{N}y}|y\rangle = Q_N|x\rangle$.

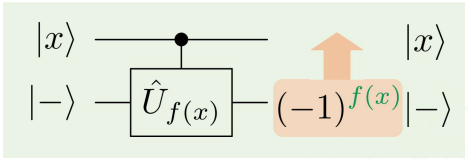
  → $\frac{1}{\sqrt{N}}\sum_{y\in\mathbb{Z}_N} e^{2\pi i\phi y}|y\rangle = Q_N|x\rangle$ — $\boxed{Q_N^{-1}}$ — $|x\rangle$ □

- Questions:
  - How to implement the (inverse) quantum Fourier transform? → *efficient* implementation for $N = 2^n$.
  - How about arbitrary phase $\phi \in [0,1)$ without promise that $\phi = x/N$? → Will talk about it next week.

Phase Estimation Algorithm

Eigenvalue Estimation (1/2)

- In the Deutsch–Jozsa algorithm, we have seen that the quantum oracle U_f can be thought of as a controlled operator c-$\hat{U}_{f(x)}$, where $\hat{U}_{f(x)}: |y\rangle \mapsto |y \oplus f(x)\rangle$.
  We saw that the state $|-\rangle$ is an eigenvector of $\hat{U}_{f(x)}$ with eigenvalue $(-1)^{f(x)}$.
  → The *phase* of the eigenvalue encodes some desired property of interest for f.



- Conversely, suppose we are given a unitary operator U (and a quantum circuits that can efficiently implements it) with an eigenvector $|\psi\rangle$ such that $U|\psi\rangle = e^{2\pi i\phi}|\psi\rangle$
- How to estimate the phase parameter ϕ?

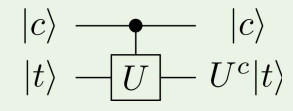
Eigenvalue Estimation (2/2)

Finding Eigenvalues of U (with Promise)

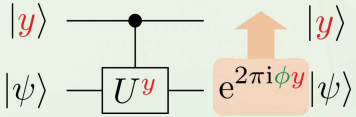
- ***Input***: A quantum circuit implementing an operator U, and an eigenstate $|\psi\rangle$ with the corresponding eigenvalue $e^{2\pi i\phi}$. Namely, $U|\psi\rangle = e^{2\pi i\phi}|\psi\rangle$.
- ***Promise***: There is an integer $x \in \mathbb{Z}_N$ such that $\phi = x/N$.
- ***Problem***: Obtain a good estimate for ϕ.

- Key ingredients: (i) controlled-U gate [§4.3, N&C]; (ii) phase kickback; (iii) QFT.

Qubit c-U:
$|c\rangle|t\rangle \mapsto |c\rangle U^c|t\rangle$

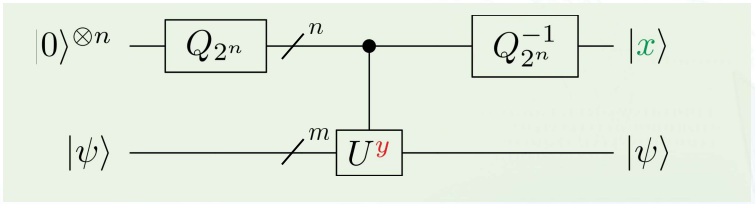


General c-U:
$|y\rangle|\psi\rangle \mapsto |y\rangle U^y|\psi\rangle$



Phase Estimation Algorithm (1/3)

1. For $N = 2^n$, start with $|0\rangle^{\otimes n}\,|\psi\rangle$.
2. Create the uniform superposition state $\frac{1}{\sqrt{2^n}}\sum_{y\in\mathbb{Z}_N}|y\rangle|\psi\rangle$ by applying $H^{\otimes n}$ (or by applying Q_N) to the first n qubits.
3. Apply the c-U^y map $|y\rangle|\psi\rangle \mapsto |y\rangle U^y|\psi\rangle = e^{2\pi i\phi y}|y\rangle|\psi\rangle = e^{2\pi i\frac{x}{N}y}|y\rangle|\psi\rangle$. (Phase kickback)
4. Apply the inverse QFT Q_N^{-1} to the first n qubits and measure the result.

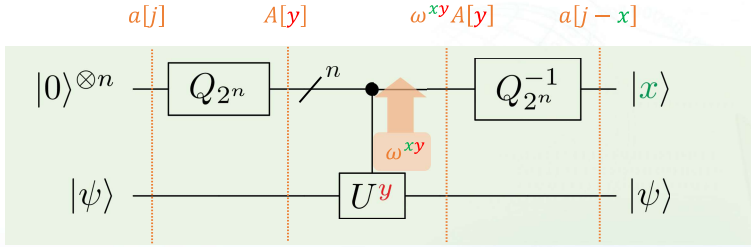


□

Phase Estimation Algorithm (2/3)

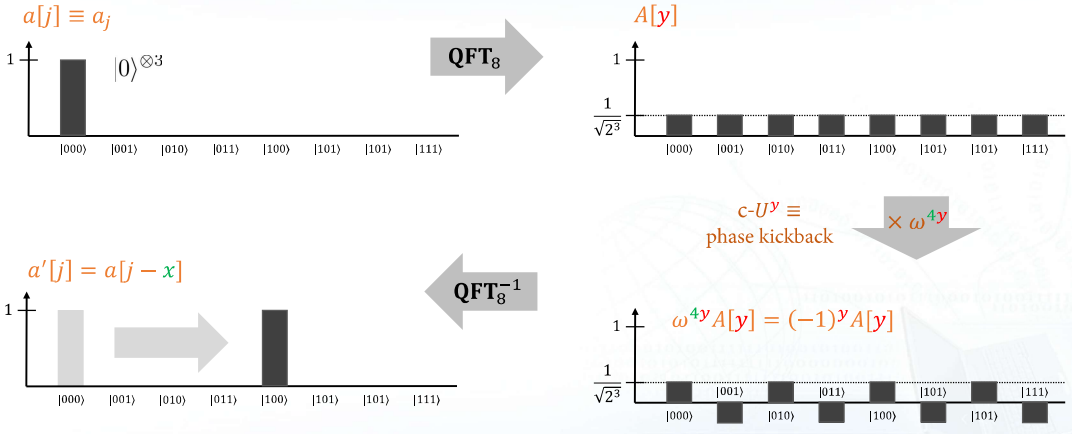
- *Phase estimation algorithm* is actually the *circular shift property* in DFT. (Let $a[j] \equiv a_j$ be the discrete-time signal of the probability amplitudes a_j's.)

$$a[j] \xleftrightarrow{\mathcal{F}} A[y] := \frac{1}{\sqrt{N}}\sum_{j\in\mathbb{Z}_N}\omega^{jy}a[j]$$
$$a[j-x] \xleftrightarrow{\text{time delay}} \omega^{xy}A[y]$$



Phase Estimation Algorithm (3/3)

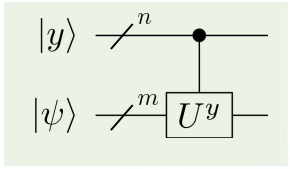
- Example of $n = 3$ and $\phi = x/2^3$ for $x = 4$.



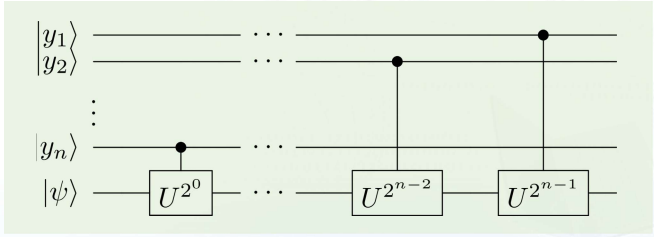
About Controlled Unitary Operations

- Binary representation of $y \in \mathbb{Z}_{2^n}$: $y =: (y_1y_2\cdots y_n) = y_1 2^{n-1} + y_2 2^{n-2} + \cdots + y_n 2^0$.

$$\begin{aligned}\text{c-}U^y : |y\rangle|\psi\rangle &\mapsto |y\rangle U^y|\psi\rangle \\ &= |(y_1\cdots y_n)\rangle U^{(y_1\cdots y_n)}|\psi\rangle \\ &= |y_1\rangle|y_2\rangle\cdots|y_n\rangle U^{y_1 2^{n-1}+y_2 2^{n-2}+\cdots y_n 2^0}|\psi\rangle \\ &= |y_1\rangle|y_2\rangle\cdots|y_n\rangle \left(U^{2^{n-1}}\right)^{y_1}\left(U^{2^{n-2}}\right)^{y_2}\cdots\left(U^{2^0}\right)^{y_n}|\psi\rangle\end{aligned}$$



≡



Period-Finding Algorithm

Periodicity Determination (1/2)

Finding the Period r of a Periodic State Given N

- ***Given***: A black box for a function $f: \mathbb{Z}_N \to Y$.

- ***Promise***: The function f is periodic with some period r, i.e. there is a smallest number r such that $f(x+r) = f(x)$ for all $x \in \mathbb{Z}_N$.

- ***Problem***: Find the period r with some constant level of probability (independent of the size N).

addition mod N

- Example: $f(x) = 7^x \bmod 15$:

Modular exponential function

	7^0	7^1	$7^2 = 49$	$7^3 = 343$	$7^4 = 2401$	7^5	7^6	7^7	7^8
mod 15	1	7	4	13	1	7	4	13	1

Period = 4

Periodicity Determination (2/2)

- We further assume that f is one-to-one in each period, i.e.

$$f(x_1) \neq f(x_2), \forall 0 \leq x_1 < x_2 < r.$$

- Classical solution: $\sqrt{N}$ queries to f are sufficient and necessary.

- Quantum solution: The period can be determined with any constant high level of probability $1-\varepsilon$ using only $O(\log \log N)$ queries and poly($\log N$) further processing steps.

- This is the key ingredient for Shor's factoring algorithm.

Algorithm (1/6)

1. Constructing an uniform superposition $\frac{1}{\sqrt{N}}\sum_{x\in\mathbb{Z}_N}|x\rangle$ by $H^{\otimes n}$, and feed it to the quantum oracle to get $|f\rangle \coloneqq \frac{1}{\sqrt{N}}\sum_{x\in\mathbb{Z}_N}|x\rangle|f(x)\rangle$.
   - *Note*: Since f is periodic, we assume r divides N *exactly*. We set the *integer* $m = N/r$ to be the number of periods.
2. Measure the second register to get some outcome, say y_0, and denote by $x_0 \in [0, r)$ the least integer such that $f(x_0) = y_0$.
3. The first register will be projected into equal superposition of the A different values $x = x_0, x_0 + r, x_0 + 2r, \ldots, x_0 + (m-1)r$ for which $f(x) = y_0$:

$$|\text{per}\rangle = \frac{1}{\sqrt{m}}\sum_{j=0}^{m-1}|x_0 + jr\rangle$$

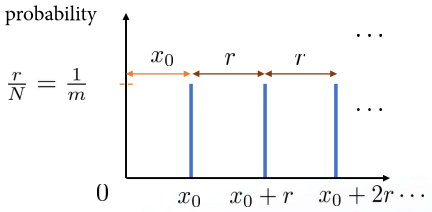
Algorithm (2/6)

x_0 is some random shift

$$|\text{per}\rangle = \frac{1}{\sqrt{m}} \sum_{j=0}^{m-1} |x_0 + jr\rangle$$

- *Note*: $0 \le x_0 < r$ has been chosen at random.
  If we measure it, we would get $x_0 + j_0 r$ where j_0 is uniformly at random.

$\Rightarrow$ Overall we get a random number between 0 and $N-1$, giving no info about r.



However, the state $|\text{per}\rangle$ seems to contain the information of r!

Algorithm (3/6)

You can also apply the inverse QFT

4. Apply $\mathbf{QFT}_N : |x\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{y \in \mathbb{Z}_N} \omega^{xy} |y\rangle$ to $|\text{per}\rangle = \frac{1}{\sqrt{m}} \sum_{j=0}^{m-1} |x_0 + jr\rangle$

$$\Rightarrow \mathbf{QFT}_N |\text{per}\rangle = \frac{1}{\sqrt{Nm}} \sum_{j=0}^{m-1} \left(\sum_{y \in \mathbb{Z}_N} \omega^{(x_0 + jr)y} |y\rangle \right) = \frac{1}{\sqrt{Nm}} \sum_{y=0}^{N-1} \omega^{x_0 y} \left[\sum_{j=0}^{m-1} \omega^{jry} \right] |y\rangle$$

Inspection:

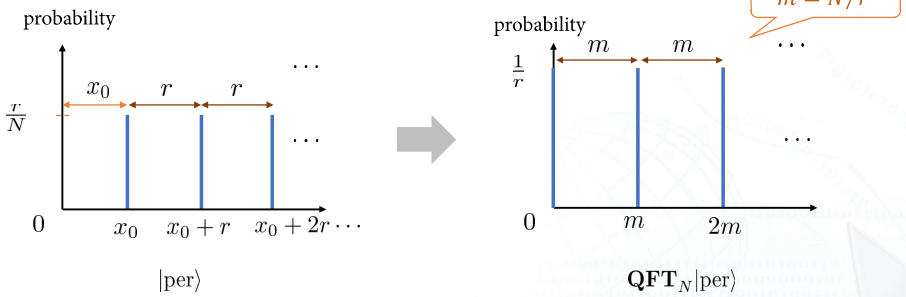
$$\sum_{j=0}^{m-1} \omega^{jry} = \sum_{j=0}^{m-1} \mathrm{e}^{\frac{2\pi \mathrm{i}}{N} jry} = \begin{cases} m & \text{if } y = \frac{\ell N}{r} = \ell m \text{ for } \ell = 0, \ldots, r-1 \\ 0 & \text{otherwise} \end{cases}$$

$\because \forall y \in \mathbb{Z}_N, \mathrm{e}^{\frac{2\pi \mathrm{i}}{N} ry}$ is an m-th root of unity

$$\Rightarrow \mathbf{QFT}_N |\text{per}\rangle = \sqrt{\frac{m}{N}} \sum_{\ell=0}^{r-1} \omega^{x_0 \ell m} |\ell m\rangle = \sqrt{\frac{1}{r}} \sum_{\ell=0}^{r-1} \omega^{x_0 \ell m} |\ell m\rangle$$

Algorithm (4/6)

$$\mathbf{QFT}_N |\text{per}\rangle = \sqrt{\frac{1}{r}} \sum_{\ell=0}^{r-1} \omega^{x_0 \ell m} |\ell m\rangle$$



Algorithm (5/6)

$$\mathbf{QFT}_N |\text{per}\rangle = \sqrt{\frac{1}{r}} \sum_{\ell=0}^{r-1} \omega^{x_0 \ell m} |\ell m\rangle$$

4. Measure it to get some outcome c which is a multiple $c = \ell_0 m = \ell_0 \frac{N}{r}$, where $0 \le \ell_0 \le r-1$ is chosen uniformly at random

$$\Rightarrow \frac{c}{N} = \frac{\ell_0}{r}$$

which is $O(\log^3 N)$ so *efficient*!

5. Calculate the fraction k_0/r (by using the classical *continued fraction* [§A4.4, N&C]) to read off the denominator r to get the desired period provided that ℓ_0 is coprime to r.
6. If k_0 is unfortunately not coprime to r, then we will get a smaller denominator $r' < r$, and hence $f(x) \neq f(x + r')\ \forall x$. But we can check it, e.g. $f(0) = f(r)$?

Algorithm (6/6)

Theorem (Coprimality)

The number of integers less than r that are coprime to r grows as $O(r/\log\log r)$ with increasing r. Hence, if $\ell_0 < r$ is chosen at random,

$$\Pr(\ell_0 \text{ is coprime to } r\,) \approx O((r/\log\log r)/r) = O(1/\log\log r).$$

$\Rightarrow$ If we repeat the whole process $O(\log\log r) \;<\; O(\log\log N)$ times we will obtain a coprime k_0 in at least one case with a constant level of probability. (Here we have used a basic fact in probability theory; see Lemma in Appendix.)

□

Circuit Implementation of QFT

Product Representation (1/5)

$\frac{1}{\sqrt{2^n}}\sum_{y\in\mathbb{Z}_{2^n}} e^{2\pi i\phi y}|y\rangle$ — $\boxed{Q_{2^n}^{-1}}$ — $|\phi\rangle$

- Assume the phase $0 \le \phi < 1$ has an n-bit binary representation:

$$\exists\, x = (x_1 x_2 \cdots x_n) \in \{0, 1, \ldots 2^n - 1\} \text{ s.t. } \phi = \frac{x}{2^n} \equiv \frac{x_1}{2} + \frac{x_2}{2^2} + \cdots \equiv (0.x_1 x_2 \cdots)$$

$$\text{e.g. } (0.011) \equiv \tfrac{0}{2} + \tfrac{1}{2^2} + \tfrac{1}{2^3} \equiv \tfrac{3}{2^3}$$

- Let's see how the inverse QFT can extract the phase ϕ from $\frac{1}{\sqrt{2^n}}\sum_{y\in\mathbb{Z}_{2^n}} e^{2\pi i\phi y}|y\rangle$.
- If the input is a 1-qubit state (so $n = 1$) and $\omega = (0.x_1)$. Then the input is

$$\frac{1}{\sqrt{2}}\sum_{y=0}^{1} e^{2\pi i(0.x_1)y}|y\rangle = \frac{1}{\sqrt{2}}\sum_{y=0}^{1} e^{2\pi i\frac{x_1}{2}y}|y\rangle = \frac{1}{\sqrt{2}}\sum_{y=0}^{1}(-1)^{x_1 y}|y\rangle = \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_1}|1\rangle)$$

$$\Rightarrow H\left(\frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_1}|1\rangle)\right) = |x_1\rangle \qquad \mathbf{QFT}_{2^1} = \mathbf{QFT}_{2^1}^{-1} = H$$

Product Representation (2/5)

- Suppose we have a 2-qubit state $\frac{1}{\sqrt{2^2}}\sum_{y\in\mathbb{Z}_{2^2}} e^{2\pi i\phi y}|y\rangle$ and $\phi = (0.x_1x_2)$.

$$\frac{1}{\sqrt{2^2}}\sum_{y=0}^{2^2-1} e^{2\pi i(0.x_1x_2)y}|y\rangle = \frac{1}{\sqrt{2^2}}\left(|0\rangle + e^{2\pi i(0.x_1x_2)}|1\rangle + e^{2\pi i(0.x_1x_2)2}|2\rangle + e^{2\pi i(0.x_1x_2)3}|3\rangle\right)$$

$e^{2\pi i(x_1.x_2)}|2\rangle$

$e^{2\pi i(0.x_2)}|2\rangle$ $\qquad$ $e^{2\pi i(0.x_2)}e^{2\pi i(0.x_1x_2)}|3\rangle$

$$= \left(\frac{|0\rangle + e^{2\pi i(0.x_2)}|1\rangle}{\sqrt{2}}\right) \otimes \left(\frac{|0\rangle + e^{2\pi i(0.x_1x_2)}|1\rangle}{\sqrt{2}}\right)$$

- Note that x_2 can be determined from the first qubit by applying Hadamard gate H. For the second qubit, if $x_2 = 0$, we can determine x_1 by applying H again.

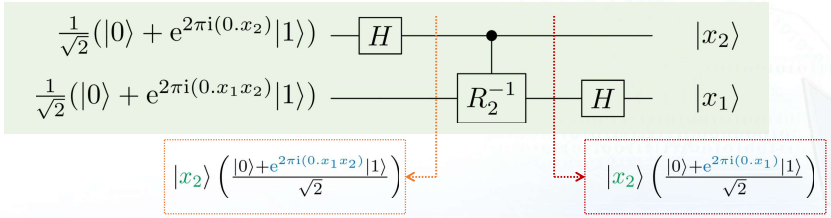
Product Representation (3/5)

- If $x_2 = 1$, the phase becomes $(0.x_1 1)$. We need to cancel $\mathrm{e}^{2\pi\mathrm{i}(0.01)}$ before H.
  This can be done by using the rotation Z gate: $R_k := \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{\frac{2\pi\mathrm{i}}{2^k}} \end{pmatrix}$ Phase rotation gate

$$\Rightarrow R_2^{-1}\left(\frac{|0\rangle + \mathrm{e}^{2\pi\mathrm{i}(0.x_1 1)}|1\rangle}{\sqrt{2}}\right) = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{-2\pi\mathrm{i}(0.01)} \end{pmatrix}\left(\frac{|0\rangle + \mathrm{e}^{2\pi\mathrm{i}(0.x_1 1)}|1\rangle}{\sqrt{2}}\right)$$

$$= \left(\frac{|0\rangle + \mathrm{e}^{2\pi\mathrm{i}(0.x_1)}|1\rangle}{\sqrt{2}}\right) \overset{H}{\mapsto} |x_1\rangle$$

Controlled-rotation



Product Representation (4/5)

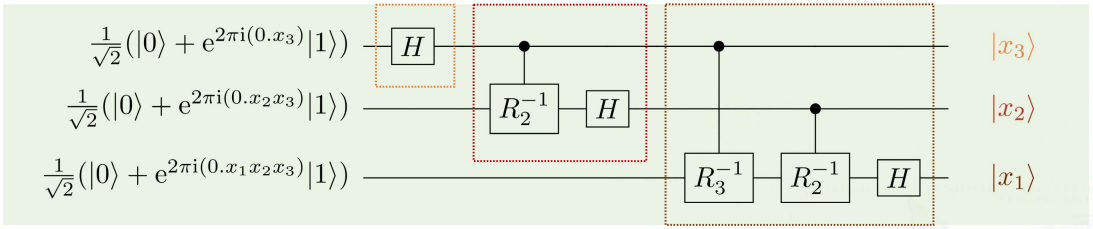
- For a general n-qubit quantum Fourier state with phase $\phi = (0.x_1x_2\cdots x_n)$.

$$\begin{aligned}
\frac{1}{\sqrt{2^n}}\sum_{y=0}^{2^n-1}\mathrm{e}^{2\pi\mathrm{i}\phi y}|y\rangle &= \frac{1}{\sqrt{2^n}}\sum_{y=0}^{2^n-1}\mathrm{e}^{2\pi\mathrm{i}\frac{x}{2^n}y}|y\rangle \\
&= \frac{1}{\sqrt{2^n}}\sum_{y=0}^{2^n-1}\mathrm{e}^{2\pi\mathrm{i}x(\sum_{\ell=1}^{n}y_\ell 2^{-\ell})}|y\rangle \\
&= \frac{1}{\sqrt{2^n}}\sum_{y=0}^{2^n-1}\prod_{\ell=1}^{n}\mathrm{e}^{2\pi\mathrm{i}y_\ell x 2^{-\ell}}|y_1\cdots y_n\rangle \\
&= \bigotimes_{\ell=1}^{n}\frac{1}{\sqrt{2}}\left(|0\rangle + \mathrm{e}^{2\pi\mathrm{i}x2^{-\ell}}|1\rangle\right) \\
&= \bigotimes_{\ell=1}^{n}\frac{1}{\sqrt{2}}\left(|0\rangle + \mathrm{e}^{2\pi\mathrm{i}(0.x_{n-\ell+1}\cdots x_n)}|1\rangle\right)
\end{aligned}$$

Product Representation (5/5)

- As an example of $n = 3$: $\frac{1}{\sqrt{2^3}}\sum_{y=0}^{2^3-1}\mathrm{e}^{2\pi\mathrm{i}(0.x_1x_2x_3)y}|y\rangle$

$$= \left(\frac{|0\rangle + \mathrm{e}^{2\pi\mathrm{i}(0.x_3)}|1\rangle}{\sqrt{2}}\right)\otimes\left(\frac{|0\rangle + \mathrm{e}^{2\pi\mathrm{i}(0.x_2x_3)}|1\rangle}{\sqrt{2}}\right)\otimes\left(\frac{|0\rangle + \mathrm{e}^{2\pi\mathrm{i}(0.x_1x_2x_3)}|1\rangle}{\sqrt{2}}\right)$$

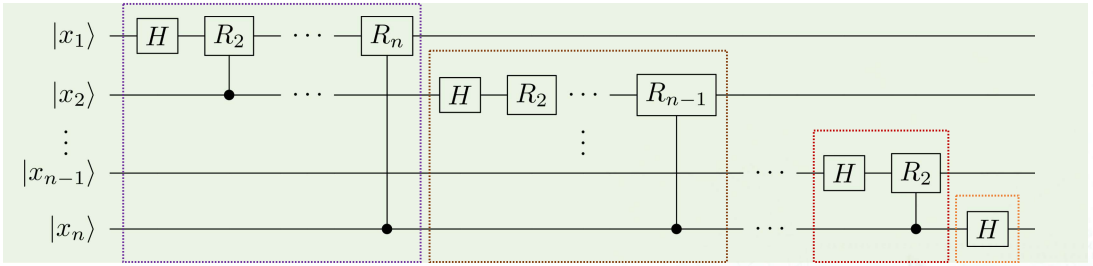


$$\begin{cases} R_k^{-1}|0\rangle = |0\rangle \\ R_k^{-1}|1\rangle = \mathrm{e}^{-2\pi\mathrm{i}(0.0\cdots01)}|1\rangle \end{cases} \quad (k)$$

Circuit Implementation of QFT

- General n-qubit quantum Fourier transform: $|\phi\rangle$ — $\mathbf{QFT}_{2^n}$ — $\frac{1}{\sqrt{2^n}}\sum_{y=0}^{2^n-1}\mathrm{e}^{2\pi\mathrm{i}\phi y}|y\rangle$

$$|\phi\rangle = |x2^{-n}\rangle = |0.x_1\cdots x_n\rangle \equiv |x_1\cdots x_n\rangle$$



- Number of gates required is $1 + 2 + \cdots + n - 1 + n = \frac{1}{2}n(n+1) = O(n^2)$, while complexity of an n-bit FFT is $O(n2^n)$ (i.e. the divide-and-conquer algorithm).

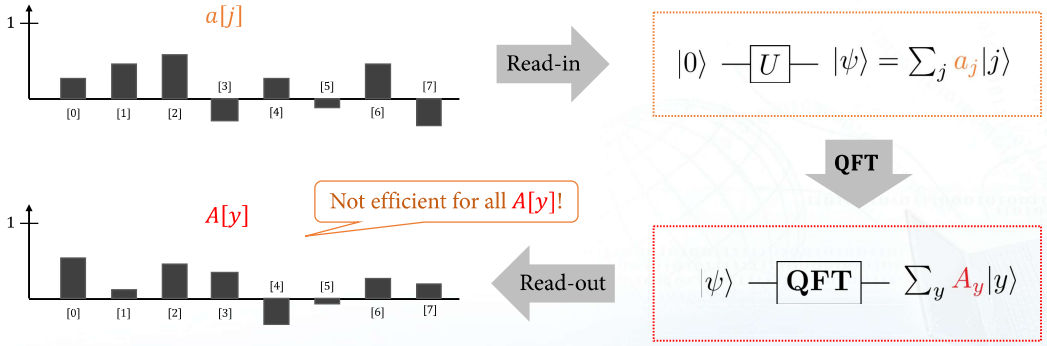
Concluding Remarks

Concluding Remarks

- In PE algorithm, what if we do not know an eigenvector of U? We may write an arbitrary state $|\varphi\rangle = \sum_j \alpha_j |\psi_j\rangle$ over eigenvectors $\{|\psi_j\rangle\}$. Then we will get an estimate of each eigenvalue ϕ_j with probability $|\alpha_j|^2$.
- We needed to apply $U^{2^m}, 0 \le m \le n-1$. If we are given U as a black box, it may be expensive. However, if we have an explicit circuit for U, we may be able to find a more efficient way of computing U^{2^m}.
- When $N = 2^n$, QFT can be efficiently implemented by a quantum circuit using $O(n^2)$ gates and $2n + 1 = O(n)$ depth.
  For large k, R_k is very close to identity gate. Hence, the complexity can be further improved to $O(n \log n)$ by allowing some error (i.e. approximate QFT).
- For N is not a power of 2, we can approximate $\mathbf{QFT}_N$ by $\mathbf{QFT}_{2^k}$ for $2^k \approx N$.
  This will incur only a small error to the whole algorithm.

After Thought

- Can we replace conventional DFT (or FFT) by just using QFT?
  Note that in the DFT regimes, the discrete-time signals are stored in arrays if RAM.
  But in QFT, they are only *probability amplitudes* of a quantum state.



Appendix

A Lemma Used in Period-Finding

Lemma

If a single trial has success probability p and we repeat the trial M times independently, then for any constant $0 < \varepsilon < 1$:

$$\Pr(\text{at leat one success in } M \text{ trials}) > 1 - \varepsilon \text{ if } M = \frac{-\log \varepsilon}{p}.$$

Hence to achieve any constant level $1 - \varepsilon$ of success probability, $O(1/p)$ trials suffice.

Proof.
1. Probability of at least one success in M runs is $1 - (1-p)^M$.
2. Then $1 - (1-p)^M = 1 - \varepsilon$ if $M = \frac{-\log \varepsilon}{-\log(1-p)}$.
3. Using the fact $p < -\log(1-p)\ \forall 0 < p < 1$. We get $M < \frac{-\log \varepsilon}{p} = O(1/p)$.

□

References

- A. Y. Kitaev, "Quantum measurements and the Abelian stabilizer problem," arXiv: quantph/9511026, 1995.
- R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, "Quantum algorithms revisited," In *Proceedings of the Royal Society of London*, A454339–354, 1998.
- D. Coppersmith, "An approximate Fourier transform useful in quantum factoring," *IBM Research Report*, No. RC19642, 1994.
- G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, New York, fifth edition, 1979.
- S. J. Hallgren. *Quantum Fourier Sampling, the Hidden subgroup Problem, and Beyond.* Ph.D. Thesis, University of California, Berkeley, 2000.
- M. Mosca and A. Ekert, "The hidden subgroup problem and eigenvalue estimation on a quantum computer," In *Proceedings of 1st NASA QCQC conference*, 1998.
- A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi. *Classical and quantum computation*, AMS, 2002.

Quantum Information and Computation
Factoring Algorithm

Hao-Chung Cheng (鄭皓中)

haochung@ntu.edu.tw

Department of Electrical Engineering

National Taiwan University

April 7, 2021

Outline

Motivation

Motivation

- So far we have seen a few *computational speedups* via quantum algorithms.
  - Deutsch–Jozsa algorithm: exponential speedup in the deterministic scenario.
  - Vazirani–Bernstein algorithm: polytime speedup in the bounded-error scenario.
  - Simon's algorithm: exponential speedup in the bounded-error scenario.
  - Grover's algorithm: quadratic speedup in the unstructured search problem.

  All in the query model
- *Goal*: To find an *efficient* quantum algorithm classically-hard & *real-life* problems.
- Given an integer N with $n = \log N$ digits, Shor's algorithm runs in polytime $O(n^3)$ to output a factor, while the best known classical algorithm (number field sieve algorithm) runs in time $e^{O(n^{1/3}\log^{2/3} n)}$. → This could break the RSA cryptosystem.
- Other cryptosystem such as the *Diffie–Hellman key exchange security* (based on the hardness of the *discrete logarithm problem*) and the *Elliptic curve cryptography* can be broke in polytime by applying Shor's idea.

← Major potential application of quantum computers

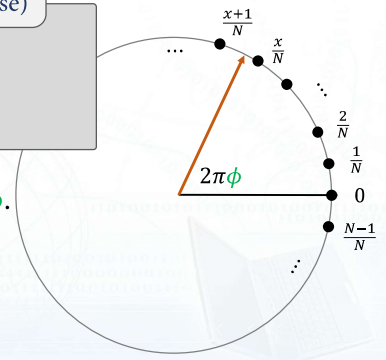
Phase Estimate (Revisited)

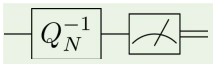
Phase Parameter of Quantum Fourier State

- Last week, we have seen how to estimate $\phi = x/N$ for some $x \in \mathbb{Z}_N$ from the state $\frac{1}{\sqrt{N}}\sum_{y\in\mathbb{Z}_N} e^{2\pi i\phi y}|y\rangle$ *with certainty*. How about arbitrary $\phi \in [0,1)$?

Finding the Phase of a Quantum Fourier State (Arbitrary Phase)

- ***Input***: The state $\frac{1}{\sqrt{N}}\sum_{y\in\mathbb{Z}_N} e^{2\pi i\phi y}|y\rangle$ for some $\phi \in [0,1)$.
- ***Problem***: Obtain a good estimate of the phase parameter ϕ.



- *Claim*: Let $\hat{\phi} = \frac{\hat{x}}{N}$ be an integer multiple of $\frac{1}{N}$ closest to ϕ.
  Applying  returns $\hat{x}$ with prob. $\geq \frac{4}{\pi^2}$.

$\Rightarrow \Pr\left(\left|\phi - \frac{\hat{x}}{N}\right| \leq \frac{1}{2N}\right) \geq \frac{4}{\pi^2}$.

Error Analysis (1/2)

1. Recall that $Q_N^{-1}: |y\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{x \in \mathbb{Z}_N} \omega^{-xy} |x\rangle$

$$\Rightarrow Q_N^{-1} \left(\frac{1}{\sqrt{N}} \sum_{y \in \mathbb{Z}_N} \mathrm{e}^{2\pi \mathrm{i} \phi} |y\rangle \right) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \sum_{y \in \mathbb{Z}_N} \mathrm{e}^{2\pi \mathrm{i} \frac{xy}{N}} \mathrm{e}^{2\pi \mathrm{i} \phi} |x\rangle$$

$$= \sum_{x \in \mathbb{Z}_N} \left[\frac{1}{N} \sum_{y \in \mathbb{Z}_N} \mathrm{e}^{2\pi \mathrm{i} y \left(\phi - \frac{x}{N}\right)} \right] |x\rangle$$

$$=: \sum_{x \in \mathbb{Z}_N} \alpha_x |x\rangle$$

2. Probability of getting x:

$$|\alpha_x|^2 = \frac{1}{N^2} \left| \frac{1 - \mathrm{e}^{2\pi \mathrm{i} y (\phi - \frac{x}{N}) N}}{1 - \mathrm{e}^{2\pi \mathrm{i} y (\phi - \frac{x}{N})}} \right| = \frac{1}{N^2} \left(\frac{\sin N\pi(\phi - \frac{x}{N})}{\sin \pi(\phi - \frac{x}{N})} \right)^2$$

$|1 - e^{i2\theta}| = |e^{-i\theta} - e^{i\theta}| = 2|\sin\theta|$

A Lemma for Estimating Probabilities

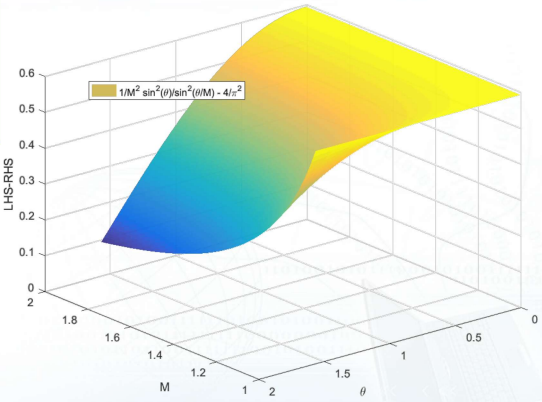
Lemma

For any $M \geq 1$ and any $|\theta| \leq \frac{\pi}{2M}$,
then $\frac{1}{M^2} \frac{\sin^2(M\theta)}{\sin^2(\theta)} \geq \frac{4}{\pi^2}$.

We omit the proof.

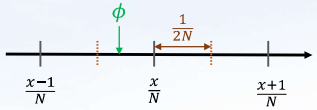
Note that $\frac{\sin^2(M\theta)}{\sin^2(\theta)}$ is decreasing in $|\theta|$.

(This is not examinable.)



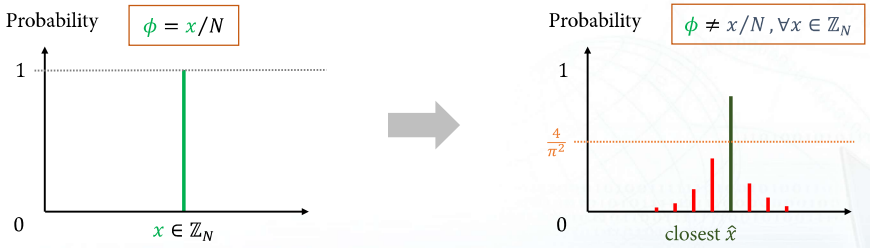
Error Analysis (2/2)

3. Lower bound for probability of getting closest x:



$$|\alpha_x|^2 = \frac{1}{N^2} \left(\frac{\sin N\pi(\phi - \frac{x}{N})}{\sin \pi(\phi - \frac{x}{N})} \right)^2 \geq \frac{4}{\pi^2} \quad \text{if} \quad \left| \phi - \frac{x}{N} \right| \leq \frac{1}{2N}$$

Only closest $\hat{x}$ satisfying it



Period-Finding (Revisited)

Periodicity Determination (1/3)

Previously we consider modular exponential function $x \mapsto a^x \bmod N$

- *Previous on estimating periods of a periodic state*:
  - Given a periodic state $|\text{per}\rangle = \frac{1}{\sqrt{m}}\sum_{j\in\mathbb{Z}_m}|x_0 + jr\rangle$ for which we are given that $N = mr$.
  - By applying $\mathbf{QFT}_{mr}$, we get equally likely measurement outcomes $\ell m \in \mathbb{Z}_{mr}$, $\ell \in \mathbb{Z}_r$.
- Let M (e.g. $M = 2^t$) be the dimension of our working space. Because of the random shift x_0, there might be truncation on the right-most label. → We don't know m!

  Moreover, it is likely that $2^t/r \notin \mathbb{N}$. We will deal with this in the following.

Finding the Period r of a Periodic State ($2^t/r \notin \mathbb{N}$)

- ***Given***: A periodic state $|\text{per}\rangle = \frac{1}{\sqrt{m_{x_0}}}\sum_{j:0\le x_0+jr<2^t-1}|x_0 + jr\rangle$.
- ***Problem***: Find the period r with some constant level of probability (independent of the size 2^t).

$x_0 \in \mathbb{Z}_r$ is an unknown shift

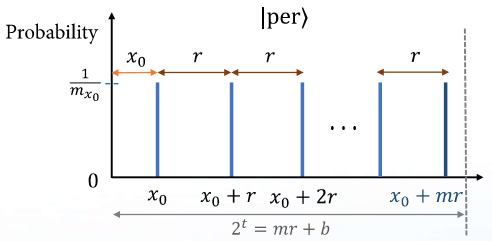
Periodicity Determination (2/3)

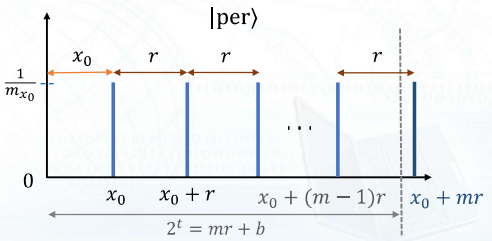
The point of this slide is that we don't know what m_{x_0} is.

$$|\text{per}\rangle = \frac{1}{\sqrt{m_{x_0}}}\sum_{j:0\le x_0+jr<2^t-1}|x_0 + jr\rangle$$

- Since $2^t/r \notin \mathbb{N}$, we let $2^t = mr + b$, i.e. $b = 2^t \bmod r$.

$$m_{x_0} = m + 1 = \left\lfloor \frac{2^t}{r} \right\rfloor + 1, \quad x_0 < b \qquad m_{x_0} = m = \left\lfloor \frac{2^t}{r} \right\rfloor, \quad x_0 \ge b$$





Periodicity Determination (3/3)

- We aim to show that the measurement outcome will yield integers y which is *close to* multiples of $2^t/r$ with suitably high probability.

Theorem (Period-Finding)

Suppose we measure the labels in $\mathbf{QFT}_{2^t}|\text{per}\rangle$. For every $\ell \in \mathbb{Z}_r$, let y be the unique integer with $\left|y - \ell\frac{2^t}{r}\right| < \frac{1}{2}$. Then $\Pr(y) \ge \frac{\gamma}{r}$ where $\gamma \approx \frac{4}{\pi^2}$.

- Proof strategy for finding the period r:
  - Standard analysis as what we did in finding phases of a quantum Fourier state (via the Lemma).
  - Using *continued fraction* on $y/2^t$ to find ℓ/r.
  - Choosing t such that $2^t \ge N^2$ to guarantee that there's only one such fraction ℓ/r within $1/2^{t+1}$.
  - Showing that performing continued fraction in time $O(t^3) = O(\log^3 N)$.

which is efficient!

Error Analysis (1/2)

- Note that x_0 is just a random shift. Thus the normalization const. m_{x_0} isn't important.
- What really matters is to calculate $\mathbf{QFT}_{2^t}|\text{per}\rangle$.
- Apply $\mathbf{QFT}_{2^t} : |x\rangle \mapsto \frac{1}{\sqrt{2^t}}\sum_{y\in\mathbb{Z}_{2^t}}\omega^{xy}|y\rangle$ to $|\text{per}\rangle = \frac{1}{\sqrt{m_{x_0}}}\sum_j |x_0 + jr\rangle$

$$\Rightarrow \mathbf{QFT}_{2^t}|\text{per}\rangle = \frac{1}{\sqrt{2^t m}}\sum_j\left(\sum_{y\in\mathbb{Z}_{2^t}}\omega^{(x_0+jr)y}|y\rangle\right) = \sum_{y\in\mathbb{Z}_{2^t}}\omega^{x_0 y}\frac{1}{\sqrt{2^t m_{x_0}}}\left[\sum_j \omega^{jry}\right]|y\rangle$$

$$=: \sum_{y\in\mathbb{Z}_{2^t}}\alpha_y|y\rangle$$

Inspection: $|\alpha_y|^2 = \frac{1}{2^t m_{x_0}}\left|\frac{1 - e^{2\pi i \frac{r}{2^t} y m_{x_0}}}{1 - e^{2\pi i \frac{r}{2^t} y}}\right|^2$

If $2^t/r \in \mathbb{N}$, then $|\alpha_y|^2 = 1/r$ for $y = \ell\frac{2^t}{r} \in \mathbb{Z}_N$, $\ell \in \mathbb{Z}_r$ as we did last week. But here $2^t/r \notin \mathbb{N}$!

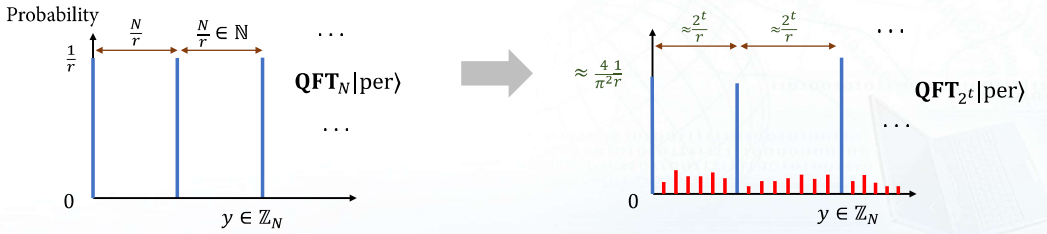
Error Analysis (2/2)

- Using the *Lemma* (as in finding phases of a quantum Fourier state) :

$$|\alpha_y|^2 = \frac{1}{2^t m_{x_0}} \left| \frac{1 - e^{2\pi i \frac{r}{2^t} y m_{x_0}}}{1 - e^{2\pi i \frac{r}{2^t} y}} \right|^2 \geq \frac{m_{x_0}}{2^t} \frac{4}{\pi^2} \approx \frac{1}{r} \frac{4}{\pi^2} \quad \text{if} \quad \left| \frac{r}{2^t} y - \ell \right| \leq \frac{1}{2m_{x_0}}$$

$\ell \in \mathbb{Z}_r$

$$\iff \boxed{\left| \frac{y}{2^t} - \frac{\ell}{r} \right| \leq \frac{1}{2^{t+1}}}$$



How to choose the qubit number t?

$$\boxed{\left| \frac{y}{2^t} - \frac{\ell}{r} \right| \leq \frac{1}{2^{t+1}}}$$

- *Claim*. Choosing t to be the least integer such that $2^t \geq N^2$ is sufficient.
  → So that ℓ/r for some $\ell \in \mathbb{Z}_r$ is *uniquely* determined from $y/2^t$.

1. Recall that $r < N$ and $2^t \geq N^2$. Hence, $\left| \frac{y}{2^t} - \frac{\ell}{r} \right| \leq \frac{1}{2N^2}$ with $r < N$.
2. We prove the claim via contradiction.
   So suppose there are ℓ_1/r_1 and ℓ_2/r_2 both lie within $1/(2N^2)$ of $y/2^t$.

$$\left| \frac{\ell_1}{r_1} - \frac{\ell_2}{r_2} \right| = \frac{|\ell_1 r_2 - r_1 \ell_2|}{r_1 r_2} \geq \frac{1}{r_1 r_2} > \frac{1}{N^2}$$

3. But ℓ_1/r_1 and ℓ_2/r_2 both lie within $1/(2N^2)$ of $y/2^t$, so they must be within $1/N^2$ of each other, contradicting the above inequality.

□

Continued Fraction (1/2) [§Appendix4.3, N&C]

- We have known that ℓ/r is *uniquely* determined by $y/2^t$.
  But how do we actually compute ℓ/r from $y/2^t$?
- *Continued fractions*:

$$\frac{y}{2^t} = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \cdots}}}}$$

We call $[a_1, a_2, a_3, a_4, \ldots]$ the *convergent* of the continued fraction

Theorem 5.1 in [N&C]

Let $0 < x < 1$ be a (ir)rational number and suppose that p/q is a rational number such that

$$\left| x - \frac{p}{q} \right| < \frac{1}{2q^2}.$$

Then p/q is a convergent of the continued fraction of x.

Continued Fraction (2/2) [§Appendix4.3, N&C]

- We have known that there is a unique ℓ/r satisfying $\left| \frac{y}{2^t} - \frac{\ell}{r} \right| \leq \frac{1}{2N^2}$.

  Then by Thm. 5.1 such ℓ/r must be a convergent of the continued fraction of $y/2^t$.

Box 5.3 in [N&C]

Consider the continued fraction $s/t = [a_1, \ldots, a_l]$. Let $p_k/q_k = [a_1, \ldots, a_k]$ be the k-th convergent for $k = 1, \ldots, l$.

If s and t (cancelled to lowest terms) are n-bit integers, then the length l of the continued fraction is $O(n)$ and this continued fraction together with its convergents can be calculated in time $O(n^3)$.

- Since $2^t = O(N^2)$, we have that y and 2^t are $O(n)$ bit integers and hence all the convergents can be calculated in time $O(n^3)$.

$n := \log N$

An Example of The Continued Fraction

- Consider the modular exponential function $f(x) = 7^x \bmod N$ with $N = 39$. N^2=1521 and $2^{10} < N^2 < 2^{11} = 2048 = 2^t$. So choose $t = 11$.

Here, $r < N = 39$

1. Suppose the measurement outcome of $\mathbf{QFT}_{2^t}$ yields $y = 853$. According to the previous theory, there is a unique fraction ℓ/r such that

$$\Rightarrow \left| \frac{853}{2048} - \frac{\ell}{r} \right| < \frac{1}{2^{m+1}} \leq \frac{1}{2N^2} = 0.000329.$$

2. Continued fraction: $\dfrac{853}{2048} = \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{42 + \cfrac{1}{4 + 0}}}}} = [2, 2, 2, 42, 4]$

$\Rightarrow$ The convergents are $[2] = \frac{1}{2}$; $[2,2] = \frac{2}{5}$; $[2,2,2] = \frac{5}{12}$; $[2,2,2,42] = \frac{212}{509}$

Shor's Factoring Algorithm

Integer Factorization – Reduction (1/3)

Integer Factorization Problem

- ***Given***: An integer N.
- ***Problem***: Output positive integers $p_1, p_2, \ldots, p_l, r_1, r_2, \ldots, r_l$ where p_i are distinct primes and $N = p_1^{r_1} p_2^{r_2} \cdots p_l^{r_l}$.

One can efficiently factor any integer that is a prime power – *AKS Primality Test* in time $O(n^3)$.

$\Rightarrow$ It can be reduced to $O(\log N)$ instances of the following problem

Splitting an Odd Non-Prime-Power Integer

- ***Given***: An odd integer N that has at least two distinct prime factors.
- ***Problem***: Output two integers N_1 and N_2 such that $N = N_1 \times N_2$.

Integer Factorization – Reduction (2/3)

- Miller showed in 1975 that the problem of splitting integers *reduces probabilistically* to the problem of *order finding*.

Order-Finding Problem

- ***Given***: Integer a and N such that $\gcd(a, N) = 1$ (i.e. a is coprime to N).
- ***Problem***: Find the *order* of a modulo N.

The gcd of two numbers can be found efficiently using the *Extended Euclidean Algorithm* in time $O(\log^2 N)$.

Euler's Theorem

If a and N are coprime, then there is a *least* power $1 < r < N$ such that $a^r = 1 \bmod N$, where such r is called the *order* of a mod N.

The *modular exponential function* with order r is also *periodic* with period r.

$\Rightarrow$ The *order-finding problem* is equivalent the *period-finding problem*.

$a^{x+r} = a^x \cdot a^r = a^x \cdot 1$

Integer Factorization – Reduction (3/3)

- Factoring any integer
  ⇓ deterministic classical polytime reduction
- Splitting odd non-prime-power N
  ⇓ probabilistic classical polytime reduction
- Finding the orders of integers modulo N
  ⇓ probabilistic classical polytime reduction
- Sampling Estimates to a random integer multiple of $1/r$ (where r is the order of some integer a mod N).

Write $a \Rightarrow b$ to mean that problem a reduces to a problem b.

The Sampling Estimate Problem is where quantum algorithm is used.

Sampling Estimates to Random Integer Multiple of $1/r$

- ***Given***: Integer a and N such that $\gcd(a, N) = 1$. Let r denote the (unknown) order of a.
- ***Problem***: Output a number $y \in \mathbb{Z}_{2^t}$ such that for each $\ell \in \mathbb{Z}_r$ we have $\Pr\left(\left|\frac{y}{2^t} - \frac{\ell}{r}\right| \leq \frac{1}{2r^2}\right) \geq \frac{c}{r}$ for some constant $c > 0$.

How to Find Factors of an Integer? (1/2)

1. Suppose we can find the order r (by the order/period-finding algorithm) of a certain modular exponential function $a^r = 1 \bmod N$ (where how a is chosen will be determined later), and suppose r comes out to be *even*.

   $$a^r - 1 = (a^{r/2} - 1)(a^{r/2} + 1) = 0 \bmod N$$

   ⇒ N exactly divides the product $(a^{r/2} - 1)(a^{r/2} + 1)$.

   Knowing r we can calculate (via *repeated squaring*) each of two terms in $O(n^3)$ time.

2. We know N does not divide $(a^{r/2} - 1)$ since r was the *order*.
3. If N does not divide $(a^{r/2} + 1)$, then N must *partly* divide into $(a^{r/2} - 1)$, and *partly* divide into $(a^{r/2} - 1)$.

   One can run EEA for gcd in time $O(n^2)$.

   ⇒ Compute $\gcd(a^{r/2} - 1, N)$ and $\gcd(a^{r/2} + 1, N)$ which are factors of N.

- Example: $f(x) = 7^x \bmod 15$ has $r = 4$. Hence, $7^4 - 1 = (7^2 - 1)(7^2 + 1) = (48)(50)$
  ⇒ $\gcd(48, 15) = 3$ and $\gcd(50, 15) = 5$ gives factors of 15.

How to Find Factors of an Integer? (2/2)

- How likely is that r is even and $a^{r/2} \neq -1 \bmod N$? How to choose a?

Theorem 5.3 in [N&C]

Suppose N is odd and not a power of a prime. If $a < N$ is chosen uniformly at random with $\gcd(a, N) = 1$, then $\Pr(r \text{ is even } \& \ a^{r/2} \neq -1 \bmod N) \geq 1/2$.

- Given any candidate factor, we can check it (in $O(n^2)$ time) by test division into N.
  Repeating the process, say 10 times, we will *fail* to get a factor only exponentially small probability $1/2^{10}$.

⇒ Only a constant number of values a needed to be tried with success probability $O(1)$.

The Factoring Algorithm (Given N)

1. Is N even? If so, output 2 and stop.
2. Use the classical polynomial time algorithm to determine if N is a prime power.
3. If N is neither even nor a prime power, choose $1 < a < N$ at random and compute $\gcd(a, N)$. If the number is not 1 (luckily), then output it and stop.
4. Find the order/period of $f(x) = a^x \bmod N$ (with any desired level of constant probability using the *order/period-finding algorithm*).

   where the quantum part – Shor's order-finding algorithm – comes in

5. If r is odd, go back to 3. If r is even, compute $p = \gcd(a^{r/2} + 1, N)$.
   If $p = 1$ or N, go back to 3. and try again. Otherwise, p is a factor of N. Output t.

Recall Shor's Order/Period-Finding Algorithm

1. Create the states $\sum_{x\in\mathbb{Z}_{2^t}} \frac{1}{\sqrt{2^t}} |x\rangle|a^x \bmod N\rangle$.
2. Measure the second (output) register gives $|\text{per}\rangle = \frac{1}{\sqrt{m_{x_0}}} \sum_{j:0\le x_0+jr<2^t-1} |x_0 + jr\rangle$. (Here $x_0 \in \mathbb{Z}_r$ is a random shift and m_{x_0} is a normalization constant.)

   If we were able to implement $\mathbf{QFT}_{m_{x_0} r}$, then we can get $\sum_{\ell\in\mathbb{Z}_r} e^{2\pi i \frac{x_0}{r}\ell} |\ell m_{x_0}\rangle$.
   So we measure values y such that $\frac{y}{r m_{x_0}} = \frac{\ell}{r}$. However, we don't know r and m_{x_0}.
3. Apply $\mathbf{QFT}_{2^t}$ (can also do the inverse QFT) to $|\text{per}\rangle$ and measure values $y \in \mathbb{Z}_{2^t}$.
4. Use continued fraction algorithm to find the convergent of ℓ/r as being within $1/2^{t+1}$ of $y/2^t$. Then to read-off the denominator $r < N$.
5. Check if the obtained r satisfies $a^r = 1 \bmod N$. If not, go back to 1.
   Repeating the whole process $O(\log\log N)$ times give the desired r with prob. $O(1)$.

Complexity of Shor's Order-Finding Algorithm

- (i) Prepare the periodic state; (ii) QFT; (iii) single qubit measurement.

1. We need to compute the function $\frac{1}{\sqrt{2^t}} \sum_{x\in\mathbb{Z}_{2^t}} |x\rangle|f(x)\rangle$ for $f(x) = a^x \bmod N$.
   - Uniform superposition requires $t = O(n)$ Hadamard gates.
   - Computing the function $f(x)$ via repeated squaring $O(\log x) = O(t) = O(n)$ & faster multiplication $O(n \log n \log\log n)$ requires $O(n^2 \log n \log\log n)$ operations in total.

   The oracle f is *efficiently* computable!
2. As mentioned last week, the QFT modulo 2^t can be efficiently implemented in $O(m^2) = O(n^2)$ steps. Also, single qubit measurement runs in time $O(n)$.
3. To get an ℓ which is coprime to r, we need to run the above process by $O(\log\log N) = O(\log n)$ repetitions to achieve a constant level of probability.

The above totally requires $O(\log n) \cdot (O(n) + O(n^2 \log n \log\log n) + O(n^2) + O(n)) = O(n^2 \log^2 n \log\log n)$.

4. To read r from ℓ/r we use classical continued fraction algorithm in $O(n^3)$ steps.

Concluding Remarks

Concluding Remarks

- With Shor's algorithm, the integer factorization problem becomes **BQP** (but we don't know if it's in **P** yet since the Miller's reduction might not be the best choice.
- The main bottleneck of the algorithm is the *classical* part – implementing modular exponentiation and the continued fraction in $O(n^3)$ time.
  → see the circuit implementation (e.g. the controlled multiplier gate) [Beauregard' 03].
- We need at least $t \ge 2n$ qubits to implement Shor's algorithm.
  (The classical record so far is 829-bit. ⇒ We would need at least 1658 qubits!
- Shor's algorithm (using QFT) can be unified as the *phase estimation algorithm*. [Appendix]
- The periodicity determination problem can be generalized from the finite field $\mathbb{Z}_N$ to arbitrary group as the so-called *hidden subgroup problem* [§5.4.3, N&C]
  (though there are no *efficient* quantum algorithms for solving most *non-Abelian* hidden subgroup problems yet such as the legendary *graph isomorphism* problem).

Appendix

Kitaev's Order-Finding Algorithm (1/3) [§5.3.1, N&C]

- Idea: Using the *Phase Estimation Algorithm* (PEA) to the order-finding problem.
- Suppose a is coprime to N. Define $U_a: |x\rangle \mapsto |xa \bmod N\rangle, \forall x \in \mathbb{Z}_N$.
  → U_a is *unitary* since a has *inverse modulo* N (hence U_a is reversible).
- Since $a^r \bmod N = 1$, we have $U_a^r: |x\rangle \mapsto |xa^r \bmod N\rangle = |x\rangle$
  (i.e. U_a is the r-th root of the identity operation).
- Let the state $|u_\ell\rangle := \frac{1}{\sqrt{r}} \sum_{x\in\mathbb{Z}_r} e^{-2\pi i \frac{\ell}{r} x} |a^x \bmod N\rangle$.

$$\Rightarrow U_a|u_\ell\rangle = \frac{1}{\sqrt{r}} \sum_{x\in\mathbb{Z}_r} e^{-2\pi i \frac{\ell}{r} x} |a^{x+1} \bmod N\rangle$$
$$= e^{2\pi i \frac{\ell}{r}} \frac{1}{\sqrt{r}} \sum_{x\in\mathbb{Z}_r} e^{-2\pi i \frac{\ell}{r}(x+1)} |a^{x+1} \bmod N\rangle$$
$$= e^{2\pi i \frac{\ell}{r}} |u_\ell\rangle$$

∃! y such that $xy = yx \bmod N = 1$ [§Appendix 4, N&C]

$\because e^{2\pi i \frac{\ell}{r}} |a^r \bmod N\rangle = e^{2\pi i 0} |a^0 \bmod N\rangle$

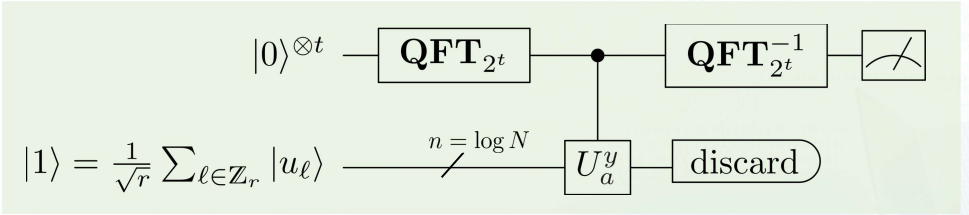
Kitaev's Order-Finding Algorithm (2/3) [§5.3.1, N&C]

- If we were given the eigenstate $|u_\ell\rangle$ of U_a, we are able to apply the PEA and perform $|0\rangle|u_\ell\rangle \mapsto |\widetilde{\ell/r}\rangle|u_\ell\rangle$ (here $|\tilde{\phi}\rangle$ means a good estimate of ϕ as before).
  → Using the error analysis as before would solve the sampling estimate problem.
- However, without knowing r, we do not know how to prepare such state $|u_\ell\rangle$.
- Key: *uniform superposition* $\Rightarrow \frac{1}{\sqrt{r}} \sum_{\ell\in\mathbb{Z}_r} |u_\ell\rangle = \frac{1}{\sqrt{r}} \sum_{\ell\in\mathbb{Z}_r} \frac{1}{\sqrt{r}} \sum_{x\in\mathbb{Z}_r} e^{-2\pi i \frac{\ell}{r} x} |a^x \bmod N\rangle = |1\rangle$
- Hence the PEA maps the input state $|0\rangle|1\rangle = \frac{1}{\sqrt{r}} \sum_{\ell\in\mathbb{Z}_r} |0\rangle|u_\ell\rangle$ to $\frac{1}{\sqrt{r}} \sum_{\ell\in\mathbb{Z}_r} |\widetilde{\ell/r}\rangle|u_\ell\rangle$.
- Measure the first register yielding equally weighted mixture of the states $|\widetilde{\ell/r}\rangle$ for $\ell \in \mathbb{Z}_r$, which is exactly the same as the last step of Shor's sampling estimate method.
- For the controlled-unitary c-$U_a^{2^j}$, $j \in \mathbb{Z}_t$ in PEA, we can use c-$U_a^{2^j}$ = c-$U_{a^{2^j}}$.

[Box 5.2, N&C]

Kitaev's Order-Finding Algorithm (3/3) [§5.3.1, N&C]

- Note that in Shor's order-finding algorithm, the modular exponential function can be implemented by first preparing the state $\frac{1}{\sqrt{2^t}} \sum_{x\in\mathbb{Z}_{2^t}} |x\rangle|1\rangle$ and performing the operation c-U_a^x: $|x\rangle|y\rangle \mapsto |y a^x \bmod N\rangle$ (note here that we take $y = 1$).
- Hence, Shor's order-finding algorithm is actually *equivalent* to Kitaev's order-finding algorithm (using PSE).
  → So are the (i) performance/error analysis; (ii) complexity; and (iii) circuit diagram.



References (1/2)

- P. W. Shor, "Algorithm for Quantum Computation: Discrete Logarithms and Factoring", in *Symposium on Foundations of Computer Science,* 1994.
- P. W. Shor. "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," *SIAM Journal on Scientific Computing*, 26(5):1484–1509, 1997.
- A. Y. Kitaev, "Quantum measurements and the Abelian stabilizer problem." arXiv:quant-ph/9511026.
- S. Beauregard, "Circuit for Shor's algorithm using $2n + 3$ qubits," *Quantum Information and Computation*, 3(2): 175-185, 2003.
- L. M. K. Vandersypen *et al.*, "Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance," *Nature*, 414, 883–887, 2001.
- S. J. Hallgren. *Quantum Fourier Sampling, the Hidden subgroup Problem, and Beyond.* Ph.D. Thesis, University of California, Berkeley, 2000.
- K. Eisenträger, S. Hallgren, A. Kitaev, and F. Song, "A quantum algorithm for computing the unit group of an arbitrary degree number field," *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, May 2014.

References (2/2)

- J. C. P. Miller, "On factorisation, with a suggested new approach," *Mathematics of Computation*, 29(129):155–172, 1975.
- N. Koblitz. *A Course in Number Theory and Cryptography*. 2nd ed. Graduate Texts in Mathematics 114. Springer-Verlag, 1994.
- M. Agrawal, N. Kayal, N, Saxena, "PRIMES is in P," *Annals of Mathematics*, 160:781-793, 2004.
- D. J. Bernstein, "Detecting perfect powers in essentially linear time." *Mathematics of Computation*, 67(223): 1253–1283, 1998.
- A. Montanaro, "Quantum algorithms: an overview," *npj Quantum Information*, 2(15023), 2016.
- A. W. Harrow and A. Montanaro, "Quantum computational supremacy," *Nature*, 549, 203-209, 2017.
- Landmark Algorithm Breaks 30-Year Impasse: https://www.quantamagazine.org/algorithm-solves-graph-isomorphism-in-record-time-20151214/
- Quantum Algorithm Zoo: https://quantumalgorithmzoo.org/
- See "*History and further reading*" in [§5, N&C]; also seee §4–§14 of Andrew Childes' lecture notes.

Quantum Information and Computation
Nonlocal Games

Hao-Chung Cheng (鄭皓中)

haochung@ntu.edu.tw

Department of Electrical Engineering

National Taiwan University

April 14, 2021

Outline

Motivation

Motivation

- Ultimate questions: Under what circumstances does quantum provide advantages?
  - Entanglement is a valuable resource for dense coding and teleportation.
  - Entanglement is necessary for quantum computational speedups.
- More questions:
  - Is there only classical physics/theory?
  - Einstein, Podolsky, and Rosen asked in 1935: "Is quantum theory *complete*?"
  - What is classical? → Local realism & the hidden-variable theory.
  - What is the capability of quantum theory? How strong quantum entanglement can be?
- We will discuss another way of quantifying the *distinction* between classical and quantum theory through the *correlations* predicted by these theories.
- Let's play a game by using classical or quantum strategies!

Nonlocal Games: The CHSH Game/Inequality

Nonlocal Games

- A number of *players* (e.g. Alice and Bob) cooperatively play against a *referee*.
- The referee runs the game and all communication is between the players and the referee – no communication directly between any of the players is permitted.
- For each player, the referee randomly selects a question and sends it to the player.
- Each player sends an *answer* back to the referee, and based on all the questions and answers the referee determines whether the player win or lose (or gives a score).
- The player's goal is to collaborate and maximize their chances of winning or score.
- Before the game, the players meet and may agree upon a joint strategy, but they move far apart from each other and cannot communicate during the game.
- The players can play with any devices (e.g. classical or quantum systems) according the laws of (classical or quantum) physics.

← device-independent games/strategies

The Clauser–Horne–Shimony–Holt Game

x	y	$a \oplus b$
0	0	0
0	1	0
1	0	0
1	1	1

- Players: Alice and Bob; judged by Referee.
- Questions: $x \in \{0,1\}$ to Alice; $y \in \{0,1\}$ to Bob.
- Answers: Alice replies $a \in \{0,1\}$; and Bob replies $y \in \{0,1\}$.
- Winning rule: $x \wedge y = a \oplus b$.

"$\wedge$" means logical AND
"$\vee$" means logical OR

$1_{x\wedge y=a\oplus b}$ can be viewed as a score function $V(x,y,a,b)$

- Winning probability: $\Pr\{\text{Win}\} = \mathbb{E}_{XYAB}\left[1_{X\wedge Y=A\oplus B}\right]$

$$= \sum_{x,y,a,b} 1_{x\wedge y=a\oplus b}\, p_{AB|XY}(a,b|x,y) p_{XY}(x,y)$$

This formula is general regardless of which strategies was chosen.

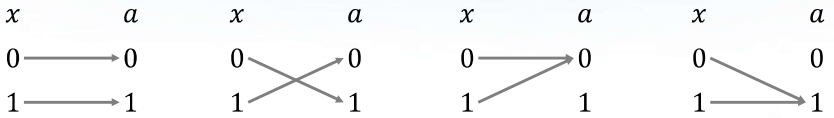
$$= \frac{1}{4}\sum_{x,y,a,b} 1_{x\wedge y=a\oplus b}\, p_{AB|XY}(a,b|x,y)$$

Uniform priors

- Strategies: $p_{AB|XY}(a,b|x,y)$
  - Classical: deterministic or with shared randomness $\rightarrow p_{AB|XY}(a,b|x,y) = \mathbb{E}_{\Lambda}\left[p_{AB|\Lambda XY}\right]$.
  - Quantum: $p_{AB|XY}(a,b|x,y) = \langle \phi_{AB}| \Pi_x^a \otimes \Pi_y^b |\phi_{AB}\rangle$, where $\{\Pi_x^a\}_a$ and $\{\Pi_y^b\}_b$ are POVMs.

Classical Deterministic Strategies (1/2)

- Alice: $a = f(x)$ for some deterministic Boolean function $f: \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$.



- Bob: $b = g(y)$ for some deterministic Boolean function $g: \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$ (just as above).
- There are 16 classical deterministic strategies for $(x,y) \in \mathbb{Z}_2^2$ and $(a,b) \in \mathbb{Z}_2^2$, so one can exhaustively check the winning probability via true table.
- Achievability of classical deterministic strategies
  An example strategy: $f(0) = f(1) = g(0) = g(1) = 0$.
  $\Rightarrow$ It satisfies the first three rows but violates the last rows.
  $\Rightarrow$ The winning probability for this strategy is 75%.

x	y	$a \oplus b$
0	0	0
0	1	0
1	0	0
1	1	1

Classical Deterministic Strategies (2/2)

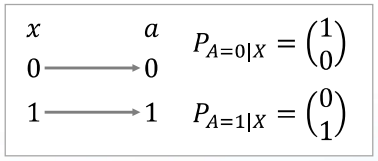
- Optimality: There is no perfect classical deterministic strategies

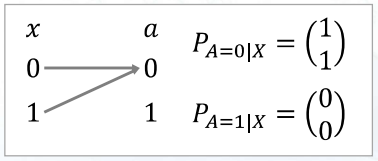
Proof. Suppose there exists Boolean functions $f(\cdot)$ & $g(\cdot)$ satisfying the four rows.

$$\Rightarrow \underbrace{f(0)\oplus g(0)}_{0} \oplus \underbrace{f(0)\oplus g(1)}_{0} \oplus \underbrace{f(1)\oplus g(0)}_{0} \oplus \underbrace{f(1)\oplus g(1)}_{1} = 0 \quad \because c \oplus c = 0$$

$$= 1 \qquad \rightarrow\leftarrow \square$$

- At most three conditions (rows) are satisfied, one is conflicted, so $\Pr\{\text{Win}\} = 75\%$.
- Note that for deterministic Boolean functions $f(\cdot)$ & $g(\cdot)$, one can write the conditional probability as $p_{AB|XY}(a,b|x,y) = \delta_{a,f(x)} \cdot \delta_{b,g(y)} = 1_{a=f(x)} \cdot 1_{b=g(y)}$.

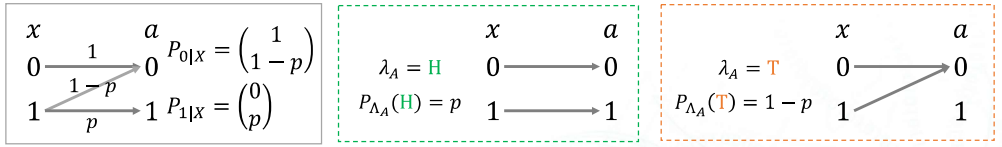




Classical Randomized Strategies

- When receiving x (resp. y), Alice (resp. Bob) *randomly* outputs answer a (resp. b).
- This can be *mathematically* (and *operationally*) described by choosing a deterministic Boolean function according to a private coin outcome, say λ_A & λ_B.

E.g.



- $\Rightarrow p_{A|X}(a|x) = \mathbb{E}_{\Lambda_A}\left[\delta_{a,f(x,\Lambda_A)}\right]$ for some deterministic function $f(\cdot,\Lambda_A)$.
- $\Rightarrow \Pr\{\text{Win}\} = \frac{1}{4}\sum_{x,y,a,b} 1_{x\wedge y=a\oplus b}\, \mathbb{E}_{\Lambda_A}\left[\delta_{a,f(x,\Lambda_A)}\right]\mathbb{E}_{\Lambda_B}\left[\delta_{a,g(x,\Lambda_B)}\right]$

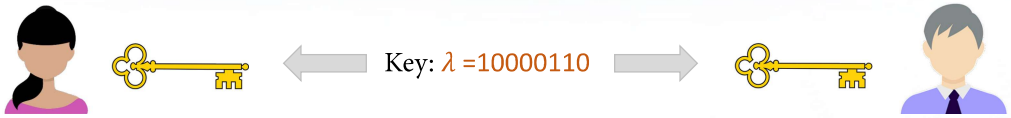
$$\leq \frac{1}{4}\sum_{x,y,a,b} 1_{x\wedge y=a\oplus b} \max_{\lambda_A} \delta_{a,f(x,\lambda_A)} \max_{\lambda_B} \delta_{b,g(x,\lambda_B)}$$

No restrictions on the hidden variable

- $\Rightarrow$ Private randomness does not provide any advantage for the winning probability.

Classical Strategies with Shared Randomness

- Before commencing the game, Alice and Bob meet and flip a coin (many times), and then record the result (say λ=10000110$\cdots$) respectively.



- Given an outcome λ, Alice (resp. Bob) plays with $p_{A|\Lambda X}(a|\lambda,x)$ (resp. $p_{B|\Lambda Y}(b|\lambda,y)$).
  ⇒ Joint strategy given λ is $p_{AB|\Lambda XY}(a,b|\lambda,x,y) = p_{A|\Lambda X}(a|\lambda,x)\cdot p_{B|\Lambda Y}(b|\lambda,y)$.

decoupled given the hidden variable λ

⇒ From outsider's perspective, $p_{AB|XY}(a,b|x,y) = \mathbb{E}_\Lambda\left[p_{AB|\Lambda XY}(a,b|\Lambda,x,y)\right]$.

- ⇒ Shared randomness does not provide any advantage for the winning probability!

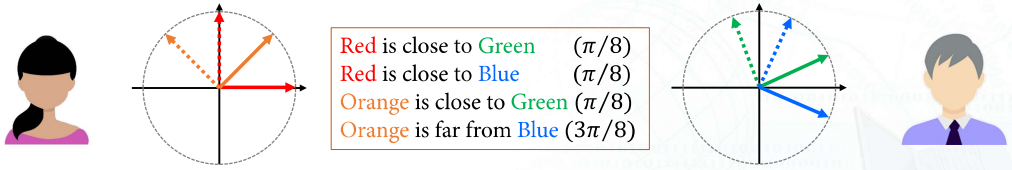
Hidden-Variable Theory & Locality

- **Hidden variables.** The results of any measurement (decision or strategy) on any individual system are *predetermined*. Any probabilities we may use to describe the system merely reflect our ignorance of these hidden definite values, which may vary from one experimental run to another.
  → Albert Einstein: "God does not play dice with the universe."
- **Locality.** Every system has its properties, independently of which interventions are carried out on other spatially separated systems.
  E.g. Alice's choice of measurement/decision does not affect the outcomes of Bob's measurement/decision, and vice versa.
- As a *classical theory*, we want such theory fulfills the above two requirements.
  ⇒ Any *classical correlation*: $p_{AB|XY}(a,b|x,y) = \mathbb{E}_\Lambda\left[p_{A|\Lambda X}(a|\lambda,x)\cdot p_{B|\Lambda Y}(b|\lambda,y)\right]$.
  ⇒ The winning probability for any *classical strategies* is at most 75%.

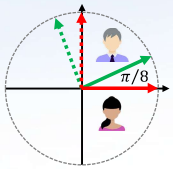
Quantum Strategies with Entanglement (1/3)

x	y	$a\oplus b$
0	0	0
0	1	0
1	0	0
1	1	1

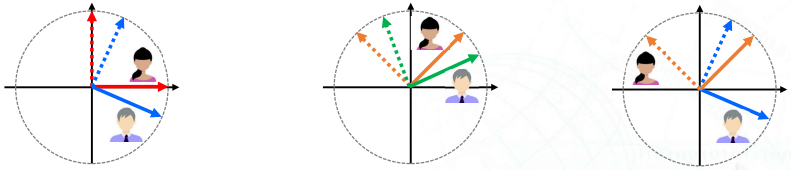
- Before commencing the game, Alice and Bob meet and prepare an EPR pair $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|++\rangle + |--\rangle)$.
  → If Alice gets measurement outcome $|0\rangle/|1\rangle/|+\rangle/|-\rangle$, Bob gets the same.
- Define basis: $\{|\alpha_0(\theta)\rangle = \cos\theta\,|0\rangle + \sin\theta\,|1\rangle,\ |\alpha_1(\theta)\rangle = -\sin\theta\,|0\rangle + \cos\theta\,|1\rangle\}$
- Alice uses basis with $\theta = 0$ when $x = 0$, with with $\theta = \pi/4$ when $x = 1$.
  Bob uses basis with $\theta = \pi/8$ when $y = 0$, with with $\theta = -\pi/8$ when $y = 1$.



Quantum Strategies with Entanglement (2/3)



1. $(x = 0, y = 0)$: Assume Alice measures with $\{|0\rangle, |1\rangle\}$ and gets $|0\rangle$.
   Bob's state collapses to $|0\rangle$ too, and then measure with $\theta = \pi/8$.
   Bob will answer $b = 0$ (so $x \wedge y = a \oplus b$) with prob. $\left|\langle\alpha_0(\frac{\pi}{8})|0\rangle\right|^2 = \cos^2\left(\frac{\pi}{8}\right)$.
2. Similar for $(x = 0, y = 1)$ and $(x = 1, y = 0)$: $\Pr\{a \oplus b = 0\} = \cos^2\left(\frac{\pi}{8}\right) \approx 85\%$.

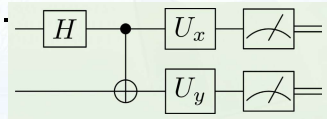


3. $(x = 1, y = 1)$: $\Pr\{a \oplus b = 1\} = \cos^2\left(\frac{\pi}{8}\right) \approx 85\%$.

- ⇒ The winning probability for this quantum strategy is $\frac{1}{2} + \frac{1}{2\sqrt{2}} \approx 85\% \geq 75\%$.

Quantum Strategies with Entanglement (3/3)

- In general for sharing a state $|\phi\rangle_{AB}$ between Alice and Bob, we have
$$p_{AB|XY}(a,b|x,y) = \langle\phi|\Pi_x^a \otimes \Pi_y^b|\phi\rangle.$$
Here, $\{\Pi_x^a\}_a$ and $\{\Pi_y^b\}_b$ are projective measurements or general POVMs.
- Hence, the *quantum strategy* for Alice and Bob is to choose $|\phi\rangle_{AB}$, $\{\Pi_x^a\}_a$ and $\{\Pi_y^b\}_b$.
- If the sharing pure state is not entangled, i.e. $|\phi\rangle_{AB} = |\alpha\rangle_A|\beta\rangle_B$,
$\Rightarrow p_{AB|XY}(a,b|x,y) = \langle\alpha|\Pi_x^a|\alpha\rangle \cdot \langle\beta|\Pi_y^b|\beta\rangle$ → Reduces to classical strategies.
- Optimality: Tsirelson's bound shows that $\Pr\{\text{Win}\} \le \cos^2\left(\frac{\pi}{8}\right)$ for quantum strategies.
- This motivates us to prepare quantum states with randomness (i.e. mixed state):
$$p_{AB|XY}(a,b|x,y) = \mathrm{Tr}\left[\rho_{AB}\Pi_x^a \otimes \Pi_y^b\right].$$
$\Rightarrow$ Not entangled (separable) state $\rho_{AB} = \mathbb{E}_\Lambda\left[\rho_A^\Lambda \otimes \rho_B^\Lambda\right]$.



The CHSH Inequality (1/2)

x	y	$a \oplus b$
0	0	0
0	1	0
1	0	0
1	1	1

$a \mapsto (-1)^a$
$b \mapsto (-1)^b$

- For any (x, y), let us consider a correlation-like quantity:
$$\langle A_x B_y\rangle := \mathbb{E}_{P_{AB|X=x,Y=y}}[(-1)^A(-1)^B].$$
- The score function is $S := \langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle$.
- Claim: $\Pr\{\text{Win}\} - \Pr\{\text{Lose}\} = S/4$. $\Rightarrow S = 8\Pr\{\text{Win}\} - 4$.
- The bound $S_{\text{classical}} \le 2$ under all classical strategies is called the *Bell-type inequality*.
- However, we've known $S^\star_{\text{quantum}} = 2\sqrt{2}$ violates this Bell-type (CHSH) inequality.
- *Proof.* For $(x,y) = (0,0)$ or $(0,1)$ or $(1,1)$,
$$P_{\text{w}} - P_{\text{l}} = \sum_{a,b} 1_{a\oplus b=0} p(a,b|x,y) - \sum_{a,b} 1_{a\oplus b=1} p(a,b|x,y)$$
$$= \sum_{a,b}(-1)^{a\oplus b} p(a,b|x,y) = \sum_{a,b}(-1)^a(-1)^b p(a,b|x,y) = \langle A_x B_y\rangle.$$
For $(x,y) = (1,1)$, $P_{\text{w}} - P_{\text{l}} = \sum_{a,b} 1_{a\oplus b=1} p(a,b|x,y) - 1_{a\oplus b=0} p(a,b|x,y)$
$$= \sum_{a,b}(-1)^{\overline{a\oplus b}} p(a,b|x,y) = -\sum_{a,b}(-1)^{a\oplus b} p(a,b|x,y) = -\langle A_x B_y\rangle.$$
□

The CHSH Inequality (2/2)

- For quantum strategies, it is written via the notation of *observables*.
$$\langle A_x B_y\rangle = \sum_{a,b}(-1)^a(-1)^b \langle\phi|\Pi_x^a \otimes \Pi_y^b|\phi\rangle$$
$$= \langle\phi|\textstyle\sum_a(-1)^a\Pi_x^a \otimes \sum_b(-1)^b\Pi_y^b|\phi\rangle$$
$$=: \langle\phi|A_x \otimes B_y|\phi\rangle$$
Linearity of inner product and tensor
- The *observable* corresponding to the $\{\Pi_x^a\}_a$ is Hermitian matrix $A_x := \sum_a(-1)^a\Pi_x^a$.
- For example, we choose the following strategy:
$$A_0 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = |0\rangle\langle 0| - |1\rangle\langle 1| = Z, \qquad A_1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = |+\rangle\langle +| - |-\rangle\langle -| = X$$
$$B_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} = H, \qquad B_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ -1 & -1\end{pmatrix}$$
$\Rightarrow \langle\Phi|A_0 \otimes B_0|\Phi\rangle = \langle\Phi|A_0 \otimes B_1|\Phi\rangle = \langle\Phi|A_1 \otimes B_0|\Phi\rangle = -\langle\Phi|A_1 \otimes B_1|\Phi\rangle = \frac{1}{\sqrt{2}}$.

No-Signaling Distribution

- From now on, we are not limited by the framework of quantum theory, but we just restrict the conditional probability $p_{AB|XY}$ so that Alice & Bob can't communicate.
- Description of one's system alone is independent of others' measurements/actions.

$$p_{A|X}(a|x) = p_{A|XY}(a|x,y) = \textstyle\sum_b p_{AB|XY}(a,b|x,y), \forall y$$
$$p_{B|Y}(b|y) = p_{B|XY}(b|x,y) = \textstyle\sum_a p_{AB|XY}(a,b|x,y), \forall x$$

- We have seen that quantum theory does satisfy the no-signaling principle.
- Is quantum correlation the strongest one satisfying the no-signaling principle?
- Is nonlocality a unique feature of quantum theory?

Popescu–Rohrlich (PR) Box

x	y	$a \oplus b$
0	0	0
0	1	0
1	0	0
1	1	1

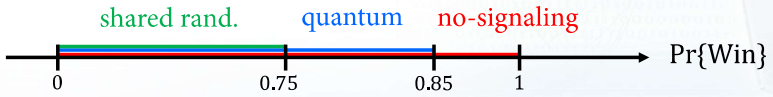
- Consider:

$$p_{AB|XY}(0,0|0,0) = p_{AB|XY}(1,1|0,0) = 1/2$$
$$p_{AB|XY}(0,0|0,1) = p_{AB|XY}(1,1|0,1) = 1/2$$
$$p_{AB|XY}(0,0|1,0) = p_{AB|XY}(1,1|1,0) = 1/2$$
$$p_{AB|XY}(0,1|1,1) = p_{AB|XY}(1,0|1,1) = 1/2$$

- They satisfy the no-signaling requirement:

$$p_{A|X}(a|x) = \sum_b p_{AB|XY}(a,b|x,0) = \sum_b p_{AB|XY}(a,b|x,1) = 1/2$$
$$p_{B|Y}(b|y) = \sum_a p_{AB|XY}(a,b|0,y) = \sum_a p_{AB|XY}(a,b|1,y) = 1/2$$

- Checking the wining rule, this strategy gives Pr{Win} = 100%!



Nonlocal Games: The GHZ Game

The Greenberger–Horne–Zeilinger Game

x	y	z	$a \oplus b \oplus c$
0	0	0	0
1	1	0	1
1	0	1	1
0	1	1	1

- Players: Alice, Bob, and Charlie; judged by Referee.
- Questions: $x, y, z \in \{0,1\}$ to players, respectively with answers.
- Answers: $a, b, c \in \{0,1\}$ to Referee.
- Winning rule: $x \vee y \vee z = a \oplus b \oplus c$.

"∧" means logical AND
"∨" means logical OR

- Winning probability: $\Pr\{\text{Win}\} = \frac{1}{4}\sum_{x,y,z,a,b,c} 1_{x\vee y\vee z=a\oplus b\oplus c}\, p_{ABC|XYZ}(a,b,c|x,y,z)$.
- Classical achievability: all choose functions with output all 1's. ⇒ Pr{Win} ≤ 75%.
- Classical optimality: No perfect classical strategies:

$$1 = 0 \oplus 1 \oplus 1 \oplus 1$$
$$= \big(f(0) \oplus g(0) \oplus h(0)\big) \oplus \big(f(1) \oplus g(1) \oplus h(0)\big)$$
$$\oplus \big(f(1) \oplus g(0) \oplus h(1)\big) \oplus \big(f(0) \oplus g(1) \oplus h(1)\big) = 0 \quad \rightarrow\leftarrow$$

□

The GHZ Game (Quantum Strategies)

- Using the idea of observables and noting that $(-1)^{a+b+c} = (-1)^{a\oplus b\oplus c}$,

  $\Rightarrow \Pr\{\text{Win}\} = \frac{1}{2} + \frac{1}{8}S$, where $S = -\langle A_0B_0C_0\rangle - \langle A_1B_1C_0\rangle - \langle A_1B_0C_1\rangle + \langle A_0\, B_1C_1\rangle$

  $\Rightarrow S = \langle\phi| - A_0 \otimes B_0 \otimes C_0 - A_1 \otimes B_1 \otimes C_0 - A_1 \otimes B_0 \otimes C_1 + A_0 \otimes B_1 \otimes C_1|\phi\rangle$

- Choose the 3-qubit sharing state $|\phi\rangle = \frac{1}{2}(|000\rangle - |110\rangle - |101\rangle - |011\rangle)$.
- Alice/Bob/Charlies measures Z (resp. X) when receiving 0 (resp. 1).

  For $(x,y,z) = (0,0,0)$, $\Rightarrow A_x \otimes B_y \otimes C_z|\phi\rangle = Z \otimes Z \otimes Z|\phi\rangle|\phi\rangle$.

  For $(x,y,z) = (1,1,0)$ or $(1,0,0)$ or $(0,1,1)$, $\Rightarrow A_x \otimes B_y \otimes C_z|\phi\rangle = -|\phi\rangle$.

- $\Rightarrow S = 4$, and then $\Rightarrow \Pr\{\text{Win}\} = 100\%$!
- We remark that no 2-qubit state can achieve perfect winning strategy.

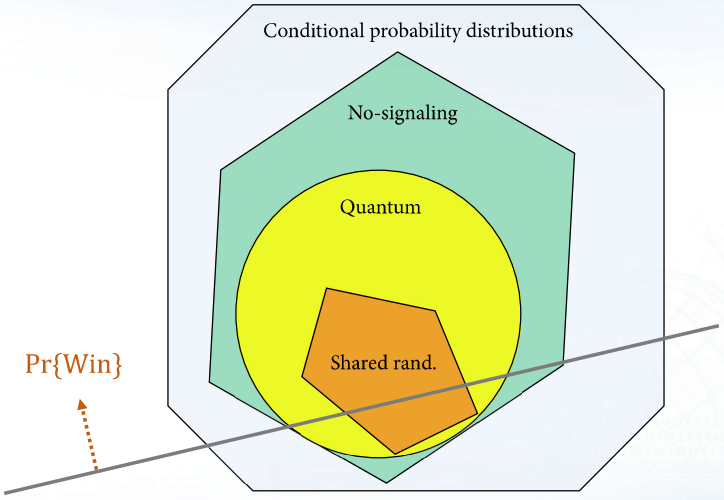
Concluding Remarks

Concluding Remarks

- Nonlocal games provide a *provable separation* between classical correlations, quantum correlations, and beyond (e.g. no-signaling correlations).
- More and more teams been demonstrating experimental separations.
- Entanglement is necessary for violating Bell's inequalities.
  → Non-entangled (product) state only leads to independence, i.e. $P_{A|X} \cdot P_{B|Y}$.
- An entangled state does not necessarily yield nonlocality (e.g. Werner states).
- Some states do not violate a Bell inequality but exhibit hidden nonlocality.
- Nonlocality is not equivalent to entanglement – quantum locality without entanglement – in other scenarios of state discrimination.
- Applications: device-independent quantum cryptography, randomness expansion.

See [References]

Correlation Hierarchy



Convex set
No-signaling: polytope
Quantum: semi-definite
Shared rand.: polytope

Game is hyperplane

Appendix

Optimality: Tsirelson's Bound (1/2)

x	y	$a \oplus b$
0	0	0
0	1	0
1	0	0
1	1	1

- Consider a general quantum strategy:

  Alice uses basis with angle θ_0 when $x = 0$, with with angle θ_1 when $x = 1$.
  Bob uses basis with angle τ_0 when $y = 0$, with with angle τ_1 when $y = 1$.

- From a geometric view, we have

$$\Pr\{\text{Win}\} = \frac{1}{4}[\cos^2(\theta_0 - \tau_0) + \cos^2(\theta_0 - \tau_1) + \cos^2(\theta_1 - \tau_0) + \sin^2(\theta_1 - \tau_1)]$$

Optimality: Tsirelson's Bound (2/2)

- By trigonometry formulas:

$$\Pr\{\text{Win}\} = \frac{1}{2} + \frac{1}{8}[\cos 2(\theta_0 - \tau_0) + \cos 2(\theta_0 - \tau_1) + \cos 2\,(\theta_1 - \tau_0) - \cos 2(\theta_1 - \tau_1)]$$

- Rewrite it as inner products of unit vectors:

$$\begin{aligned}\Pr\{\text{Win}\} &= \frac{1}{2} + \frac{1}{8}[u_0 \cdot v_0 + u_0 \cdot v_1 + u_1 \cdot v_0 - u_1 \cdot v_1] \\ &= \frac{1}{2} + \frac{1}{8}[u_0 \cdot (v_0 + v_1) + u_1 \cdot (v_0 - v_1)] \\ &\le \frac{1}{2} + \frac{1}{8}[\,\|v_0 + v_1\| + \|v_0 - v_1\|] \\ &\le \frac{1}{2} + \frac{1}{8}\sqrt{2\|v_0 + v_1\|^2 + 2\|v_0 - v_1\|^2} \le \frac{1}{2} + \frac{1}{8}\sqrt{8}\end{aligned}$$

Cauchy–Swartz inequality

Parallelogram inequality

□

References (1/4)

- Original Bell-type inequalities and nonlocal games
  - A. Einstein, B. Podolsky, N. Rosen, "Can quantum-mechanical description of reality be considered complete?" *Physical Review*, 47, 777–780, 1935.
  - J. S. Bell, "One the Einstein–Podolsky–Rosen paradox," *Physics*, 1, 195–200, 1964.
  - J. F. Clauser, M. A. Horne, A. Shimony, R. A. olt, "Proposed experiment to test local hidden-variable theories," *Physical Review Letters*, 23, 880–884, 1969.
  - D. M. Greenberger, M. A. Horne, A. Zeilinger, "Going Beyond Bell's Theorem," in *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, Springer, Dordrecht, 69–72, 1989; "Bell's theorem without inequalities," *American Journal of Physics* 58(12):1131–1143, 1990.
- Optimality (see also https://www.tau.ac.il/~tsirel/Research/bellopalg/main.html)
  - B. S. Tsirelson, "Quantum generalizations of Bell's inequality," *Letters in Mathematical Physics*, 4 (2): 93–100, 1980; "Quantum analogues of the Bell inequalities. The case of two spatially separated domains," *Journal of Soviet Mathematics*, 36 (4): 557–570, 1987.
  - M. Junge, M. Navascués, C. Palazuelos. D. Pérez-García, V. B. Scholz, R. F. Werner, "Connes' embedding problem and Tsirelson's problem," *Journal of Mathematical Physics,* 52(1): 012102, 2011.

References (2/4)

- Review papers/books
  - R. F. Werner, M. M. Wolf, "Bell Inequalities and Entanglement," *Quantum Information and Computation*, 1, 1–25, 2001.
  - N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, S. Wehner, "Bell nonlocality" *Reviews of Modern Physics*, 86, 419, 2014.
  - R. A. Bertelmann, A. Zeilinger (Ed.), *Quantum [un]speakables – From Bell to quantum information.* Springer-Verlag, Berlin, 2002.
  - I. Šupić, and J. Bowles, "Self-testing of quantum systems: a review," *Quantum*, 4(337), 2020.
- Applications in quantum cryptography & randomness expansion & nonlocal computation
  - A. Ekert, "Quantum cryptography based on Bell's theorem," *Physical Review Letters*, 67, 661–663, 1991.
  - R. Colbeck, "Quantum and Relativistic Protocols for Secure Multi-Party Computation," arXiv:0911.3814.
  - A. Acín, L. Masanes, "Certified randomness in quantum physics," *Nature,* 540(7632):213–219, 2016.
  - A. Acín, N. Gisin, L. Masanes, "From Bell's Theorem to Secure Quantum Key Distribution," *Physical Review Letters*, 97, 120405, 2006.

References (3/4)

- Hierarchy of correlations & PR-boxes & nonlocality
  - S. Popescu and D. Rohrlich, "Quantum nonlocality as an axiom," *Foundations of Physics*, 24, 379–385, 1994.
  - J. Barrett, S. Pironio, "Popescu–Rohrlich correlations as a unit of nonlocality," *Physical Review Letters*, 95, 140401, 2005.
  - M. Navascues, S. Pironio, A. Acín, "Bounding the set of quantum correlations," *Physical Review Letters*, 98, 010401, 2007.
  - S. Popescu, "Nonlocality beyond quantum mechanics," *Nature Physics*, 10, 264–270, 2014.
  - M. Pawlowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, M. Zukowski, "Information Causality as a Physical Principle," *Nature*, 461(1101), 2009.
  - N. Linden, S. Popescu, A. J. Short, A, Winter, "Quantum Nonlocality and Beyond: Limits from Nonlocal Computation," *Physical Review Letters*, 99(180502), 2007.
  - C. Palazuelos, "Super-activation of quantum non-locality," *Physical Review Letters*, 109, 190401, 2012.
- Other extensions (games)
  - N. Brunner, N. Linden, "Connection between Bell nonlocality and Bayesian game theory," *Nature Communications*, 4(2057), 2013.
  - C. A. Melo-Luna *et al.*, "Quantum Locality in Game Strategy," *Scienific Reports*, 7(44730), 2017.

References (4/4)

- Experimental aspects
  - J. F. Clauser, A. Shimony, "Bell's theorem. Experimental tests and implications," *Reports on Progress in Physics*, 41(12):1881, 1978.
  - A. Aspect, P. Grangier, G. Roger, "Experimental realization of Einstein–Podolsky–Rosen–Bohm Gedankenexperiment: a new violation of Bell's inequalities. *Physical review letters*, *49*(2):91, 1982.
  - G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, "Violation of Bell's inequality under strict Einstein locality conditions," *Physical Review Letters*, 81(23):5039, 1998.
  - A. Aspect, "Bell's inequality test: more ideal than ever," *Nature*, 398, 189–190, 1999.
  - M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, D. J. Wineland, "Experimental violation of a Bell's inequality with efficient detection," *Nature*, 409, 791–794, 2001.
- Interpretations (not really recommended since you might get drawn there)
  - G. Auletta, *Foundations and Interpretation of Quantum Mechanics: In the Light of a Critical-Historical Analysis of the Problems and of a Synthesis of the Results*. World Scientific, 2000.

Quantum Information and Computation
Open Quantum Systems

Hao-Chung Cheng (鄭皓中)

haochung@ntu.edu.tw

Department of Electrical Engineering

National Taiwan University

April 28, 2021

Motivation

Motivation

Preparation → Evolution → Measurement → Data

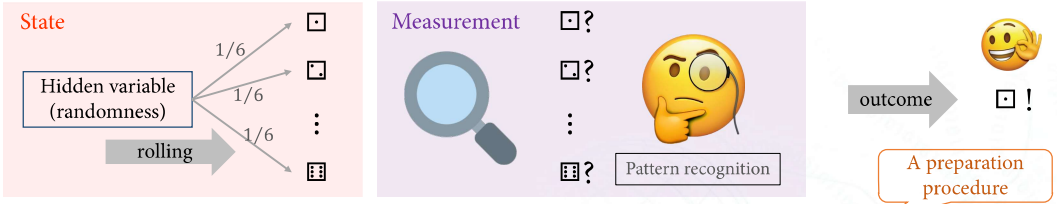
- What we have learnt so far (for closed quantum systems):
  - Pure states: unit vectors (up to a global phase) in a Hilbert space.
  - Reversible evolutions: unitary operations on the Hilbert space.
  - Projection measurements: often measuring with respect to an orthonormal basis.
- What if the underlying closed quantum system is interacting with an environment or there is some hidden noise/randomness that we are not in control of?
- Does quantum theory fully capture a statistical framework?
- Today: We will incorporate *probability reasoning* into the previous formalism.
- We will find a correspondence between the closed & open quantum systems.
- If there is such a correspondence, what's the point of studying open quantum stuff ?

  → More resources to *purify* everything! Essentially, closed quantum ↔ God's view.

Postulate 1 (State)

Probabilistic Classical Systems and States

- To a *probabilistic/randomized* classical system where each symbol in the alphabet occurs with certain probability, we use a *probability model* to characterize it.
  - $\mathcal{X} = \{\boxdot, \boxdot, \boxdot, \boxdot, \boxdot, \boxdot\}$ for a *unbiased* dice: $\Pr(x) = 1/6, \forall x \in \mathcal{X}$.



- The *state* of a dice is associated with a *probability distribution*, say P_X.
  - A *biased* dice: $P_X = [1/2, 1/3, 1/6, 0, 0, 0]^\dagger$; a *deterministic* dice: $P_X = [1, 0, 0, 0, 0, 0]^\dagger$.
- An *observation* on $\mathcal{X}$ gives an *realization* of the random variable X.

  → This is called a *measurement outcome*.

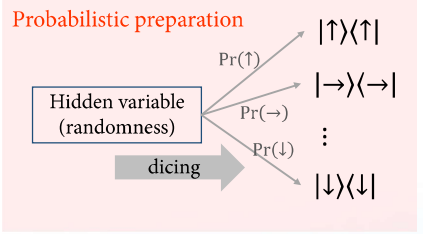
Once a measurement is performed, the state collapses!

Open Quantum Systems and States (1/2)

- Previously, a *pure* quantum state (of a closed quantum system) is associated with a unit vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$.

A deterministic preparation

- A unit vector $|\psi\rangle$ (up to a global phase, i.e. $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$) is by a one-to-one correspondence to a *rank-one projection* $|\psi\rangle\langle\psi|$ to the subspace spanned by $|\psi\rangle$.
- One can roll a dice to prepare each rank-one projection with certain probability. This is a (probabilistic) *preparation procedure* for a quantum system.



Mathematically, we use convex combination of rank-one projection to describe such a probabilistic preparation.

Mixed quantum state:

$\Pr(\uparrow)\,|\uparrow\rangle\langle\uparrow| + \Pr(\rightarrow)\,|\rightarrow\rangle\langle\rightarrow| + \cdots + \Pr(\downarrow)\,|\downarrow\rangle\langle\downarrow|$

Open Quantum Systems and States (2/2)

Postulate 1⋆ (State)

- To any physical (quantum-mechanical) system, there is associated a Hilbert space, i.e. a complex inner product space.
- The system is completely described by its *density operator*, which is a positive semi-definite operator with unit trace acting on the Hilbert space.

- Density operator/matrix or *mixed state*: $\rho \geq 0, \mathrm{Tr}[\rho] = 1$.
- A classical probability distribution is $p_i \geq 0, \sum_i p_i = 1$.

Probability vector → Matrix
Summation→ Trace

Remarks on Density Operators

Pure state

Mixed state

- Any convex combination of rank-one projections gives a density operator.

  Conversely, any density operator ρ can be written in the form:

off-diagonal phase terms

eigenvectors

$$\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i| = U \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_d \end{pmatrix} U^\dagger = \begin{pmatrix} \rho_{11} & \cdots & \rho_{1d} \\ \vdots & \ddots & \vdots \\ \rho_{d1} & \cdots & \rho_{dd} \end{pmatrix}$$

eigenvalues

Will give its *entropy*

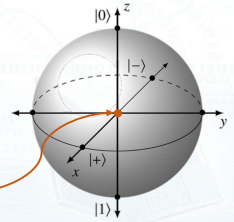
  The eigenvalues $\{\lambda_i\}_i$ form a probability distribution, i.e. $\lambda_i \geq 0, \sum_i \lambda_i = 1$.

- Such decomposition is *not unique* due to the multiplicity of eigenvalues (but the decomposition into eigenprojections/eigenspaces is unique).
  Note that the decomposition (into a *simplex*) of a probability distribution is unique.
- A classical distribution P_X can be *embedded* into a (diagonal) density operator ρ_X:
  Fix a computation basis $\{|x\rangle\}_{x\in\mathcal{X}}$; write $\rho_X = \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x|$.

Classical state

Properties of Quantum States

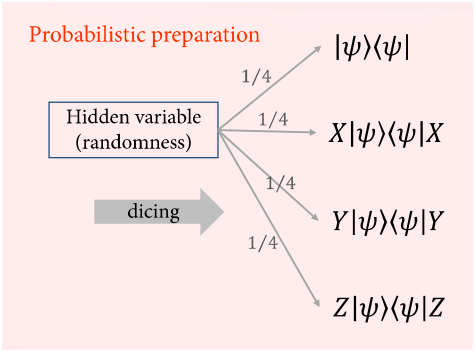
- The set of quantum states $\mathcal{S}(\mathcal{H})$ is *convex* (just as the set of probability distributions).
  - Given $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ and $t \in [0,1]$, then $t\rho + (1-t)\sigma \in \mathcal{S}(\mathcal{H})$ is a gain a state.
  - In general, for any ensemble $\{p_i, \rho_i\}_i$, the mixture $\rho = \sum_i p_i \rho_i$ is a gain a state.
  - Conversely, a state ρ admits *infinitely many* decompositions $\rho = \sum_i p_i \rho_i$.
- Extremal elements of $\mathcal{S}(\mathcal{H})$ are the pure states;
  state that has eigenvalue 0 belongs to the boundary of $\mathcal{S}(\mathcal{H})$.
- Let $\mathcal{H} = \mathbb{C}^d$. There are at most d mutually orthogonal pure states, whereas the dimension of $\mathcal{S}(\mathcal{H})$ is $d^2 - 1$.
  → Bloch-vector representation: ∃ traceless operators $E_1, \ldots E_{d^2-1}$ such that $\langle E_i, E_j \rangle = d\delta_{ij}$ and $\rho = \frac{1}{d}\left(I + \sum_{i=1}^{d^2-1} r_i E_i\right)$.
  [Bengtsson–Życzkowski, 2006]
- E.g. Maximally mixed state I/d.
  → Eigenvalues form a uniform distribution $[1/d, \ldots, 1/d]^\dagger$.



An Example of Maximally Mixed State

$|\psi\rangle = a|0\rangle + b|1\rangle$

- Suppose a quantum device outputs $I|\psi\rangle$, $X|\psi\rangle$, $Y|\psi\rangle$, and $Z|\psi\rangle$ equally probable:



$$= \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} a^* & b^* \end{pmatrix} = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}$$

$$= \begin{pmatrix} b \\ a \end{pmatrix} \begin{pmatrix} b^* & a^* \end{pmatrix} = \begin{pmatrix} |a|^2 & a^*b \\ ab^* & |b|^2 \end{pmatrix}$$

$$= \begin{pmatrix} b \\ -a \end{pmatrix} \begin{pmatrix} b^* & -a^* \end{pmatrix} = \begin{pmatrix} |b|^2 & -a^*b \\ -ab^* & |a|^2 \end{pmatrix}$$

$$= \begin{pmatrix} a \\ -b \end{pmatrix} \begin{pmatrix} a^* & -b^* \end{pmatrix} = \begin{pmatrix} |a|^2 & -ab^* \\ -a^*b & |b|^2 \end{pmatrix}$$

$$\Rightarrow \rho = \frac{1}{4}|\psi\rangle\langle\psi| + \frac{1}{4}X|\psi\rangle\langle\psi|X + \frac{1}{4}Y|\psi\rangle\langle\psi|Y + \frac{1}{4}Z|\psi\rangle\langle\psi|Z = \frac{1}{2}I$$

Maximally mixed qubit state

Postulate 2 (Composite System)

Composite Classical Systems (1/2)

- *State* of a composition of classical system is given by a *joint probability distribution*.
  - $\mathcal{X} \times \mathcal{Y} = \{\boxdot, \boxdot, \boxdot, \boxdot, \boxdot, \boxdot\} \times \{\spadesuit, \clubsuit, \heartsuit, \diamondsuit\}$, and
  - For a P_{XY}, we have $P_{XY}(x,y) \geq 0$ and $\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} P_{XY}(x,y) = 1$.
  - Marginal distribution: $P_X(x) \coloneqq \sum_{y\in\mathcal{Y}} P_{XY}(x,y)$.
- Fix bases $\{|x\rangle\}_{x\in\mathcal{X}}$, $\{|y\rangle\}_{y\in\mathcal{Y}}$, and $\{|xy\rangle\}_{(x,y)\in\mathcal{X}\times\mathcal{Y}}$.
  Let's write P_{XY} in terms of a density operator ρ_{XY} on Hilbert space spanned by $\{|xy\rangle\}$.

$|xy\rangle \coloneqq |x\rangle \otimes |y\rangle$

$$\rho_{XY} = \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} P_{XY}(x,y)|xy\rangle\langle xy| = \begin{pmatrix} P_{XY}(0,0) & & & \\ & P_{XY}(0,1) & & \\ & & \ddots & \\ & & & P_{XY}(|\mathcal{X}|-1,|\mathcal{Y}|-1) \end{pmatrix}$$

$$\rho_X = \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x| = \begin{pmatrix} P_X(0) & & & \\ & P_X(1) & & \\ & & \ddots & \\ & & & P_X(|\mathcal{X}|-1) \end{pmatrix}$$

Composite Classical Systems (1/2)

i: index; X/Y: systems

- *Classical correlation*: Any classical joint state can be written as $\sum_i p_i \rho_X^i \otimes \rho_Y^i$.

$$\rho_{XY} = \sum_{(x,y)} P_{XY}(x,y)|xy\rangle\langle xy| = \sum_{(x,y)} P_{XY}(x,y)|x\rangle \otimes |y\rangle\langle x| \otimes \langle y|$$
$$= \sum_{(x,y)} P_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y|$$

$|x\rangle\langle x|$ projects onto span$\{|x\rangle\}$

- *Independence*: $P_{XY}(x,y) = P_X(x) \cdot P_Y(y)$ for all $(x,y) \in \mathcal{X} \times \mathcal{Y}$.

Independence

$$\rho_{XY} = \sum_{(x,y)} P_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| = \sum_{(x,y)} P_X(x) \cdot P_Y(y)|x\rangle\langle x| \otimes |y\rangle\langle y|$$

Linearity

$$= \sum_{(x,y)} P_X(x)|x\rangle\langle x| \otimes P_Y(y)|y\rangle\langle y| = \sum_x P_X(x)|x\rangle\langle x| \otimes \sum_y P_Y(y)|y\rangle\langle y|$$
$$= \rho_X \otimes \rho_Y$$

⇒ The product state $\rho_X \otimes \rho_Y$ means that system X is *independent* of system Y.

Examples

- Perfect (*maximally correlated*) classical state on $\mathcal{X} \times \mathcal{Y} = \mathbb{Z}_2 \times \mathbb{Z}_2$:
  $P_{XY}(0,0) = P_{XY}(1,1) = 1$.

$$\Rightarrow \rho_{XY} = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11| = \begin{pmatrix} 1/2 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1/2 \end{pmatrix}$$

- Uniform (*maximally mixed*) classical state on $\mathcal{X} \times \mathcal{Y} = \mathbb{Z}_2 \times \mathbb{Z}_2$:
  $P_{XY}(0,0) = P_{XY}(0,1) = P_{XY}(1,0) = P_{XY}(1,1) = 1/4$. ⇒ $P_X(x) = P_Y(y) = 1/2$.

Independence

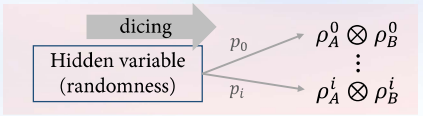
$$\rho_{XY} = \begin{pmatrix} 1/4 & & & \\ & 1/4 & & \\ & & 1/4 & \\ & & & 1/4 \end{pmatrix} = \begin{pmatrix} 1/2 & \\ & 1/2 \end{pmatrix} \otimes \begin{pmatrix} 1/2 & \\ & 1/2 \end{pmatrix} = \frac{I}{2} \otimes \frac{I}{2}$$

Composite Quantum Systems

Postulate 2* (Composition)

- For a joint system composed of two subsystems with Hilbert space $\mathcal{H}_A$ and $\mathcal{H}_B$, the Hilbert space is the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ (e.g. $\mathbb{C}^d \otimes \mathbb{C}^d = \mathbb{C}^{d\times d}$).
- If the system i is *independently* prepared in the state ρ_i, the joint state of the total system is $\bigotimes_i \rho_i$, called the *product state*.
- In general, a quantum state on a joint system $\mathcal{H}_A \otimes \mathcal{H}_B$ is given by an associated density operator ρ_{AB}.
  The state describing subsystem $\mathcal{H}_A$ is the *reduced state* $\rho_A = \mathrm{Tr}_B[\rho_{AB}]$.

Entangled States



- Definition (Separability). A state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$ is separable if $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$ for some prob. distribution $\{p_i\}_i$ and states $\rho_{A/B}^i \in \mathcal{S}(\mathcal{H}_{A/B})$.
- Remark. Every classical joint distribution is separable, but a separable state is not necessary classical (diagonal) $\because \rho_A^i$ and ρ_B^i might not *commute*. i.e. $AB \neq BA$
- Definition (Entanglement). A state ρ_{AB} on $\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is entangled if $\rho_{AB} \neq \sum_i p_i \rho_A^i \otimes \rho_B^i$ for any prob. distribution $\{p_i\}_i$ and states $\rho_{A/B}^i \in \mathcal{S}(\mathcal{H}_{A/B})$.
- Property: A pure state (rank-one projection) is *separable* if and only if it is *product*.

Maximally entangled state:

$$|\Phi_{AB}^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix} \Rightarrow \rho_{AB} = |\Phi_{AB}^+\rangle\langle\Phi_{AB}^+| = \begin{pmatrix}1/2 & & & 1/2\\ & 0 & & \\ & & 0 & \\ 1/2 & & & 1/2\end{pmatrix}$$

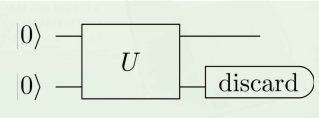
Partial Trace and Reduced States

- Suppose we are given a quantum state ρ_{AB} on a quantum system AB composed of two subsystems A and B with overall Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.
  Which state should we use to describe the state of subsystem A along?
- Classically, recall that the *marginal* distribution is given by $P_X(x) \coloneqq \sum_{y\in\mathcal{Y}} P_{XY}(x,y)$.
- Definition (Partial trace). For every operator M_{AB} on $\mathcal{H}_{AB}$,

$$\mathrm{Tr}_B[M_{AB}] \coloneqq \sum_b (I_A \otimes \langle b|) M_{AB} (I_A \otimes |b\rangle)$$

  Linear operation

  where $\{|b\rangle\}_b$ is any *orthonormal* basis of $\mathcal{H}_B$.
- Definition (Reduced states). Given a state ρ_{AB} on a quantum system AB, we define its *reduced state* on subsystem A by $\rho_A \coloneqq \mathrm{Tr}_B[\rho_{AB}]$.
- Operationally, it is equivalent to ignoring the other systems.



Examples (1/2)

$\rho_{XY} = |x_0y_0\rangle\langle x_0y_0| = |x_0\rangle\langle x_0| \otimes |y_0\rangle\langle y_0|$

- Reduced state of a classical deterministic state $P_{XY}(x_0,y_0) = 1$ for a (x_0,y_0):

$$\begin{aligned}\Rightarrow \mathrm{Tr}_Y[\rho_{XY}] &= \textstyle\sum_y (I_X \otimes \langle y|)|x_0\rangle\langle x_0| \otimes |y_0\rangle\langle y_0|(I_X \otimes |y\rangle)\\ &= \textstyle\sum_y (|x_0\rangle\langle x_0| \otimes \langle y|y_0\rangle\langle y_0|)(I_X \otimes |y\rangle) = \sum_y |x_0\rangle\langle x_0| \otimes \langle y|y_0\rangle\langle y_0|y\rangle\\ &= |x_0\rangle\langle x_0| \otimes \langle y_0|y_0\rangle\langle y_0|y_0\rangle = |x_0\rangle\langle x_0|\end{aligned}$$

The marginal of a joint deterministic (pure) classical state is still deterministic (pure).

- Reduced state of a product state $\rho_{AB} = \rho_A \otimes \rho_B$: Classically, $\sum_y P_X(x)P_Y(y) = P_X(x)$

$$\begin{aligned}\Rightarrow \mathrm{Tr}_B[\rho_{AB}] &= \textstyle\sum_b (I_A \otimes \langle b|)\rho_A \otimes \rho_B (I_A \otimes |b\rangle) = \sum_b (\rho_A \otimes \langle b|\rho_B)(I_A \otimes |b\rangle)\\ &= \textstyle\sum_b \rho_A \otimes \langle b|\rho_B|b\rangle = \rho_A \otimes \mathrm{Tr}[\rho_B] = \rho_A.\end{aligned}$$

- Question: Are the reduced states of a pure joint state pure as well?

Examples (2/2)

- Reduced state of a maximally entangled state $\rho_{AB} = |\Phi^+_{AB}\rangle\langle\Phi^+_{AB}|$:

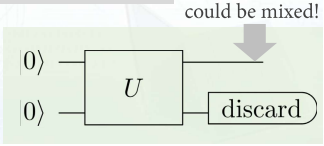
$$= \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|)$$

$$= \frac{1}{2}(|0\rangle\langle 0| \otimes \underbrace{|0\rangle\langle 0|}_{\mathrm{Tr}=1} + |1\rangle\langle 0| \otimes \underbrace{|1\rangle\langle 0|}_{\mathrm{Tr}=0} + |0\rangle\langle 1| \otimes \underbrace{|0\rangle\langle 1|}_{\mathrm{Tr}=0} + |1\rangle\langle 1| \otimes \underbrace{|1\rangle\langle 1|}_{\mathrm{Tr}=1})$$

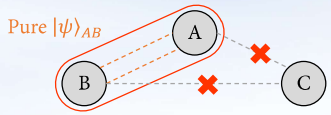
Using the *linearity* of partial trace, we directly take Tr on each second system.
$\Rightarrow \mathrm{Tr}_B[|\Phi^+_{AB}\rangle\langle\Phi^+_{AB}|] = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = I/2$, ← *maximally mixed state*.

The reduced state of a maximally entangled pure state is not pure but mixed!

- Question: Given ρ_A and ρ_B, what are the possible ρ_{AB}?



Purifications



- Theorem (Schmidt decomposition). For each vector $|\psi\rangle_{AB} \in \mathcal{H}_{AB}$, there exists orthonormal bases $\{|e_j\rangle\}_j$ of $\mathcal{H}_A$ and $\{|f_j\rangle\}_j$ of $\mathcal{H}_B$ such that

$$|\psi\rangle_{AB} = \sum_{j=1}^{\min\{|\mathcal{H}_A|,|\mathcal{H}_B|\}} \sqrt{\lambda_j}|e_j\rangle \otimes |f_j\rangle$$

$\{\sqrt{\lambda_j}\}_j$ are Schmidt coefficients

- Properties:
  - The number of non-zero λ_j is called the *Schmidt rank*.
    $\Rightarrow |\psi\rangle_{AB}$ is entangled if and only if the Schmidt rank ≥ 2.
  - The reduced states are $\rho_A = \sum_j \lambda_j |e_j\rangle\langle e_j|$ and $\rho_B = \sum_j \lambda_j |f_j\rangle\langle f_j|$, respectively.
  - Let $\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$. Then, ρ_A is pure $\Leftrightarrow$ ρ_B is pure $\Leftrightarrow$ ρ_{AB} is a product state.
  - The Schmidt decomposition gives ρ_A a *purification* $|\psi\rangle_{AB}$.
    However, this is not unique since $(I_A \otimes U_B)|\psi\rangle_{AB}$ for all U_B are purified states, too.
  - *Monogamy*: If Alice & Bob share a pure state $|\psi\rangle_{AB}$, then its extension must be $\rho_{ABC} = |\psi\rangle\langle\psi|_{AB} \otimes \rho_C$. Hence, Alice & Bob are completely uncorrelated with others.

They have the same entropy!

We call $|\psi\rangle_{AB}$ a purification of ρ_A

Properties of Partial Trace

1. $\mathrm{Tr}_B: M_{AB} \mapsto \sum_b (I_A \otimes \langle b|) M_{AB} (I_A \otimes |b\rangle)$ is a *linear superoperator*.
2. $\mathrm{Tr}_B[X_A \otimes Y_B] = X_A \otimes \mathrm{Tr}[Y_B] = \mathrm{Tr}[Y_B] \cdot X_A$.
   (We take the trace of Y_B but leave X_A untouched.)
3. If $M_{AB} \geq 0$, then $\mathrm{Tr}_B[M_{AB}] \geq 0$. ← positivity-preserving
4. $\mathrm{Tr}[M_{AB}] = \mathrm{Tr}[\mathrm{Tr}_B[M_{AB}]]$.
5. $\mathrm{Tr}[M_{AB}(X_A \otimes I_B)] = \mathrm{Tr}[\mathrm{Tr}_B[M_{AB}]X_A]$.

$= \sum_{a,b}(\langle a| \otimes \langle b|) M_{AB} (X_A \otimes I_B)(|a\rangle \otimes |b\rangle) = \sum_{a,b}(\langle a| \otimes \langle b|) M_{AB} (X_A|a\rangle \otimes |b\rangle)$

$= \sum_{a,b} \langle a|(I_A \otimes \langle b|) M_{AB} (I_A \otimes |b\rangle) X_A |a\rangle$

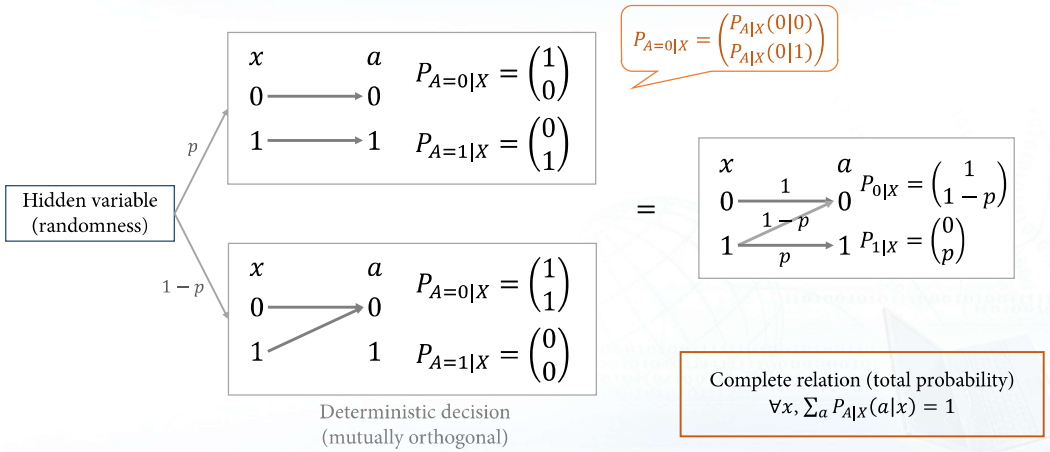
$= \sum_a \langle a| \sum_b (I_A \otimes \langle b|) M_{AB} (I_A \otimes |b\rangle)\, X_A |a\rangle$

$= \sum_a \langle a|\mathrm{Tr}_B[M_{AB}]X_A|a\rangle = \mathrm{Tr}[\mathrm{Tr}_B[M_{AB}]X_A]$ □

Postulate 4 (Measurement)

Classical Randomized Decisions

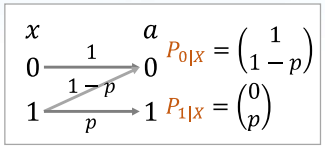
- When you read '6', what's your claim of observation? Is it '6', or 'b', or 'b'?



Classical Born's Rule

State $P_X \xrightarrow{P_{A|X}}$ outcome P_A

- Suppose that the classical system is in the state $P_X = [q, 1-q]^\dagger$, and we use the randomized decision:



- The probability of getting observation outcome $A = 0$ is:

$$\begin{aligned} \Pr(A=0) &= q \cdot 1 + (1-q) \cdot (1-p) \\ &= \sum_{x \in \{0,1\}} P_X(x) P_{A|X}(0|x) \\ &= \mathbb{E}_X\left[P_{A|X}(0|X)\right] \\ &= \langle P_X, P_{A|X}(0|X) \rangle \quad = \mathrm{Tr}\left[\begin{pmatrix} q & \\ & 1-q \end{pmatrix}\begin{pmatrix} 1 & \\ & 1-p \end{pmatrix}\right] \end{aligned}$$

P_X: priori distribution
$P_{A|X}$: conditional dist.

Measuring a Quantum System

Postulate 4⋆ (Measurement)

Positive operator-valued measure (POVM)

- A *general measurement* is described by a collection of positive semi-definite operators $\{\Pi^a\}_a$ (on the underlying Hilbert space) that satisfies the completeness relation $\sum_a \Pi^a = I$.
- The index a refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is ρ, then

$$\Pr_\rho(\text{outcome } a) = \mathrm{Tr}[\rho \Pi^a],$$

The Born Rule

and the *post-measurement state* is

$$\tilde{\rho}^a = \frac{\sqrt{\Pi^a}\rho\sqrt{\Pi^a}}{\sqrt{\Pr(a)}}.$$

Quantum Born's Rule

Quantum $\rho_X \xrightarrow{\{\Pi^a\}_a}$ classical P_A

- Deterministic decision → mutually orthogonal projections (PVMs e.g. with a basis)
  Randomized decision → POVM elements not necessary orthogonal or commute.
- Given a state ρ_X, the probability of getting the measurement outcome $A = a$ using POVM $\{\Pi^a\}_a$ is given by $\Pr(A=a) = \mathrm{Tr}[\rho_X \Pi^a] \equiv \mathbb{E}_{\rho_X}[\Pi^a] = \langle \rho_X, \Pi^a \rangle$.
- If the quantum system AB is in state ρ_{AB} and we want to measure on subsystem A using POVM $\{\Pi_A^\omega\}_\omega$ is given by $\Pr(\omega) = \mathrm{Tr}[\rho_{AB}(\Pi_A^\omega \otimes I_B)] = \mathrm{Tr}[\rho_A \Pi_A^\omega]$.
  → This means that the state ρ_A reproduces the statistics of all possible measurement on A but contains no information about B (since it is a state on A alone).

Other Examples

- Suppose a quantum device outputs $I|\psi\rangle$, $X|\psi\rangle$, $Y|\psi\rangle$, and $Z|\psi\rangle$ equally probable:

  $\Rightarrow \rho = \frac{1}{4}|\psi\rangle\langle\psi| + \frac{1}{4}X|\psi\rangle\langle\psi|X + \frac{1}{4}Y|\psi\rangle\langle\psi|Y + \frac{1}{4}Z|\psi\rangle\langle\psi|Z = I/2$ (maximally mixed).

  $\Rightarrow$ For any POVM $\{\Pi^\omega\}_\omega$, we get $\Pr(\omega) = \mathrm{Tr}[I/2 \cdot \Pi^\omega] = \mathrm{Tr}[\Pi^\omega]/2$.

$$\Pr(\omega) = \frac{1}{4}\langle\psi|\Pi^\omega|\psi\rangle + \frac{1}{4}\langle\psi|X^\dagger\Pi^\omega X|\psi\rangle + \frac{1}{4}\langle\psi|Y^\dagger\Pi^\omega Y|\psi\rangle + \frac{1}{4}\langle\psi|Z^\dagger\Pi^\omega Z|\psi\rangle = \mathrm{Tr}[\rho\Pi^\omega]$$

- Non-local games. Suppose Alice & Bob prepare state ρ_{AB} before game, and apply local measurements $\{\Pi_x^a\}_a$ and $\{\Pi_y^b\}_b$ on their systems, respectively.

  $\Rightarrow \Pr_{AB|XY}(a,b|x,y) = \mathrm{Tr}\left[\rho_{AB}\Pi_x^a \otimes \Pi_y^b\right]$.

  If the state is *separable*, i.e. $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$, then by linearity of Tr & tensors

  $\Rightarrow \Pr_{AB|XY}(a,b|x,y) = \sum_i p_i \mathrm{Tr}\left[\rho_A^i\Pi_x^a\right] \cdot \mathrm{Tr}\left[\rho_B^i\Pi_y^b\right] = \sum_i p_i \Pr_{A|X}(a|x,i) \Pr_{B|Y}(b|y,i)$.

  Classical hidden-variable theory

Concluding Remarks

Concluding Remarks (1/2)

- You need a dictionary:
- Embedding vectors to diagonal matrices.
- Classical information science for manipulating probability distributions naturally extend to that of density operators → QIS or QIP.

Classical	Quantum
Probability vector	Density operator
Sum	Trace
Marginal distribution	Reduced state
Deterministic	Pure
Probabilistic	Convex mixture
$\mathbb{E}_X[P_{\Omega\|X}(\omega\|X)]$	$\mathrm{Tr}[\rho_X\Pi^\omega]$

- A statistical framework
  - Preparation: A preparation procedure determines the *state* of a system.
  - Evolution: How is a quantum state *evolving*?
  - Measurement: A measurement procedure produces some observation outcomes.

Concluding Remarks (2/2)

- Allowing *off-diagonal phases* yields the possibility of *superposition*, which is a property of quantum states of a single system that exhibits 'quantumness'.
- Allowing *entangled* vectors or operators yields *entanglement*, which is a property of quantum states of a composite quantum system that exhibits 'quantumness'.
- *Operational meaning* of separable states: They are the largest class of bipartite states that can be prepared by Alice and Bob in their own labs by using local quantum operations and classical communications (LOCC).
- We only talked about pure entangled states $|\psi\rangle_{AB}$. How about other entangled states? Mixed & multipartite entangled states are less well understood.
- Why mixed states? (i) QM + randomness → QIS; (ii) Reduced states could be mixed.
- Why POVM? (i) randomized decisions; (ii) with ancillas.

Quantum Information and Computation
Open Quantum Systems

Hao-Chung Cheng (鄭皓中)
haochung@ntu.edu.tw

Department of Electrical Engineering
National Taiwan University

April 28, 2021

Motivation

Motivation

Preparation → Evolution → Measurement → Data

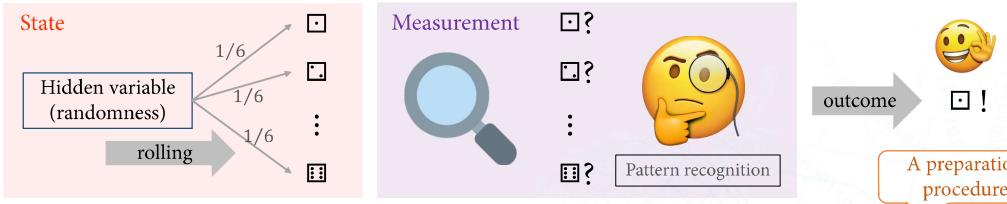
- What we have learnt so far (for closed quantum systems):
  - Pure states: unit vectors (up to a global phase) in a Hilbert space.
  - Reversible evolutions: unitary operations on the Hilbert space.
  - Projection measurements: often measuring with respect to an orthonormal basis.
- What if the underlying closed quantum system is interacting with an environment or there is some hidden noise/randomness that we are not in control of?
- Does quantum theory fully capture a statistical framework?
- Today: We will incorporate *probability reasoning* into the previous formalism.
- We will find a correspondence between the closed & open quantum systems.
- If there is such a correspondence, what's the point of studying open quantum stuff ?

  → More resources to *purify* everything! Essentially, closed quantum ↔ God's view.

Postulate 1 (State)

Probabilistic Classical Systems and States

- To a *probabilistic/randomized* classical system where each symbol in the alphabet occurs with certain probability, we use a *probability model* to characterize it.
  - $\mathcal{X} = \{\boxdot, \boxdot, \boxdot, \boxdot, \boxdot, \boxdot\}$ for a *unbiased* dice: $\Pr(x) = 1/6, \forall x \in \mathcal{X}$.



A preparation procedure

- The *state* of a dice is associated with a *probability distribution*, say P_X.
  - A *biased* dice: $P_X = [1/2, 1/3, 1/6, 0, 0, 0]^\dagger$; a *deterministic* dice: $P_X = [1, 0, 0, 0, 0, 0]^\dagger$.
- An *observation* on $\mathcal{X}$ gives an *realization* of the random variable X.
  → This is called a *measurement outcome*.

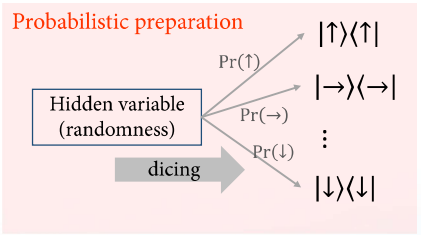
Once a measurement is performed, the state collapses!

Open Quantum Systems and States (1/2)

- Previously, a *pure* quantum state (of a closed quantum system) is associated with a unit vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$.

A deterministic preparation

- A unit vector $|\psi\rangle$ (up to a global phase, i.e. $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$) is by a one-to-one correspondence to a *rank-one projection* $|\psi\rangle\langle\psi|$ to the subspace spanned by $|\psi\rangle$.
- One can roll a dice to prepare each rank-one projection with certain probability. This is a (probabilistic) *preparation procedure* for a quantum system.



Mathematically, we use convex combination of rank-one projection to describe such a probabilistic preparation.

Mixed quantum state:

$$\Pr(\uparrow)\,|\uparrow\rangle\langle\uparrow| + \Pr(\rightarrow)\,|\rightarrow\rangle\langle\rightarrow| + \cdots + \Pr(\downarrow)\,|\downarrow\rangle\langle\downarrow|$$

Open Quantum Systems and States (2/2)

Postulate 1⋆ (State)

- To any physical (quantum-mechanical) system, there is associated a Hilbert space, i.e. a complex inner product space.
- The system is completely described by its *density operator*, which is a positive semi-definite operator with unit trace acting on the Hilbert space.

- Density operator/matrix or *mixed state*: $\rho \geq 0$, $\mathrm{Tr}[\rho] = 1$.
- A classical probability distribution is $p_i \geq 0$, $\sum_i p_i = 1$.

Probability vector → Matrix
Summation → Trace

Remarks on Density Operators

Pure state

Mixed state

- Any convex combination of rank-one projections gives a density operator.

  Conversely, any density operator ρ can be written in the form:

off-diagonal phase terms

eigenvectors

$$\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i| = U \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_d \end{pmatrix} U^\dagger = \begin{pmatrix} \rho_{11} & \cdots & \rho_{1d} \\ \vdots & \ddots & \vdots \\ \rho_{d1} & \cdots & \rho_{dd} \end{pmatrix}$$

eigenvalues

Will give its *entropy*

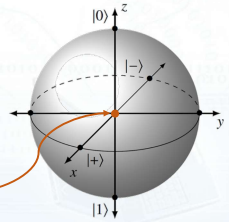
  The eigenvalues $\{\lambda_i\}_i$ form a probability distribution, i.e. $\lambda_i \geq 0, \sum_i \lambda_i = 1$.

- Such decomposition is *not unique* due to the multiplicity of eigenvalues (but the decomposition into eigenprojections/eigenspaces is unique). Note that the decomposition (into a *simplex*) of a probability distribution is unique.
- A classical distribution P_X can be *embedded* into a (diagonal) density operator ρ_X: Fix a computation basis $\{|x\rangle\}_{x\in\mathcal{X}}$; write $\rho_X = \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x|$.

Classical state

Properties of Quantum States

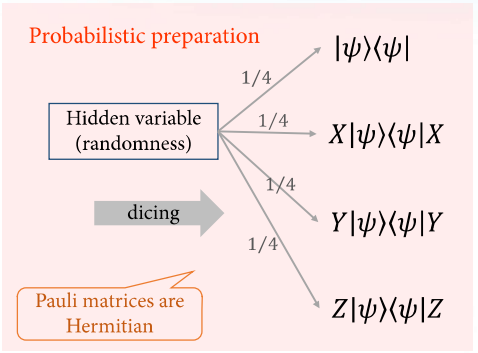
- The set of quantum states $\mathcal{S}(\mathcal{H})$ is *convex* (just as the set of probability distributions).
  - Given $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ and $t \in [0,1]$, then $t\rho + (1-t)\sigma \in \mathcal{S}(\mathcal{H})$ is a gain a state.
  - In general, for any ensemble $\{p_i, \rho_i\}_i$, the mixture $\rho = \sum_i p_i \rho_i$ is a gain a state.
  - Conversely, a state ρ admits *infinitely many* decompositions $\rho = \sum_i p_i \rho_i$.
- Extremal elements of $\mathcal{S}(\mathcal{H})$ are the pure states;
  state that has eigenvalue 0 belongs to the boundary of $\mathcal{S}(\mathcal{H})$.
- Let $\mathcal{H} = \mathbb{C}^d$. There are at most d mutually orthogonal pure states, whereas the dimension of $\mathcal{S}(\mathcal{H})$ is $d^2 - 1$.
  → Bloch-vector representation: ∃ traceless operators $E_1, \dots E_{d^2-1}$
  such that $\langle E_i, E_j \rangle = d\delta_{ij}$ and $\rho = \frac{1}{d}\left(I + \sum_{i=1}^{d^2-1} r_i E_i\right)$.
  [Bengtsson–Życzkowski, 2006]
- E.g. Maximally mixed state I/d.
  → Eigenvalues form a uniform distribution $[1/d, \dots, 1/d]^\dagger$.



An Example of Maximally Mixed State

$|\psi\rangle = a|0\rangle + b|1\rangle$

- Suppose a quantum device outputs $I|\psi\rangle$, $X|\psi\rangle$, $Y|\psi\rangle$, and $Z|\psi\rangle$ equally probable:



$$= \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} a^* & b^* \end{pmatrix} = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}$$

$$= \begin{pmatrix} b \\ a \end{pmatrix} \begin{pmatrix} b^* & a^* \end{pmatrix} = \begin{pmatrix} |b|^2 & a^*b \\ ab^* & |a|^2 \end{pmatrix}$$

$$= \begin{pmatrix} b \\ -a \end{pmatrix} \begin{pmatrix} b^* & -a^* \end{pmatrix} = \begin{pmatrix} |b|^2 & -a^*b \\ -ab^* & |a|^2 \end{pmatrix}$$

$$= \begin{pmatrix} a \\ -b \end{pmatrix} \begin{pmatrix} a^* & -b^* \end{pmatrix} = \begin{pmatrix} |a|^2 & -ab^* \\ -a^*b & |b|^2 \end{pmatrix}$$

$$\Rightarrow \rho = \frac{1}{4}|\psi\rangle\langle\psi| + \frac{1}{4}X|\psi\rangle\langle\psi|X + \frac{1}{4}Y|\psi\rangle\langle\psi|Y + \frac{1}{4}Z|\psi\rangle\langle\psi|Z = \frac{1}{2}I$$

Maximally mixed qubit state

Postulate 2 (Composite System)

Composite Classical Systems (1/2)

- *State* of a composition of classical system is given by a *joint probability distribution*.
  - $\mathcal{X} \times \mathcal{Y} = \{\boxdot, \boxdot, \boxdot, \boxdot, \boxdot, \boxdot\} \times \{\spadesuit, \clubsuit, \heartsuit, \diamondsuit\}$, and
  - For a P_{XY}, we have $P_{XY}(x,y) \geq 0$ and $\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} P_{XY}(x,y) = 1$.
  - Marginal distribution: $P_X(x) \coloneqq \sum_{y\in\mathcal{Y}} P_{XY}(x,y)$.

$|xy\rangle \coloneqq |x\rangle \otimes |y\rangle$

- Fix bases $\{|x\rangle\}_{x\in\mathcal{X}}$, $\{|y\rangle\}_{y\in\mathcal{Y}}$, and $\{|xy\rangle\}_{(x,y)\in\mathcal{X}\times\mathcal{Y}}$.
  Let's write P_{XY} in terms of a density operator ρ_{XY} on Hilbert space spanned by $\{|xy\rangle\}$.

$$\rho_{XY} = \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} P_{XY}(x,y)|xy\rangle\langle xy| = \begin{pmatrix} P_{XY}(0,0) & & & \\ & P_{XY}(0,1) & & \\ & & \ddots & \\ & & & P_{XY}(|\mathcal{X}|-1, |\mathcal{Y}|-1) \end{pmatrix}$$

$$\rho_X = \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x| = \begin{pmatrix} P_X(0) & & & \\ & P_X(1) & & \\ & & \ddots & \\ & & & P_X(|\mathcal{X}|-1) \end{pmatrix}$$

Composite Classical Systems (2/2)

i: index; X/Y: systems

- *Classical correlation*: Any classical joint state can be written as $\sum_i p_i \rho_X^i \otimes \rho_Y^i$.

$$\rho_{XY} = \sum_{(x,y)} P_{XY}(x,y)|xy\rangle\langle xy| = \sum_{(x,y)} P_{XY}(x,y)|x\rangle \otimes |y\rangle\langle x| \otimes \langle y|$$
$$= \sum_{(x,y)} P_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y|$$

$|x\rangle\langle x|$ projects onto span$\{|x\rangle\}$

- *Independence*: $P_{XY}(x,y) = P_X(x) \cdot P_Y(y)$ for all $(x,y) \in \mathcal{X} \times \mathcal{Y}$.

Independence

$$\rho_{XY} = \sum_{(x,y)} P_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| = \sum_{(x,y)} P_X(x) \cdot P_Y(y)|x\rangle\langle x| \otimes |y\rangle\langle y|$$

Linearity

$$= \sum_{(x,y)} P_X(x)|x\rangle\langle x| \otimes P_Y(y)|y\rangle\langle y| = \sum_x P_X(x)|x\rangle\langle x| \otimes \sum_y P_Y(y)|y\rangle\langle y|$$
$$= \rho_X \otimes \rho_Y$$

$\Rightarrow$ The product state $\rho_X \otimes \rho_Y$ means that system X is *independent* of system Y.

Examples

- Perfect (*maximally correlated*) classical state on $\mathcal{X} \times \mathcal{Y} = \mathbb{Z}_2 \times \mathbb{Z}_2$:
  $P_{XY}(0,0) = P_{XY}(1,1) = 1/2$.

$$\Rightarrow \rho_{XY} = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11| = \begin{pmatrix} 1/2 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1/2 \end{pmatrix}$$

- Uniform (*maximally mixed*) classical state on $\mathcal{X} \times \mathcal{Y} = \mathbb{Z}_2 \times \mathbb{Z}_2$:
  $P_{XY}(0,0) = P_{XY}(0,1) = P_{XY}(1,0) = P_{XY}(1,1) = 1/4$. $\Rightarrow P_X(x) = P_Y(y) = 1/2$.

Independence

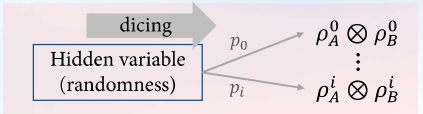
$$\rho_{XY} = \begin{pmatrix} 1/4 & & & \\ & 1/4 & & \\ & & 1/4 & \\ & & & 1/4 \end{pmatrix} = \begin{pmatrix} 1/2 & \\ & 1/2 \end{pmatrix} \otimes \begin{pmatrix} 1/2 & \\ & 1/2 \end{pmatrix} = \frac{I}{2} \otimes \frac{I}{2}$$

Composite Quantum Systems

Postulate 2* (Composition)

- For a joint system composed of two subsystems with Hilbert space $\mathcal{H}_A$ and $\mathcal{H}_B$, the Hilbert space is the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$ (e.g. $\mathbb{C}^d \otimes \mathbb{C}^d = \mathbb{C}^{d \times d}$).
- If the system i is *independently* prepared in the state ρ_i, the joint state of the total system is $\bigotimes_i \rho_i$, called the *product state*.
- In general, a quantum state on a joint system $\mathcal{H}_A \otimes \mathcal{H}_B$ is given by an associated density operator ρ_{AB}.
  The state describing subsystem $\mathcal{H}_A$ is the *reduced state* $\rho_A := \mathrm{Tr}_B[\rho_{AB}]$.

Entangled States



- Definition (Separability). A state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$ is separable if $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$ for some prob. distribution $\{p_i\}_i$ and states $\rho_{A/B}^i \in \mathcal{S}(\mathcal{H}_{A/B})$.
- Remark. Every classical joint distribution is separable, but a separable state is not necessary classical (diagonal) $\because \rho_A^i$ and ρ_B^i might not *commute*.

i.e. $AB \neq BA$

- Definition (Entanglement). A state ρ_{AB} on $\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is entangled if $\rho_{AB} \neq \sum_i p_i \rho_A^i \otimes \rho_B^i$ for any prob. distribution $\{p_i\}_i$ and states $\rho_{A/B}^i \in \mathcal{S}(\mathcal{H}_{A/B})$.
- Property: A pure state (rank-one projection) is *separable* if and only if it is *product*.

Maximally entangled state:

$$|\Phi_{AB}^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow \rho_{AB} = |\Phi_{AB}^+\rangle\langle\Phi_{AB}^+| = \begin{pmatrix} 1/2 & & & 1/2 \\ & 0 & & \\ & & 0 & \\ 1/2 & & & 1/2 \end{pmatrix}$$

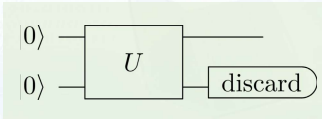
Partial Trace and Reduced States

- Suppose we are given a quantum state ρ_{AB} on a quantum system AB composed of two subsystems A and B with overall Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.
  Which state should we use to describe the state of subsystem A alone?
- Classically, recall that the *marginal* distribution is given by $P_X(x) := \sum_{y \in \mathcal{Y}} P_{XY}(x, y)$.
- Definition (Partial trace). For every operator M_{AB} on $\mathcal{H}_{AB}$,

$$\mathrm{Tr}_B[M_{AB}] := \sum_b (I_A \otimes \langle b|) M_{AB} (I_A \otimes |b\rangle)$$

Linear operation

  where $\{|b\rangle\}_b$ is any *orthonormal* basis of $\mathcal{H}_B$.
- Definition (Reduced states). Given a state ρ_{AB} on a quantum system AB, we define its *reduced state* on subsystem A by $\rho_A := \mathrm{Tr}_B[\rho_{AB}]$.
- Operationally, it is equivalent to ignoring the other systems.



Examples (1/2)

$\rho_{XY} = |x_0y_0\rangle\langle x_0y_0| = |x_0\rangle\langle x_0| \otimes |y_0\rangle\langle y_0|$

- Reduced state of a classical deterministic state $P_{XY}(x_0, y_0) = 1$ for a (x_0, y_0):

$$\begin{aligned}\Rightarrow \mathrm{Tr}_Y[\rho_{XY}] &= \textstyle\sum_y (I_X \otimes \langle y|)|x_0\rangle\langle x_0| \otimes |y_0\rangle\langle y_0|(I_X \otimes |y\rangle) \\ &= \textstyle\sum_y (|x_0\rangle\langle x_0| \otimes \langle y|y_0\rangle\langle y_0|)(I_X \otimes |y\rangle) = \sum_y |x_0\rangle\langle x_0| \otimes \langle y|y_0\rangle\langle y_0|y\rangle \\ &= |x_0\rangle\langle x_0| \otimes \langle y_0|y_0\rangle\langle y_0|y_0\rangle = |x_0\rangle\langle x_0|\end{aligned}$$

The marginal of a joint deterministic (pure) classical state is still deterministic (pure).

- Reduced state of a product state $\rho_{AB} = \rho_A \otimes \rho_B$:

Classically, $\sum_y P_X(x)P_Y(y) = P_X(x)$

$$\begin{aligned}\Rightarrow \mathrm{Tr}_B[\rho_{AB}] &= \textstyle\sum_b (I_A \otimes \langle b|)\rho_A \otimes \rho_B (I_A \otimes |b\rangle) = \sum_b (\rho_A \otimes \langle b|\rho_B)(I_A \otimes |b\rangle) \\ &= \textstyle\sum_b \rho_A \otimes \langle b|\rho_B|b\rangle = \rho_A \otimes \mathrm{Tr}[\rho_B] = \rho_A.\end{aligned}$$

- Question: Are the reduced states of a pure joint state pure as well?

Examples (2/2)

- Reduced state of a maximally entangled state $\rho_{AB} = |\Phi^+_{AB}\rangle\langle\Phi^+_{AB}|$:

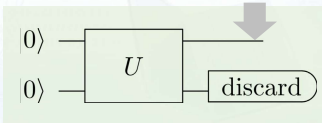
$$\begin{aligned}&= \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|) \\ &= \frac{1}{2}(|0\rangle\langle 0| \otimes \underbrace{|0\rangle\langle 0|}_{\mathrm{Tr}=1} + |1\rangle\langle 0| \otimes \underbrace{|1\rangle\langle 0|}_{\mathrm{Tr}=0} + |0\rangle\langle 1| \otimes \underbrace{|0\rangle\langle 1|}_{\mathrm{Tr}=0} + |1\rangle\langle 1| \otimes \underbrace{|1\rangle\langle 1|}_{\mathrm{Tr}=1})\end{aligned}$$

Using the *linearity* of partial trace, we directly take Tr on each second system.
$\Rightarrow \mathrm{Tr}_B[|\Phi^+_{AB}\rangle\langle\Phi^+_{AB}|] = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = I/2$, ← *maximally mixed state*.

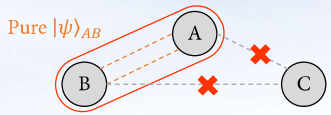
The reduced state of a maximally entangled pure state is not pure but mixed!

- Question: Given ρ_A and ρ_B, what are the possible ρ_{AB}?

could be mixed!



Purifications



- Theorem (Schmidt decomposition). For each vector $|\psi\rangle_{AB} \in \mathcal{H}_{AB}$, there exists orthonormal bases $\{|e_j\rangle\}_j$ of $\mathcal{H}_A$ and $\{|f_j\rangle\}_j$ of $\mathcal{H}_B$ such that

$$|\psi\rangle_{AB} = \sum_{j=1}^{\min\{|\mathcal{H}_A|, |\mathcal{H}_B|\}} \sqrt{\lambda_j}\,|e_j\rangle \otimes |f_j\rangle$$

$\{\sqrt{\lambda_j}\}_j$ are Schmidt coefficients

- Properties:
  - The number of non-zero λ_j is called the *Schmidt rank*.
    $\Rightarrow |\psi\rangle_{AB}$ is entangled if and only if the Schmidt rank ≥ 2.
  - The reduced states are $\rho_A = \sum_j \lambda_j |e_j\rangle\langle e_j|$ and $\rho_B = \sum_j \lambda_j |f_j\rangle\langle f_j|$, respectively.

They have the same entropy!

  - Let $\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$. Then, ρ_A is pure $\Leftrightarrow$ ρ_B is pure $\Leftrightarrow$ ρ_{AB} is a product state.
  - The Schmidt decomposition gives ρ_A a *purification* $|\psi\rangle_{AB}$.
    However, this is not unique since $(I_A \otimes U_B)|\psi\rangle_{AB}$ for all U_B are purified states, too.

We call $|\psi\rangle_{AB}$ a purification of ρ_A

  - *Monogamy*: If Alice & Bob share a pure state $|\psi\rangle_{AB}$, then its extension must be $\rho_{ABC} = |\psi\rangle\langle\psi|_{AB} \otimes \rho_C$. Hence, Alice & Bob are completely uncorrelated with others.

Properties of Partial Trace

1. $\mathrm{Tr}_B : M_{AB} \mapsto \sum_b (I_A \otimes \langle b|) M_{AB} (I_A \otimes |b\rangle)$ is a *linear superoperator*.
2. $\mathrm{Tr}_B[X_A \otimes Y_B] = X_A \otimes \mathrm{Tr}[Y_B] = \mathrm{Tr}[Y_B] \cdot X_A$.
   (We take the trace of Y_B but leave X_A untouched.)
3. If $M_{AB} \geq 0$, then $\mathrm{Tr}_B[M_{AB}] \geq 0$. ← positivity-preserving
4. $\mathrm{Tr}[M_{AB}] = \mathrm{Tr}\big[\mathrm{Tr}_B[M_{AB}]\big]$.
5. $\mathrm{Tr}[M_{AB}(X_A \otimes I_B)] = \mathrm{Tr}[\mathrm{Tr}_B[M_{AB}]X_A]$.

$= \sum_{a,b} (\langle a| \otimes \langle b|) M_{AB} (X_A \otimes I_B)(|a\rangle \otimes |b\rangle) = \sum_{a,b} (\langle a| \otimes \langle b|) M_{AB} (X_A|a\rangle \otimes |b\rangle)$

$= \sum_{a,b} \langle a|(I_A \otimes \langle b|) M_{AB} (I_A \otimes |b\rangle) X_A|a\rangle$

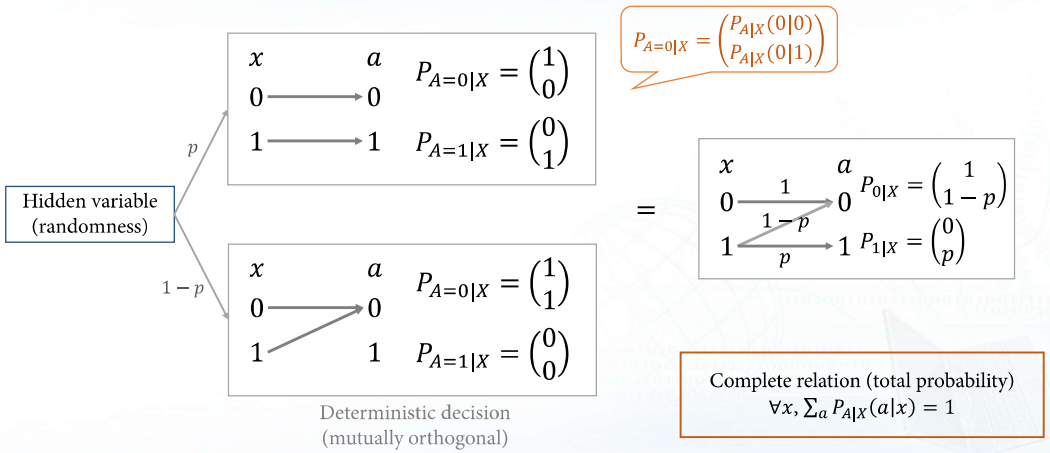
$= \sum_a \langle a| \sum_b (I_A \otimes \langle b|) M_{AB} (I_A \otimes |b\rangle)\, X_A|a\rangle$

$= \sum_a \langle a|\mathrm{Tr}_B[M_{AB}]X_A|a\rangle \;\; = \mathrm{Tr}[\mathrm{Tr}_B[M_{AB}]X_A]$ □

Postulate 4 (Measurement)

Classical Randomized Decisions

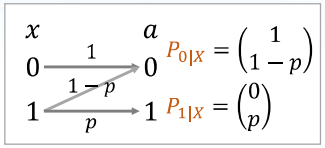
- When you read '6', what's your claim of observation? Is it '6', or 'b', or 'ɓ'?



Classical Born's Rule

State $P_X \xrightarrow{P_{A|X}}$ outcome P_A

- Suppose that the classical system is in the state $P_X = [q, 1-q]^\dagger$, and we use the randomized decision:



- The probability of getting observation outcome $A = 0$ is:

$$\Pr(A=0) = q \cdot 1 + (1-q) \cdot (1-p)$$
$$= \sum_{x \in \{0,1\}} P_X(x) P_{A|X}(0|x)$$
$$= \mathbb{E}_X\big[P_{A|X}(0|X)\big]$$
$$= \langle P_X, P_{A|X}(0|X)\rangle \;\; = \mathrm{Tr}\left[\begin{pmatrix} q & \\ & 1-q \end{pmatrix}\begin{pmatrix} 1 & \\ & 1-p \end{pmatrix}\right]$$

P_X: priori distribution
$P_{A|X}$: conditional dist.

Measuring a Quantum System

Postulate 4* (Measurement)

Positive operator-valued measure (POVM)

- A *general measurement* is described by a collection of positive semi-definite operators $\{\Pi^a\}_a$ (on the underlying Hilbert space) that satisfies the completeness relation $\sum_a \Pi^a = I$.
- The index a refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is ρ, then

$$\Pr_\rho(\text{outcome } a) = \mathrm{Tr}[\rho\Pi^a],$$

The Born Rule

and the *post-measurement state* is

$$\tilde{\rho}^a = \frac{\sqrt{\Pi^a}\rho\sqrt{\Pi^a}}{\sqrt{\Pr(a)}}.$$

Quantum Born's Rule

Quantum $\rho_X \xrightarrow{\{\Pi^a\}_a}$ classical P_A

- Deterministic decision → mutually orthogonal projections (PVMs e.g. with a basis)
  Randomized decision → POVM elements not necessary orthogonal or commute.
- Given a state ρ_X, the probability of getting the measurement outcome $A = a$ using POVM $\{\Pi^a\}_a$ is given by $\Pr(A = a) = \mathrm{Tr}[\rho_X\Pi^a] \equiv \mathbb{E}_{\rho_X}[\Pi^a] = \langle\rho_X, \Pi^a\rangle$.
- If the quantum system AB is in state ρ_{AB} and we want to measure on subsystem A using POVM $\{\Pi_A^\omega\}_\omega$ is given by $\Pr(\omega) = \mathrm{Tr}[\rho_{AB}(\Pi_A^\omega \otimes I_B)] = \mathrm{Tr}[\rho_A\Pi_A^\omega]$.
  → This means that the state ρ_A reproduces the statistics of all possible measurement on A but contains no information about B (since it is a state on A alone).

Other Examples

- Suppose a quantum device outputs $I|\psi\rangle$, $X|\psi\rangle$, $Y|\psi\rangle$, and $Z|\psi\rangle$ equally probable:
  $\Rightarrow \rho = \frac{1}{4}|\psi\rangle\langle\psi| + \frac{1}{4}X|\psi\rangle\langle\psi|X + \frac{1}{4}Y|\psi\rangle\langle\psi|Y + \frac{1}{4}Z|\psi\rangle\langle\psi|Z = I/2.$ maximally mixed state
  $\Rightarrow$ For any POVM $\{\Pi^\omega\}_\omega$, we get $\Pr(\omega) = \mathrm{Tr}[I/2 \cdot \Pi^\omega] = \mathrm{Tr}[\Pi^\omega]/2.$

$$\Pr(\omega) = \frac{1}{4}\langle\psi|\Pi^\omega|\psi\rangle + \frac{1}{4}\langle\psi|X^\dagger\Pi^\omega X|\psi\rangle + \frac{1}{4}\langle\psi|Y^\dagger\Pi^\omega Y|\psi\rangle + \frac{1}{4}\langle\psi|Z^\dagger\Pi^\omega Z|\psi\rangle = \mathrm{Tr}[\rho\Pi^\omega]$$

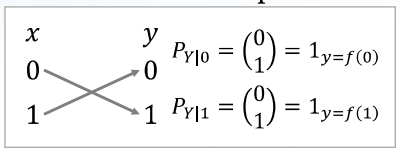
- Non-local games. Suppose Alice & Bob prepare state ρ_{AB} before game, and apply local measurements $\{\Pi_x^a\}_a$ and $\{\Pi_y^b\}_b$ on their systems, respectively.
  $\Rightarrow \Pr_{AB|XY}(a, b|x, y) = \mathrm{Tr}[\rho_{AB}\Pi_x^a \otimes \Pi_y^b].$
  If the state is *separable*, i.e. $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$, then by linearity of Tr & tensors
  $\Rightarrow \Pr_{AB|XY}(a, b|x, y) = \sum_i p_i \mathrm{Tr}[\rho_A^i\Pi_x^a] \cdot \mathrm{Tr}[\rho_B^i\Pi_y^b] = \sum_i p_i \Pr_{A|X}(a|x, i) \Pr_{B|Y}(b|y, i).$
  classical hidden-variable theory

Postulate 3 (Evolution)

Classical Channels/Operations

Input $P_X \xrightarrow{P_{Y|X}}$ Output P_A

- A *classical channel/operation* can be viewed as a black-box which takes input a value $x \in \mathcal{X}$ and output some value $y \in \mathcal{Y}$.
  - Deterministic function: $y = f(x)$.
    Given $P_X = [p, 1-p]^T$, $P_Y(y) = [1-p, p]^T$.
    → To any f, we can only focus on how $P_X \mapsto P_Y$.

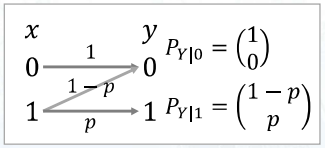


- The channel can be inherently *random*, mathematically it can be described by a *conditional probability distribution* $P_{Y|X}$, i.e. to each input x, it outputs $P_{Y|X}(\cdot|x)$.

  Hence, given an input distribution, we calculate $P_Y(y) = \sum_{x\in\mathcal{X}} P_X(x)P_{Y|X}(y|x)$.

- Equivalently, one can write $\{P_{Y|X}(y|x)\}$ as a probabilistic transition matrix, so it take probability vector P_X and outputs a probability vector P_X.

  → A channel $P_{Y|X}$ takes input P_X and outputs P_Y.



Quantum Channels/Operations

- A *quantum channel/operation*, say $\mathcal{N}^{A\to B}: \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}_B)$, is the most general transformation that can be performed on a given ensemble of quantum systems.
- It maps a quantum state to another quantum state.

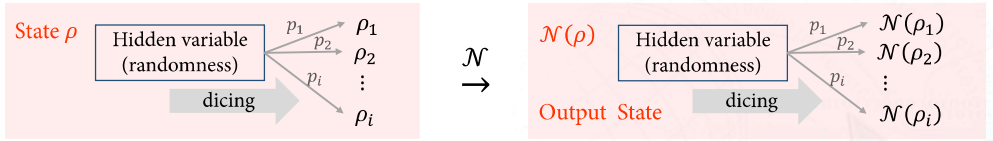
Sometimes could be subnormalized.

Also called a *superoperator*

ρ_A → Evolution $\mathcal{N}^{A\to B}$ → ρ_B

Any linear map satisfies that.

1. It preserves convex combinations: $\mathcal{N}(\sum_i p_i\rho_i) = \sum_i p_i\mathcal{N}(\rho_i)$.



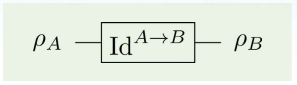
2. Trace preserving: $\mathrm{Tr}[\mathcal{N}(\rho)] = \mathrm{Tr}[\rho]$.
3. Positive: $\mathcal{N}(\rho) \geq 0, \forall\rho \geq 0$.

A quantum state ρ should be $\rho \geq 0$ and $\mathrm{Tr}[\rho] = 1$.

Examples of Quantum Channels (1/2)

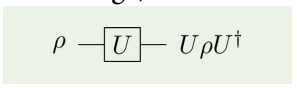
Used in teleportation

- Identity (noiseless) channel: The box does not change the input at all (or no box).

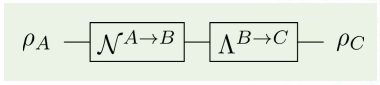


Closed quantum evolution

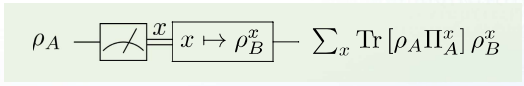
- Unitary transformation (basis change): For a fixed unitary or isometry,



- Concatenating: Given two channels $\mathcal{N}^{A\to B}$ and $\Lambda^{B\to C}$, $\rho_C = \Lambda^{B\to C}(\mathcal{N}^{A\to B}(\rho_A))$:



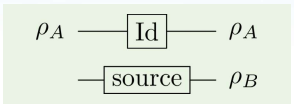
- Measure and Prepare: Given a POVM $\{\Pi_A^\omega\}_\omega$, and collection of states $\{\rho_B^\omega\}_\omega$,



Examples of Quantum Channels (2/2)

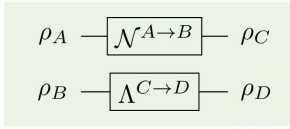
- Appending state: $\mathcal{N}^{A\to AB}(\rho_A) = \rho_A \otimes \rho_B$.

The initial state of the environment is independent of the channel input.

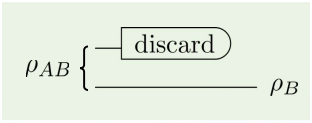


- Product of channels: Given $\mathcal{N}^{A\to B}$ and $\Lambda^{C\to D}$,

The channel acts on $\mathcal{H}_A \otimes \mathcal{H}_B$.



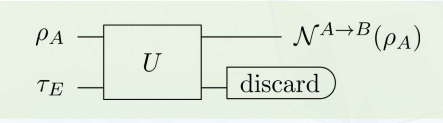
- Partial Trace: $\mathcal{N}^{AB\to B} = \mathrm{Tr}_A$.



Counterexamples of positivity: for $A, B \geq 0$,

- Neither AB nor A/B are Hermitian.
- Symmetric product: $B \mapsto (AB + BA)/2 \not\geq 0$.
- Exponential differentiation: $\frac{\partial}{\partial t}e^{A+tB}|_{t=0} = \int_0^1 e^{uA}Be^{(1-u)A}\,du \not\geq 0$.
- Antiunitary operator U such that $\langle Ux, Uy\rangle = \overline{\langle x, y\rangle}$ is positive but not CP.

- $\mathcal{N}^{A\to B}(\rho_A) \coloneqq \mathrm{Tr}_E[U\rho_A \otimes \tau_E U^\dagger]$.



Partial Transposition

- Definition (Transposition). Given a fixed basis, say $\{|i\rangle\}_{i=1}^{d}$, *transposition* is a linear mapping: $T: |i\rangle\langle j| \mapsto |j\rangle\langle i|$, or equivalently, $\left[M_{ij}\right]^T = \left[M_{ji}\right]$ for any matrix M.
  The *partial transpose* (on the first system) is defined as $T \otimes \mathrm{Id}$.

Positive and trace-preserving

$|\Phi^+_{AB}\rangle\langle\Phi^+_{AB}|$ { —[T]— ?

$\frac{1}{2}\begin{pmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{pmatrix}$

$$|\Phi^+_{AB}\rangle\langle\Phi^+_{AB}| = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|)$$
$$= \frac{1}{2}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)$$
$$\xrightarrow{T \otimes \mathrm{Id}} = \frac{1}{2}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 1| \otimes |1\rangle\langle 0| + |1\rangle\langle 0| \otimes |0\rangle\langle 1| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|)$$
$$= \frac{1}{2}\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{pmatrix}$$

← This has eigenvalue $-\frac{1}{2}$ and hence not a state!

Complete Positivity

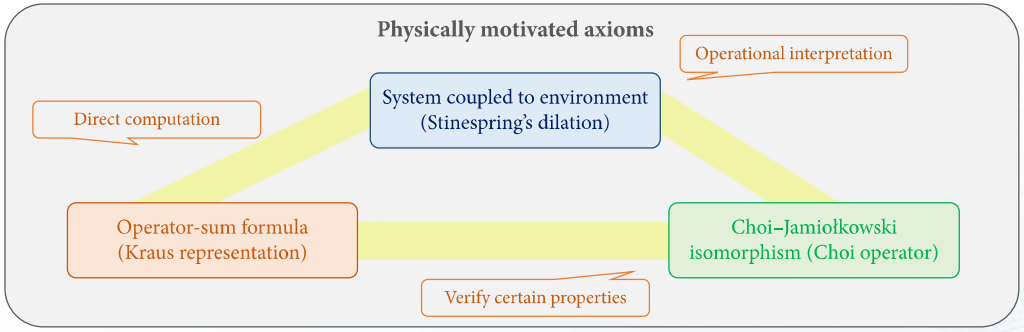
- Definition (Complete positivity). A linear map $\mathcal{N}^{A\to B}$ is *completely positive* if the map $\mathcal{N}^{A\to B} \otimes \mathrm{Id}_R$ is positive for all extensions $\mathcal{H}_R$.

ρ_{AR} { —[$\mathcal{N}^{A\to B}$]—

- Definition (Quantum channel). A quantum channel is a linear superoperator and that is completely positive and trace-preserving (CPTP).
- *Remark.* In a classical world, if $P_{Y|X}$ is a conditional probability distribution, then so is $P_{YZ|XZ}(y, z'|x, z) = P_{Y|X}(y|x)\delta_{z,z'}$ (tensoring the transition matrix with I_Z).
- Exercise:
  - To show all the above-mentioned ones (except the partial transpose) are CPTP maps.
  - Let K be a matrix from $\mathcal{H}_A$ to $\mathcal{H}_B$, then the simple operation $\rho \mapsto K\rho K^\dagger$ is CP.
- When is a superoperator completely positive? How to check it?

Characterizations

- We have introduced quantum channels as transformations of quantum states that satisfy certain mathematical properties in order to preserve the basic statistical framework of a quantum system (physically motivated axioms).
  In the following, we will give different characterizations.



System Coupled To Environment

- Theorem (Stinespring's dilation). For any quantum channel $\mathcal{N}: \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}_B)$, there exists a Hilbert space $\mathcal{H}_E$ and an isometry $V: \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ such that $\mathcal{N}(M) = \mathrm{Tr}_E\left[VMV^\dagger\right]$ for all linear operators M on $\mathcal{H}_A$.

  Equivalently, there exists a pure state $\tau_E \in \mathcal{S}(\mathcal{H}_E)$ and a unitary U on $\mathcal{H}_B \otimes \mathcal{H}_E$ such that $\mathcal{N}(M) = \mathrm{Tr}_E\left[U(\rho_A \otimes \tau_E)U^\dagger\right]$ for all $\rho_A \in \mathcal{S}(\mathcal{H}_A)$.

ρ_A —
τ_E — [U] — $\mathcal{N}^{A\to B}(\rho_A)$
discard

- One can view U as a *purification* (*not unique*) as the noisy quantum channel $\mathcal{N}$.
- Why noisy channel? → The evolution is coupled (*entangled*) to some inaccessible environment, and we only reach to *partial information* on the output system.

Operator-Sum Formula

- Theorem (Kraus representation). For any quantum channel $\mathcal{N}: \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}_B)$, there exists operators $X_1, \dots, X_r: \mathcal{H}_A \to \mathcal{H}_B$ with $r \le \dim(\mathcal{H}_A)\dim(\mathcal{H}_B)$ such that $\mathcal{N}(M) = \sum_{i=1}^r X_i M X_i^\dagger$ for all linear operators M on $\mathcal{H}_A$.

  $X_1, \dots, X_r$ are called *Kraus operators*

  Proof (Equivalence to Stinespring's dilation).

  "⇐": Define $\mathcal{H}_E = \mathbb{C}^r$. Let $V \coloneqq \sum_{i=1}^r X_i \otimes |i\rangle$.

  Then, $\mathrm{Tr}_E\left[VMV^\dagger\right] = \mathrm{Tr}_E\left[\sum_{i=1}^r X_i M \otimes |i\rangle V^\dagger\right]$

  $= \mathrm{Tr}_E\left[\sum_{i,j=1}^r X_i M X_j^\dagger \otimes |i\rangle\langle j|\right] = \sum_{i=1}^r X_i M X_i^\dagger$.

  "⇒": Let $X_i \coloneqq (I_B \otimes \langle i|)V$.

  Then, $\sum_{i=1}^r X_i M X_i^\dagger = \sum_{i=1}^r (I_B \otimes \langle i|)VMV^\dagger(I_B \otimes |i\rangle) = \mathrm{Tr}_E\left[VMV^\dagger\right]$. □

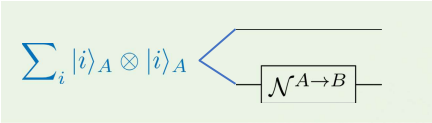
Choi–Jamiołkowski isomorphism (1/2)

- Definition (Choi operator). For a superoperator $\mathcal{N}^{A\to B}$, we define the *Choi operator* by $J_{AB}^{\mathcal{N}} \coloneqq (I_A \otimes \mathcal{N}^{A\to B})\left(\sum_{i,j}|ii\rangle\langle jj|\right) = \sum_{i,j}|i\rangle\langle j| \otimes \mathcal{N}^{A\to B}(|i\rangle\langle j|)$,

  where $\{|i\rangle\}_i$ denotes any orthonormal basis of $\mathcal{H}_A$,

  and $\sum_i |ii\rangle \in \mathcal{H}_A \otimes \mathcal{H}_A$ is an *unnormalized* maximally entangled state (MES).

  called *Choi state* if normalized



- *Intuition*. One can view a linear operator on $\mathcal{H}_A$ as vectors in $\mathcal{H}_A \otimes \mathcal{H}_A^*$. Likewise, a superoperator on top of that is *isomorphic* to operators from $\mathcal{H}_A \otimes \mathcal{H}_A^*$ to $\mathcal{H}_B \otimes \mathcal{H}_B^*$.
- *Example*. For a completely dephasing channel $\Delta(\rho) \coloneqq \sum_i \langle i|\rho|i\rangle\, |i\rangle\langle i|$, the corresponding Choi operator is $J_{AB}^{\Delta} = \sum_i |i\rangle\langle i| \otimes |i\rangle\langle i|$.

  unnormalized maximally correlated state

Choi–Jamiołkowski isomorphism (2/2)

- Theorem (Choi–Jamiołkowski isomorphism). The following map is an *isomorphism*, $\mathcal{N}^{A\to B} \mapsto J_{AB}^{\mathcal{N}}$, with inverse given by $\mathcal{N}^{A\to B}(M_A) = \mathrm{Tr}_A\left[(M_A^T \otimes I_B)J_{AB}^{\mathcal{N}}\right]$, $\forall M_A: \mathcal{H}_A \to \mathcal{H}_A$ (where the transpose is taken in the same basis used in the definition of $J_{AB}^{\mathcal{N}}$).

  Proof. $\mathrm{Tr}_A\left[(M_A^T \otimes I_B)J_{AB}^{\mathcal{N}}\right] = \sum_{i,j} \mathrm{Tr}_A\left[(M_A^T \otimes I_B)|i\rangle\langle j| \otimes \mathcal{N}^{A\to B}(|i\rangle\langle j|)\right]$

  $= \sum_{i,j} \mathrm{Tr}_A\left[M_A^T|i\rangle\langle j| \otimes \mathcal{N}^{A\to B}(|i\rangle\langle j|)\right]$

  $= \sum_{i,j} \underbrace{\mathrm{Tr}\left[M_A^T|i\rangle\langle j|\right]}_{= \langle j|M_A^T|j\rangle = \langle i|M_A^T|j\rangle} \otimes \mathcal{N}^{A\to B}(|i\rangle\langle j|)$

  $= \sum_{i,j} \mathcal{N}^{A\to B}\left(|i\rangle\langle i|M_A^T|j\rangle\langle j|\right)$

  $= \mathcal{N}^{A\to B}(M_A)$. □

When is a superoperator completely positive? (1/2)

- Theorem. For a superoperator $\mathcal{N}^{A\to B}: \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}_B)$, the following are equivalent:

1. $\mathcal{N}^{A\to B}$ is *completely positive*, i.e. $(\mathcal{N}^{A\to B} \otimes \mathrm{Id}_R)(M_{AR}) \ge 0$, $\forall M_{AR} \ge 0$ and $\forall \mathcal{H}_R$.
2. $J_{AB}^{\mathcal{N}} \ge 0$, i.e. the *Choi operator* of $\mathcal{N}^{A\to B}$ is positive semi-definite.
3. *Kraus representation*: There exists operators $X_1, \dots, X_r: \mathcal{H}_A \to \mathcal{H}_B$ such that $\mathcal{N}^{A\to B}(M_A) = \sum_{i=1}^r X_i M_A X_i^\dagger$ for all linear operators M_A on $\mathcal{H}_A$.
4. *Stinespring representation*: There exists a $\mathcal{H}_E$ and an isometry $V: \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ such that $\mathcal{N}^{A\to B}(M_A) = \mathrm{Tr}_E\left[VM_AV^\dagger\right]$ for all linear operators M_A on $\mathcal{H}_A$.

Moreover, r in 3 and in 4 can be chosen as $\mathrm{rank}\left(J_{AB}^{\mathcal{N}}\right) \le \dim(\mathcal{H}_A)\dim(\mathcal{H}_B)$ (or larger).

Proof. The implication 1 ⇒ 2 is immediate, and 4 ⇒ 1 is easy to check .

We have proven 3 ⇔ 4. It remains to show 2 ⇒ 3 with $r = \mathrm{rank}\left(J_{AB}^{\mathcal{N}}\right)$.

When is a superoperator completely positive? (2/2)

Proof. Since $J_{AB}^{\mathcal{N}} \geq 0$, we use spectral decomposition to write $J_{AB}^{\mathcal{N}} = \sum_{\ell=1}^{r} |v_\ell\rangle\langle v_\ell|$.

Define Kraus operators $X_\ell := \sum_{i,j} \langle ij|v_\ell\rangle |j\rangle\langle i|$ on $\mathcal{H}_A$ to $\mathcal{H}_B$.

$$\begin{aligned}
\mathcal{N}^{A\to B}(M_A) &= \mathrm{Tr}_A\left[(M_A^T \otimes I_B) J_{AB}^{\mathcal{N}}\right] = \sum_\ell \mathrm{Tr}_A\left[(M_A^T \otimes I_B)|v_\ell\rangle\langle v_\ell|\right] \\
&= \sum_\ell \sum_{i,j} \sum_{i',j'} \mathrm{Tr}_A\left[(M_A^T \otimes I_B)|ij\rangle\langle ij||v_\ell\rangle\langle v_\ell||i'j'\rangle\langle i'j'|\right] \\
&= \sum_\ell \sum_{i,j} \sum_{i',j'} \langle ij|v_\ell\rangle\langle v_\ell|i'j'\rangle \mathrm{Tr}_A\left[(M_A^T \otimes I_B)|ij\rangle\langle i'j'|\right] \\
&= \sum_\ell \sum_{i,j} \sum_{i',j'} \langle ij|v_\ell\rangle\langle v_\ell|i'j'\rangle \underbrace{\mathrm{Tr}_A[M_A^T|i\rangle\langle i'|}_{=\langle i'|M_A^T|i\rangle = \langle i|M_A|i\rangle} \otimes |j\rangle\langle j'|] \\
&= \sum_\ell \sum_{i,j} \sum_{i',j'} \langle ij|v_\ell\rangle |j\rangle\langle i|M_A|i\rangle\langle j'|\langle v_\ell|i'j'\rangle \\
&= \sum_{i=1}^{r} X_i M_A X_i^\dagger
\end{aligned}$$

□

When is a superoperator trace-preserving?

- Theorem. For a completely positive superoperator $\mathcal{N}^{A\to B}$, the following are equivalent:

1. $\mathcal{N}$ is *trace-preserving*, i.e. $\mathrm{Tr}[\mathcal{N}(M)] = \mathrm{Tr}[M], \forall M_A$ (hence a quantum channel).
2. *Choi operator*: $\mathrm{Tr}_B\left[J_{AB}^{\mathcal{N}}\right] = I_A$.
3. *Kraus representation*: $\sum_{i=1}^{r} X_i^\dagger X_i = I_A$ for one/every Kraus operators.
4. *Stinespring representation*: $V^\dagger V = I_A$ (i.e. V is an isometry) for every dilation.

In fact, the equivalence between 1 and 2 holds for arbitrary superoperators (CP or not).

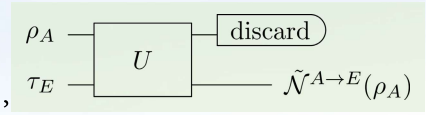
Proof. (i) Using the fact that $\mathrm{Tr}[XM] = \mathrm{Tr}[YM]$ for all $M \Leftrightarrow X = Y$. (ii) Cyclicity of Tr.

$$\mathrm{Tr}[\mathcal{N}^{A\to B}(M_A)] = \mathrm{Tr}\left[\mathrm{Tr}_A\left[(M_A^T \otimes I_B) J_{AB}^{\mathcal{N}}\right]\right] = \mathrm{Tr}\left[M_A^T \mathrm{Tr}_B\left[J_{AB}^{\mathcal{N}}\right]\right]$$

$$\mathrm{Tr}\left[\sum_i X_i M_A X_i^\dagger\right] = \mathrm{Tr}\left[M_A \sum_i X_i^\dagger X_i\right] \qquad \mathrm{Tr}\left[\mathrm{Tr}_E\left[V M_A V^\dagger\right]\right] = \mathrm{Tr}\left[M_A V^\dagger V\right]$$

□

Properties of Quantum Channels



- Suppose $(\mathcal{H}_E, U, \tau_E)$ is a dilation of a channel $\mathcal{N}^{A\to B}$, the *conjugate channel* $\tilde{\mathcal{N}}^{A\to E}$ is route to the environment defined as $\mathrm{Tr}_B \circ U \circ \tau_E$.
- The set of quantum channels is convex.
  (One can study the *extremal channels* and the properties.)
  In general, $\mathcal{N}_1 \circ \mathcal{N}_2 \neq \mathcal{N}_2 \circ \mathcal{N}_1$
- A channel $\mathcal{N}$ has inverse, i.e. there exists a channel $\bar{\mathcal{N}}$ such that $\bar{\mathcal{N}} \circ \mathcal{N} = \mathrm{Id}$, if and only if is a unitary channel.
  Sufficiency of a quantum channel $\mathcal{N}$
  Recovery map: for states ρ and σ,$\exists \mathcal{R}$ such that $\rho = \mathcal{R} \circ \mathcal{N}(\rho)$ and $\rho = \mathcal{R} \circ \mathcal{N}(\sigma)$.
- Quantum Markov chains: $\rho_{t+1} := \mathcal{N}(\rho_t) = \mathcal{N} \circ \mathcal{N}(\rho_{t-1}) = \cdots$.
  - Continuous time: quantum Markov semigroups.
- Tensor product: For channels $\mathcal{N}$ & $\mathcal{M}$ with Kraus operators $\{X_i\}_i$ and $\{Y_j\}_j$, the Kraus operators of $\mathcal{N} \otimes \mathcal{M}$ are $\{X_i \otimes Y_j\}_{i,j}$.

Other Examples of Quantum Channels (1/2)

- Unital channels: $\mathcal{N}(I) = I$. (Classically, $P_{Y|X}$ is *doubly stochastic*.)
  I, X, Y, Z
- Qubit Pauli channels: $\mathcal{N}(\rho) = \sum_i q_i \sigma_i \rho \sigma_i$ for a prob. distribution q and Pauli's σ_i.
  (The set of all unital channels is unitarily equivalent to the set of Pauli channels.)
  - Depolarizing channels: $q_1 = q_2 = q_3 = p/4$; then $\mathcal{N}(\rho) = (1-p)\rho + p\frac{I}{2}$. [§8.3, N&C]
    (The Bloch sphere is symmetrically shrunk to a ball of radius $1-p$.
  - Phase-damping channels: $q_0 = 1-p$, $q_3 = p$; then $\mathcal{N}(\rho) = (1-p)\rho + pZ$.
    → Its power series $\mathcal{N}, \mathcal{N}^2, \cdots$, converges to the channel $\mathrm{diag}_{\{|i\rangle\}_i}(\rho) = \sum_i \langle i|\rho|i\rangle |i\rangle\langle i|$.
    (The z-th component the Bloch vector is preserved but not others.)
- Amplitude damping: $X_0 = |0\rangle\langle 0| + \sqrt{1-\gamma}|1\rangle\langle 1|, X_1 = \sqrt{\gamma}|0\rangle\langle 1|$,

$$\mathcal{N}(\rho) = \begin{pmatrix} \rho_{00} + \gamma\rho_{11} & \sqrt{1-\gamma}\rho_{01} \\ \sqrt{1-\gamma}\rho_{10} & (1-\gamma)\rho_{11} \end{pmatrix}.$$ The Bloch vector damps to the fixed point $|0\rangle\langle 0|$.

Other Examples of Quantum Channels (2/2)

- Random unitary channels: $\mathcal{N}(\rho) = \sum_i q_i U_i \rho U_i^\dagger$ → flips a coin and chose U_i.

  Average unitary (*Twirling*): $\mathcal{N}(M) = \int_{\mathcal{U}(\mathcal{H})} UMU^\dagger \mathrm{d}U = \frac{\mathrm{Tr}[M]}{d} I$, for all M.

  (An example: $\mathcal{N}(\rho) = \frac{1}{4}\sum_i \sigma_i \rho \sigma_i = \frac{I}{2}$ in random Pauli channels.)
- POVM measurement (quantum-classical): $\mathcal{N}(\rho) = \mathrm{Tr}_\omega[\rho \Pi^\omega] \otimes |\omega\rangle\langle\omega|$.
- Classical-quantum channels: $\mathcal{N}(\sigma) = \sum_x \langle x|\sigma|x\rangle \rho_x$ for diagonal σ.
- Entanglement-breaking channels: $(\mathcal{N} \otimes I)(\rho)$ is always separable for all ρ.
  Any entanglement-breaking channel can be written as *measurement and prepare*.
- Degradable channels: $\mathcal{N}$ and its conjugate channel $\widetilde{\mathcal{N}}$, there exists a channel $\mathcal{M}$ such that $\mathcal{M} \circ \mathcal{N} = \widetilde{\mathcal{N}}$ (which is important in the channel capacity problems).

Concluding Remarks

Concluding Remarks (1/3)

- You need a dictionary:
- Embedding vectors to diagonal matrices.
- Classical information science for manipulating probability distributions naturally extend to that of density operators → QIS or QIP.

Classical	Quantum
Probability vector	Density operator
Sum	Trace
Marginal distribution	Reduced state
Deterministic	Pure
Probabilistic	Convex mixture
$\mathbb{E}_X[P_{\Omega\|X}(\omega\|X)]$	$\mathrm{Tr}[\rho_X \Pi^\omega]$

- A statistical framework
  - Preparation: A preparation procedure determines the *state* of a system.
  - Evolution: How is a quantum state *evolving*?
  - Measurement: A measurement procedure produces some observation outcomes.

Preparation → Evolution → Measurement → Data

Concluding Remarks (2/3)

- Allowing *off-diagonal phases* yields the possibility of *superposition*, which is a property of quantum states of a single system that exhibits 'quantumness'.
- Allowing *entangled* vectors or operators yields *entanglement*, which is a property of quantum states of a composite quantum system that exhibits 'quantumness'.
- *Operational meaning* of separable states: They are the largest class of bipartite states that can be prepared by Alice and Bob in their own labs by using local quantum operations and classical communications (LOCC).
- We only talked about pure entangled states $|\psi\rangle_{AB}$. How about other entangled states?

  Mixed & multipartite entangled states are less well understood.
- Why mixed states? (i) QM + randomness → QIS; (ii) Reduced states could be mixed.
- Why POVM? (i) randomized decisions; (ii) with ancillas.

Concluding Remarks (3/3)

- Any quantum channel $\mathcal{N}^{A\to B}$ as introduced above can be realized *physically* (by Stinespring's dilation). That is, in principle, there exists a device that outputs the state $\rho_B = \mathcal{N}^{A\to B}(\rho_A)$ given an arbitrary state ρ_A as input.
- Why noisy quantum channels (instead of just unitary evolution as before)?
  (i) hidden variables as in classical probabilistic transitions (via Kraus' representation);
  (ii) coupled to environment (e.g. energy leakage) and partial trace.
- For unitary channels the environment is dynamically independent of the system. (This can be seen from the conjugate channel of U, i.e. it is a constant channel τ_E.)
- Both the Stinespring and Kraus representations are not unique.
- The concepts of classical channels ($P_{Y|X}: P_X \mapsto P_Y$) naturally extends to quantum channels. Hence, what we study for the classical ones (e.g. noisy channel coding) extends the quantum ones – the central topic in quantum information processing.

Quantum Information and Computation
Distance Measures and Entropies

Hao-Chung Cheng (鄭皓中)
haochung@ntu.edu.tw

Department of Electrical Engineering
National Taiwan University

May 5, 2021

Motivation

Motivation

- How close are two quantum states?

  When states are both pure, we can calculate their overlap (e.g. via the Swap Test).

  Since mixed states are operators, we are essentially manipulating with operators.
- In classical probability/statistics, there are numerous distance measures on probability distributions. How to definite the associated noncommutative quantities?
- Is there any *operational interpretation* for such quantities?
- How *uncertain* a quantum state is? → Entropies for quantum systems.
- In this lecture, we will introduce various distance measures, divergences, and entropies for quantum systems, and collect most of their properties.
- Welcome to the *Quantum Entropy Zoo*!

Trace Distance and Fidelity

Norms of Operators/Matrices

- Definition (Matrix functional). For any Hermitian matrix with eigenvalue decomposition $M = \sum_i \lambda_i |v_i\rangle\langle v_i|$ and any real-valued function $f: \mathbb{R} \to \mathbb{R}$, the associated matrix function is defined by $f(M) \coloneqq \sum_i f(\lambda_i)|v_i\rangle\langle v_i|$.
- Definition (Schatten p-norm). For a Hermitian operator $M = \sum_i \lambda_i |v_i\rangle\langle v_i|$, we define $\|M\|_p \coloneqq (\mathrm{Tr}[|M|^p])^{1/p} \equiv (\sum_i |\lambda_i|^p)^{1/p} =: \|\boldsymbol{\lambda}\|_p$ for all $p \in [1, \infty]$.
  - $\|M\|_1 \coloneqq \mathrm{Tr}[|M|]$ is called the *trace norm* (or nuclear norm).
  - $\|M\|_2 \coloneqq \langle M, M\rangle^{1/2}$ is the *Hilbert–Schmidt norm* (or Frobenius norm).
  - $\|M\|_\infty \coloneqq \max_i \lambda_i = \max_{\||\psi\rangle\|_2=1} \|M|\psi\rangle\|_2$ is the operator norm (or spectral norm).
- For general (non-Hermitian) operators, the Schatten p-norm is the ℓ_p-norm of the singular values, and $|M| \coloneqq \sqrt{M^\dagger M}$.
- Other norms: e.g. *Ky Fan k-norm* is the sum of the k largest singular values.

Properties of Norms

- Invariant under adjoint, conjugation, and transposition (w.r.t. any ONB).
- Unitary-invariance: $\|M\|_p = \left\|UMU^\dagger\right\|_p$ for any unitary U.
- Monotonically decreasing: $\|M\|_1 \geq \|M\|_2 \geq \|M\|_\infty$. Reversely, there will be dimension factors: $\|M\|_1 \leq \sqrt{d}\|M\|_2$ and $\|M\|_p \leq d^{1/p-1/q}\|M\|_q$ for $q \geq p$.
- Triangle (Minkowski) inequality: $\|M + N\|_p \leq \|M\|_p + \|N\|_p$.
- Convexity: $\|(1-t)M + tN\|_p \leq (1-t)\|M\|_p + t\|N\|_p$ for all $t \in [0,1]$.
- Holder's inequality: $|\langle M, N\rangle| \leq \|M\|_p \|M\|_q$ for $\frac{1}{p} + \frac{1}{q} = 1$ and $p \in [1, \infty]$.
- Duality: $\|M\|_p = \max_{M:\|M\|_q=1} |\langle M, N\rangle|$ for $\frac{1}{p} + \frac{1}{q} = 1$.
- Submultiplicativity: $\|MN\|_p \leq \|M\|_p \|N\|_q$.
- Additivity under direct sum: $\|M \oplus N\|_p = \|M\|_p + \|N\|_p$.

Trace Distance

- Definition (Trace distance). The (normalized) trace distance between two states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ is defined as $T(\rho, \sigma) \coloneqq \frac{1}{2}\|\rho - \sigma\|_1$.
- Properties:
  - $T(\rho, \sigma) \in [0,1]$: $T(\rho, \sigma) = 0 \Leftrightarrow \rho = \sigma$ and $T(\rho, \sigma) = 1 \Leftrightarrow \rho \perp \sigma$.
  - For pure $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\phi\rangle\langle\phi|$, $T(\rho, \sigma) = \sqrt{1 - |\langle\phi|\psi\rangle|^2}$.
  - When $[\rho, \sigma] = 0$ (classical), it is the *total variation distance* $\sup_{A \in \mathcal{F}} |P(A) - Q(A)|$.
  - Isometry/unitary-invariance: $T\left(U\rho U^\dagger, U\sigma U^\dagger\right) = T(\rho, \sigma)$ for any unitary U.
  - Invariant under ancilla: $T(\rho, \sigma) = T(\rho \otimes \tau, \sigma \otimes \tau)$ for all ancilla state τ.
  - Convexity: $T(\sum_i p_i \rho_i, \sum_i p_i \sigma_i) \leq \sum_i p_i T(\rho_i, \sigma_i)$ for all probability distribution p.
  - Contractive under CPTP: $T\left(\mathcal{N}(\rho), \mathcal{N}(\sigma)\right) \leq T(\rho, \sigma)$.

  $(T(\rho_A, \sigma_A) \leq T(\rho_{AB}, \sigma_{AB})$ under partial trace.)

What is the operational meaning?

Holevo–Helström Theorem

- Theorem: Given two states ρ_0 and ρ_1 with prior probability p_0 and p_1, the maximal success probability $P_s^\star := \max_{0\le\Pi\le I}\{p_0\mathrm{Tr}[(I-\Pi)\rho_0] + p_1\mathrm{Tr}[\Pi\rho_1]\}$ is given by $P_s^\star = \frac{1}{2} + T(p_0\rho_0, p_1\rho_1)$. ←This gives the trace distance an *operational interpretation*.
- Remarks:
  - Perfect discrimination if and only if $\rho_0 \perp \rho_1$ (as the classical case).
  - For $p_0 = p_1 = 1/2$, $\rho_0 = |\psi_0\rangle\langle\psi_0|$ and $\rho_1 = |\psi_1\rangle\langle\psi_1|$, $P_s^\star = \frac{1}{2} + \frac{1}{2}\sqrt{1-|\langle\psi_0|\psi_1\rangle|^2}$.
  - The duality $T(p_0\rho_0, p_1\rho_1) = \max_{0\le\Pi\le I}\mathrm{Tr}\,[\Pi(p_0\rho_0 - p_1\rho_1)]$ already suggests that the optimal measurement is $\Pi^\star = \{p_0\rho_0 - p_1\rho_1\}$.
  - How about $P_s^\star(n) = \frac{1}{2} + T\left(p_0\rho_0^{\otimes n}, p_1\rho_1^{\otimes n}\right)$ ← Asymptotic binary hypothesis testing.

    $\Rightarrow P_e^\star(n) := 1 - P_s^\star(n) \simeq e^{-nC}$, where $C := \max_{\alpha\in[0,1]}\mathrm{Tr}\,[\rho_0^\alpha\rho_1^{1-\alpha}]$.

    [Audenaert *et al.*' 18]

    Chernoff bound

Proof of Holevo–Helström Theorem

1. Achievability (Sufficiency).

$$\begin{aligned} \mathrm{P_s} &= p_0\mathrm{Tr}\,[\rho_0\Pi_0] + p_1\mathrm{Tr}\,[\rho_1\Pi_1] \\ &= \tfrac{1}{2}\mathrm{Tr}\,[(p_0\rho_0 + p_1\rho_1)(\Pi_0+\Pi_1)] + \tfrac{1}{2}\mathrm{Tr}\,[(p_0\rho_0 - p_1\rho_1)(\Pi_0-\Pi_1)] \\ &= \tfrac{1}{2} + \tfrac{1}{2}\mathrm{Tr}\,[(p_0\rho_0 - p_1\rho_1)(\Pi_0-\Pi_1)] \end{aligned}$$

   Let $A := p_0\rho_0 - p_1\rho_1$ and choose $\Pi_0 = \{p_0\rho_0 > p_1\rho_1\}$.

   $\Rightarrow P_s = \frac{1}{2} + \frac{1}{2}\mathrm{Tr}\,[A\{A>0\} - A\{A\le 0\}] = \frac{1}{2} + \frac{1}{2}\mathrm{Tr}\,[|A|]$

2. Optimality (Necessity).

   Hölder's inequality: $|\langle A, B\rangle| \le \|A\|_p\|B\|_q, \quad \frac{1}{p}+\frac{1}{q} = 1$

   $\Rightarrow \mathrm{Tr}\,[A(\Pi_0-\Pi_1)] \le \|A\|_1\underbrace{\|\Pi_0-\Pi_1\|_\infty}_{\le 1} \le \|A\|_1$ □

Quantum Fidelity (1/2)

- Definition (Fidelity). Given states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, we define their fidelity by

  $F(\rho,\sigma) := \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2 = \left(\mathrm{Tr}\left[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right]\right)^2$.

  It is a similarity or "overlap" rather than distance

  (Some books and literature used definitions without the square.)
- Properties:
  - $F(\rho,\sigma) \in [0,1]$: $F(\rho,\sigma) = 0 \Leftrightarrow \rho = \sigma$ and $F(\rho,\sigma) = 1 \Leftrightarrow \rho \perp \sigma$.
  - If $\sigma = |\phi\rangle\langle\phi|$, then $F(\rho,\sigma) = \langle\phi|\rho|\phi\rangle$; if further $\rho = |\psi\rangle\langle\psi|$, $F(\rho,\sigma) = |\langle\phi|\psi\rangle|^2$.
  - When $[\rho,\sigma] = 0$ (classical), it is the Bhattacharyya coefficient $\left(\sum_x\sqrt{P(x)Q(x)}\right)^2$.
  - Isometry/unitary-invariance: $F\left(U\rho U^\dagger, U\sigma U^\dagger\right) = F(\rho,\sigma)$ for any unitary U.
  - Invariant under ancilla: $F(\rho,\sigma) = F(\rho\otimes\tau, \sigma\otimes\tau)$ for all ancilla state τ.
  - Concavity: $F(\sum_i p_i\rho_i, \sum_i p_i\sigma_i) \ge \sum_i p_i F(\rho_i,\sigma_i)$ for all probability distribution p.
  - Never-decrease under CPTP: $F\left(\mathcal{N}(\rho), \mathcal{N}(\sigma)\right) \ge F(\rho,\sigma)$.

Quantum Fidelity (2/2)

- Properties:
  - Symmetry: $F(\rho,\sigma) = F(\sigma,\rho)$.
  - Multiplicativity under tensorization: $F\left(\rho^{\otimes n}, \sigma^{\otimes n}\right) = n\,F(\rho,\sigma)$.
  - Relations: $1 - \sqrt{F(\rho,\sigma)} \le T(\rho,\sigma) \le \sqrt{1-F(\rho,\sigma)}$.
  - (Operational meaning) Equivalence to choosing the best measurement:

    $F(\rho,\sigma) = \min_{\mathrm{POVM}\,\{\Pi^x\}_x} F(P_X, Q_X) = \min_{\mathrm{POVM}\,\{\Pi^x\}_x}\left(\sum_x\sqrt{\mathrm{Tr}[\rho\Pi^x]\mathrm{Tr}[\sigma\Pi^x]}\right)^2$.
  - Definition (*Purified distance*). $P(\rho,\sigma) := \sqrt{1-F(\rho,\sigma)}$.
- Theorem (Uhlmann). Let $\rho_A, \sigma_A \in \mathcal{S}(\mathcal{H}_A)$ and $\mathcal{H}_B$ a Hilbert space such that both states admit *purifications* on $\mathcal{H}_A\otimes\mathcal{H}_B$. Then,

  $F(\rho_A,\sigma_A) = \max\{|\langle\psi|\phi\rangle_{AB}|^2 : |\psi\rangle_{AB}, |\phi\rangle_{AB} \in \mathcal{H}_A\otimes\mathcal{H}_B$ purifications of $\rho_A, \sigma_A\}$.

  (It is the maximal overlap between any pair of purifications.)

Proof of Uhlmann's Theorem (For $\mathcal{H}_A = \mathcal{H}_B$)

1. From the Schmidt decomposition, we choose two *standard purifications*:
   $|\psi\rangle_{AB} \coloneqq \left(\sqrt{\rho_A} \otimes I_B\right) \sum_i |i\rangle \otimes |i\rangle$, $|\phi\rangle_{AB} \coloneqq \left(\sqrt{\sigma_A} \otimes I_B\right) \sum_i |i\rangle \otimes |i\rangle$.
2. Simply fix $|\psi\rangle_{AB}$ and vary $|\phi\rangle_{AB}$. Since all $\{|\phi\rangle_{AB}\}$ are related by $I_A \otimes U_B$, $\therefore$

$$\begin{aligned}
|\langle\psi|(I_A \otimes U_B)|\phi\rangle| &= \sum_{i,j} (\langle i| \otimes \langle i|)\left(\sqrt{\rho_A}\sqrt{\sigma_A} \otimes U_B\right)(|j\rangle \otimes |j\rangle) \\
&= \sum_{i,j} \langle i|\sqrt{\rho_A}\sqrt{\sigma_A}|j\rangle\langle i|U_B|j\rangle \\
&= \sum_{i,j} \langle i|\sqrt{\rho_A}\sqrt{\sigma_A}|j\rangle\langle j|U_B^T|i\rangle \\
&= \sum_i \langle i|\sqrt{\rho_A}\sqrt{\sigma_A}U_A^T|i\rangle \\
&= \mathrm{Tr}[\sqrt{\rho_A}\sqrt{\sigma_A}U_A^T]
\end{aligned}$$

Duality of $\|\cdot\|_1$

3. Maximizing over U_A: $\max_{U_A} \mathrm{Tr}[\sqrt{\rho_A}\sqrt{\sigma_A}U_A^T] = \max_{U_A} \mathrm{Tr}[\sqrt{\rho_A}\sqrt{\sigma_A}U_A] = F(\rho_A, \rho_B)^{1/2}$. □

References

- Matrix Analysis
  - F. Hiai, D. Petz, *Introduction to Matrix Analysis and Applications,* Springer Publisher, 2014.
  - R. Bhatia, *Matrix Analysis*, Springer Publisher, 1997.
  - N. Higham, *Functions of Matrices: Theory and Computation*, Society for Industrial and Applied Math. 2008.
- Quantum Hypothesis Testing
  - K. M. R. Audenaert, M. Nussbaum, A. Szkoła, F. Verstraete, "Asymptotic Error Rates in Quantum Hypothesis Testing," *Communications in Mathematical Physics*, 279, 251–283, 2008.
  - M. Nussbaum, A. Szkoła, "The Chernoff lower bound for symmetric quantum hypothesis testing," *The Annals of Statistics*, 37(2), 2009.
- Books on Quantum Information Theory
  - M. Tomamichel, *Quantum Information Processing with Finite Resources*, Springer Publisher, 2015.
  - M. Wilde, *Quantum Information Theory* (2nd Edition), Cambridge University Press, 2017.
  - J. Watrous, *The Theory of Quantum Information*, Cambridge University Press, 2019.

Quantum Information and Computation
Quantum Data Compression

Hao-Chung Cheng (鄭皓中)
haochung@ntu.edu.tw

Department of Electrical Engineering
National Taiwan University

May 12, 2021

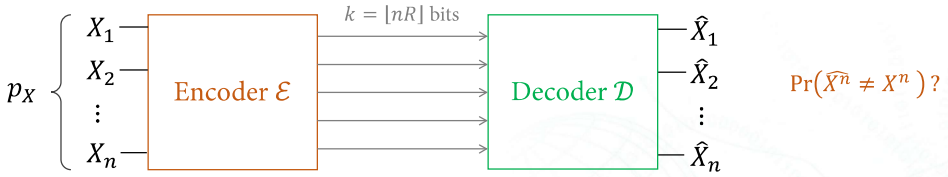
Classical Data Compression

Classical Data Compression (1/2)

Independently and identically distributed (IID)

- A *discrete memoryless source* (DMS) means that the source (modeled by a probability distribution $p_X \in \mathcal{P}(\mathcal{X})$) emits a sequence of symbols $x^n := x_1 x_2 \cdots x_n \in \mathcal{X}^n$ with probability $p_{X^n}(x^n) = P_{X_1}(x_1) P_{X_2}(x_2) \cdots P_{X_n}(x_n)$.

Assume p_X is known



- The goal of (lossless) source coding is to compress a sequence of n-symbols X^n to a k-bit string (with *compression rate* $R := k/n$), and then later recover it with vanishing *probability of error*, say $\Pr(\hat{X}^n \neq X^n) \geq 1 - \varepsilon$.
(For example, the Huffman code is an *optimal* lossless code.)

Classical Data Compression (2/2)

- Definition (Source code). An (n, R, ε)-code for probability distribution $p_X \in \mathcal{P}(\mathcal{X})$ is a pair of functions $\mathcal{E}: \mathcal{X}^n \to \{0,1\}^{\lfloor nR \rfloor}$ and $\mathcal{D}: \{0,1\}^{\lfloor nR \rfloor} \to \mathcal{X}^n$ such that $\Pr(\mathcal{D} \circ \mathcal{E}(X^n) \neq X^n) = \sum_{x^n \in \mathcal{X}^n} p_{X^n}(x^n) \mathbf{1}_{\mathcal{D} \circ \mathcal{E}(x^n) \neq x^n}$, for $X^n \sim p_X$ IID.

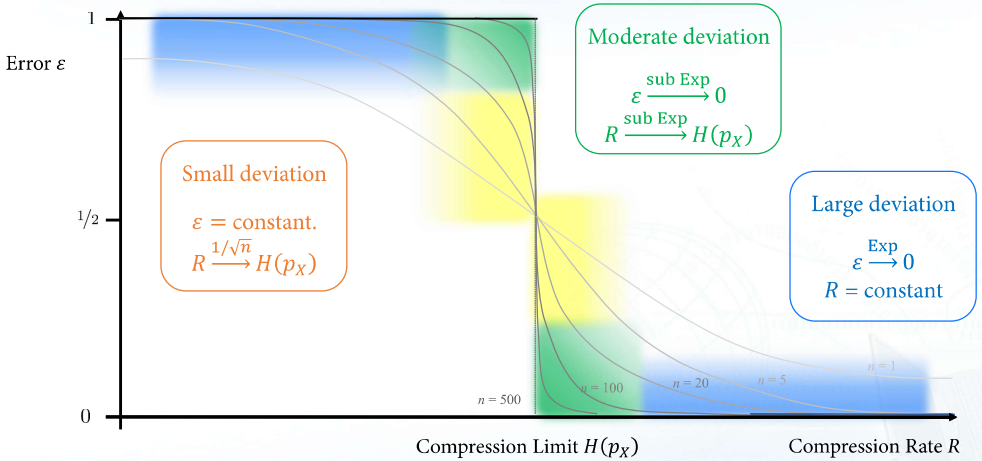
as an Average Fidelity

Shannon's (Noiseless) Source Coding Theorem

Let $p_X \in \mathcal{P}(\mathcal{X})$ and $\varepsilon \in (0,1)$. Then:

- (Achievability) If $R > H(p_X)$, $\exists N_0 \in \mathbb{N}$ such that $\exists$ an (n, R, ε)-code for all $n \geq N_0$.
- (Optimality) If $R < H(p_X)$, $\exists N_0 \in \mathbb{N}$ such that no (n, R, ε)-codes exist for all $n \geq N_0$.

Characterization



Typical Sequences

- Definition (Typical set). For $p_X \in \mathcal{P}(\mathcal{X})$, $n \in \mathbb{N}$, and $\delta > 0$, define the *typical set*:
$$\mathcal{T}_{n,\delta}(p_X) := \left\{x^n \in \mathcal{X}^n : \left| p_X(x_1) \cdots p_X(x_n) - 2^{-nH(p_X)} \right| \leq \delta \right\}.$$

Asymptotic Equipartition Property (AEP)

Typical Sequence Theorem

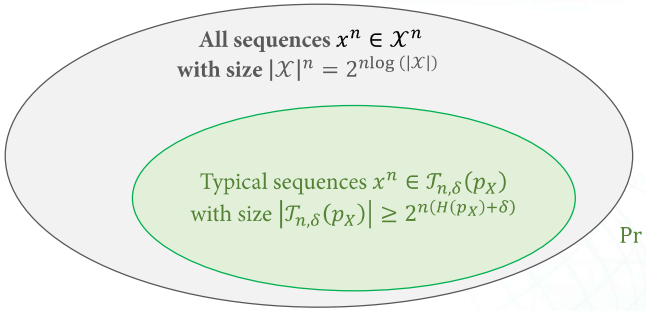
- $2^{-n(H(p_X)+\delta)} \leq p_{X^n}(x^n) \leq 2^{-n(H(p_X)-\delta)}$ for all $x^n \in \mathcal{T}_{n,\delta}(p_X)$.
- $\left|\mathcal{T}_{n,\delta}(p_X)\right| \leq 2^{n(H(p_X)+\delta)}$.
- $\Pr\left(X^n \in \mathcal{T}_{n,\delta}(p_X)\right) \geq 1 - \frac{\sigma^2}{n\delta^2}$, where $\sigma^2 = \mathrm{Var}\left(\log \frac{1}{p_X(X)}\right)$ only depends on p_X.

Interpretation

In the asymptotic limit

- Not all sequences $x^n \in \mathcal{X}^n$ are happening equally likely, but for large n, the number of occurrences of any symbol x appears in $\mathcal{X}^n$ is $\approx n(p_X(x) \pm \delta)$.

Large deviation regime



$$\Pr(X^n \in \mathcal{X}^n) = 1$$

$$\Pr\left(X^n \in \mathcal{T}_{n,\delta}(p_X)\right) \geq 1 - \frac{\sigma^2}{n\delta^2}$$

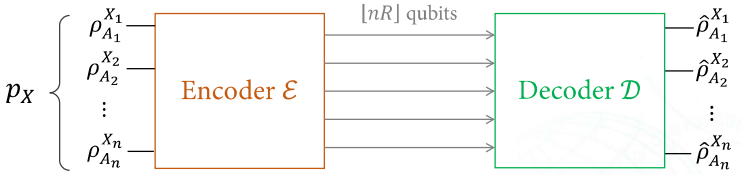
- Source Coding Theorem: We only encode the *typical sequences* with rate $\approx H(p_X)$. *Atypical sequences* incur errors but they are only rare events (arbitrarily small prob.).

Quantum Data Compression

Quantum Data Compression (1/3)

- A *quantum source* emits the states ρ_A^x with probability $p_X(x)$. Further, we consider the *IID* (*memoryless*) source, which emits states $\rho_{A^n}^{x^n} := \rho_{A_1}^{x_1} \otimes \rho_{A_2}^{x_2} \otimes \cdots \otimes \rho_{A_n}^{x_n}$ with probability $p_{X^n}(x^n) = P_{X_1}(x_1)P_{X_2}(x_2)\cdots P_{X_n}(x_n)$.

Again assume p_X is known

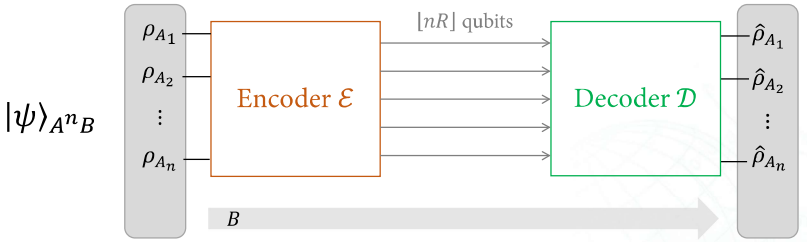


$\bar{F}$ or F_c?

- The goal of (lossless) quantum data compression is to compress the system $\mathcal{H}_A^{\otimes n}$ to $\lfloor nR \rfloor$-qubits (with *compression rate* R), and then later recover it with high fidelity, e.g. $F_c\left(\mathcal{D} \circ \mathcal{E}, \rho_A^{\otimes n}\right) \geq 1 - \varepsilon$ or $\bar{F} = \sum_{x^n} p_{X^n}(x^n) F\left(\rho_{A^n}^{x^n}, \mathcal{D} \circ \mathcal{E}\left(\rho_{A^n}^{x^n}\right)\right) \geq 1 - \varepsilon$.

Quantum Data Compression (2/3)

- Why IID? Let $\rho_A := \sum_x p_X(x)\rho_A^x$. Then, $\sum_{x^n} p_{X^n}(x^n)\rho_{A^n}^{x^n} = \sum_{x_1\cdots x_n} p_{X_1}(x_1)\cdots p_{X_1}(x_1)\rho_{A_1}^{x_1} \otimes \cdots \otimes \rho_{A_1}^{x_1} = \rho_{A_1} \otimes \rho_{A_2} \otimes \cdots \otimes \rho_{A_n}$.



$$(\mathcal{D} \circ \mathcal{E} \otimes \mathrm{Id}_B)(|\psi\rangle\langle\psi|_{A^nB}) \approx |\psi\rangle\langle\psi|_{A^nB}$$

- Consider a purification $|\psi\rangle_{A^nB}$ of $\rho_A^{\otimes n}$. We want to design quantum operations $\mathcal{E}$ and $\mathcal{D}$ such that $F\left(|\psi\rangle\langle\psi|, (\mathcal{D} \circ \mathcal{E} \otimes \mathrm{Id})(|\psi\rangle\langle\psi|)\right) = F_c\left(\mathcal{D} \circ \mathcal{E}, \rho_A^{\otimes n}\right) \geq 1 - \varepsilon$.

Quantum Data Compression (3/3)

- Definition (Quantum code). An (n, R, ε)-quantum code for $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ is a pair of operations $\mathcal{E}$ from $\mathcal{H}_A^{\otimes n}$ to $\left(\mathbb{C}^2\right)^{\otimes \lfloor nR \rfloor}$ and $\mathcal{D}$ from $\left(\mathbb{C}^2\right)^{\otimes \lfloor nR \rfloor}$ to $\mathcal{H}_A^{\otimes n}$ such that $F_{\mathrm{c}}\left(\mathcal{D} \circ \mathcal{E}, \rho_A^{\otimes n}\right) \geq 1 - \varepsilon$.

Schumacher Compression [Schmacher'95]

Let $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ and $\varepsilon \in (0,1)$. Then:

- (Achievability) If $R > H(\rho_A)$, $\exists N_0 \in \mathbb{N}$ such that $\exists$ an (n, R, ε)-qcode for all $n \geq N_0$.
- (Optimality) If $R < H(\rho_A)$, $\exists N_0 \in \mathbb{N}$ such that no (n, R, ε)-qcodes exist for all $n \geq N_0$.

The von Neumann entropy $H(\rho_A)$ serves as a fundamental limit for quantum data compression.

Quantum Typical Projection

- Definition (Typical projection). For $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ with eigen-decomposition $\rho_A = \sum_{y=1}^{\dim \mathcal{H}_A} q(y)|e_y\rangle\langle e_y|$, $n \in \mathbb{N}$, and $\delta > 0$, define the *typical projection*:
$$\Pi_{n,\delta} \coloneqq \sum_{y^n \in \mathcal{T}_{n,\delta}(q)} |e_{y_1}\rangle\langle e_{y_1}| \otimes \cdots \otimes |e_{y_n}\rangle\langle e_{y_n}|.$$

$\Pi_{n,\delta}$ commutes with $\rho_A^{\otimes n}$

Quantum Asymptotic Equipartition Property (AEP)

Quantum Typical Projection Theorem

- The nonzero eigenvalues of $\Pi_{n,\delta}\rho_A^{\otimes n}\Pi_{n,\delta}$ are within $2^{-n(H(\rho_A)\pm\delta)}$.
- $\operatorname{rank}\Pi_{n,\delta} = \left|\mathcal{T}_{n,\delta}(q)\right| \leq 2^{n(H(\rho_A)+\delta)}$.
- $\operatorname{Tr}\left[\rho_A^{\otimes n}\Pi_{n,\delta}\right] \geq 1 - \frac{\sigma^2}{n\delta^2}$, where σ^2 only depends on eigenvalues of ρ_A.

Proof (Achievability)

- High-level idea: We project the state in system A^n into the typical subspace (called the coding space): $|\psi\rangle_{A^nB} \mapsto |\psi_{\|}\rangle_{A^nB} + |\psi_{\perp}\rangle_{A^nB}$, and this happens with high prob.

For sufficiently large $n \geq \delta^{-1}$

1. Let $\delta = \frac{R-H(q)}{2} = \frac{R-H(\rho_A)}{2}$. Then, $\operatorname{rank}\Pi_{n,\delta} \leq 2^{n(H(\rho_A)+\delta)} = 2^{n(R-\delta)} \leq 2^{\lfloor nR \rfloor}$.
2. Construct isometry $V: \mathcal{H}_A^{\otimes n} \to \left(\mathbb{C}^2\right)^{\otimes \lfloor nR \rfloor}$ such that $V^\dagger V = \Pi_{n,\delta}$.
   E.g. $V \coloneqq \sum_{i=1}^{\operatorname{rank}\Pi_{n,\delta}} |\psi_i\rangle\langle\varphi_i|$, where $\{|\varphi_i\rangle\}_i$ is the basis of the typical subspace.
3. Codes: $\mathcal{E}(M) \coloneqq VMV^\dagger + \operatorname{Tr}\left[\sqrt{I - \Pi_{n,\delta}}M\sqrt{I - \Pi_{n,\delta}}\right]\alpha$,
   $\mathcal{D}(M) \coloneqq V^\dagger MV + \operatorname{Tr}\left[\sqrt{I - VV^\dagger}M\sqrt{I - VV^\dagger}\right]\beta$.
   $\uparrow \{|\psi_i\rangle\}_i$ is an ONB of $\left(\mathbb{C}^2\right)^{\otimes \lfloor nR \rfloor}$
   $\leftarrow \alpha$ & β don't matter
4. $\mathcal{E}$ has Kraus operator V and $\mathcal{D}$ has Kraus operator $V^\dagger$, $\Rightarrow \mathcal{D} \circ \mathcal{E}$ has $V^\dagger V = \Pi_{n,\delta}$.
   $\Rightarrow$ By Lemma (Channel Fidelity), $F\left(\mathcal{D} \circ \mathcal{E}, \rho_A^{\otimes n}\right) \geq \left(\operatorname{Tr}[\rho_A^{\otimes n}\Pi_{n,\delta}]\right)^2 \geq 1 - \varepsilon$. □

Proof (Optimality)

- High-level idea: If we project (say using P) to too small subspace (say with rank $P \leq 2^{nR}$) then the probability of "triggering" such event vanishes.

1. $$\begin{aligned}\operatorname{Tr}\left[\rho_A^{\otimes n}P\right] &= \operatorname{Tr}\left[\rho_A^{\otimes n}\Pi_{n,\delta}P\right] + \operatorname{Tr}\left[\rho_A^{\otimes n}(I - \Pi_{n,\delta})P\right] \\ &\leq \underbrace{\left\|\rho_A^{\otimes n}\Pi_{n,\delta}\right\|_\infty}_{\leq 2^{-n(H(\rho_A)-\delta)}} \underbrace{\|P\|_1}_{\leq 2^{nR}} + \underbrace{\operatorname{Tr}\left[\rho_A^{\otimes n}(I - \Pi_{n,\delta})P\right]}_{= 1 - \operatorname{Tr}\left[\rho_A^{\otimes n}\Pi_{n,\delta}\right]} \\ &\leq 2^{-n\delta} + \left(1 - \operatorname{Tr}\left[\rho_A^{\otimes n}\Pi_{n,\delta}\right]\right) \to 0.\end{aligned}$$
   Hölder's inequality
   Choose $\delta = (H(\rho_A) - R)/2$
2. Let $\mathcal{D} \circ \mathcal{E}$ has Kraus operators $\{Z_k\}_k$ with each rank $\leq 2^{nR}$.
   $\Rightarrow F\left(\mathcal{D} \circ \mathcal{E}, \rho_A^{\otimes n}\right) = \sum_k \left|\operatorname{Tr}[\rho_A^{\otimes n}Z_k]\right|^2 = \sum_k \left|\operatorname{Tr}[\rho_A^{\otimes n}Z_k\{Z_k \geq 0\}]\right|^2$
   $\leq \sum_k \operatorname{Tr}\left[Z_k^\dagger Z_k \rho_A^{\otimes n}\right] \operatorname{Tr}\left[\rho_A^{\otimes n}\{Z_k \geq 0\}\right] \leq 2^{-n\delta} + \left(1 - \operatorname{Tr}\left[\rho_A^{\otimes n}\Pi_{n,\delta}\right]\right)$ □

Concluding Remarks

- We studied the trade-off between (i) compression rate R, (ii) block-length n, and (iii) performance – Fidelity F_c. Hence, Schumacher's compression shows that $H(\rho_A)$ is a fundamental limit on R, for which $F_c \to 0$ or 1 (in the asymptotic limit).
  - Fix $F_c \geq 1 - \varepsilon$, study R: second-order analysis.
  - Fix $R > H(\rho_A)$, study how fast $F_c \to 1$: error-exponent analysis (still open).
- Various protocols:
  - Quantum data compression with quantum side information (Slepian–Wolf coding).
  - Classical data compression with quantum side information.
  - Quantum data compression with classical side information.
  - Distributed quantum data compression (Wyner–Ahlswede–Körner coding).
  - Quantum *lossy* data compression (rate-distortion function).

References (1/3)

- Original Quantum Data Compression
  - B. Schumacher, "Quantum coding," *Physical Review A*, 51:2738–2747, 1995.
  - R. Jozsa and B. Schumacher, "A new proof of the quantum noiseless coding theorem," *Journal of Modern Optics*, 41:2343–2349, 1994.
  - M. A. Nielsen. *Quantum Information Theory*. Ph.D. Thesis, University of New Mexico, 1998.
  - H. Barnum, C. A. Fuchs, R. Jozsa, and B. Schumacher, "General fidelity limit for quantum channels," *Physical Review A*, 54:4707, 1996.
  - M. Horodecki, "Limits for compression of quantum information carried by ensembles of mixed states," *Physical Review A*, 57:3364–3369, 1997.
  - S. L. Braunstein, C. A. Fuchs, D. Gottesman, and H. Lo, "A quantum analog of Huffman coding," *IEEE Transactions on Information Theory*, 46(4):1644–1649, 2000.
- Second-Order Analysis of Quantum Data Compression
  - D. Abdelhadi, J. M. Renes, "Second-order asymptotics of quantum data compression from partially-smoothed conditional entropy," 2020 *IEEE International Symposium on Information Theory* (ISIT).

References (2/3)

- Classical Data Compression with Quantum Side Information
  - I. Devetak, A. Winter, "Classical Data Compression with Quantum Side Information," *Physical Review A*, 68(5), 2003.
  - M. Tomamichel and M. Hayashi, "A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks," *IEEE Transactions on Information Theory*, 59(11):7693–7710, 2013.
  - H.-C. Cheng, E. P. Hanson, N. Datta, and M.-H. Hsieh, "Non-Asymptotic Classical Data Compression With Quantum Side Information," *IEEE Transactions on Information Theory*, 67(2):902–930, 2021.
- Quantum Data Compression with Classical Side Information
  - Z. B. Khanian, and A. Winter, "Distributed Compression of Correlated Classical-Quantum Sources or: The Price of Ignorance," *IEEE Transactions on Information Theory*, 66(9):5620–5633, 2020.
- Quantum Slepian–Wolf Coding
  - A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, "The mother of all protocols: restructuring quantum information's family tree," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 465(2108): 2537–2563, 2009.

References (3/3)

- Lossy Quantum Data Compression
  - H. Barnum, "Quantum rate-distortion coding," *Physical Review A*, 62:4, 2000.
  - N. Datta, M.-H. Hsieh, M. M. Wilde, and A. Winter, "Quantum-to-classical rate distortion coding," *Journal of Mathematical Physics*, 54(4):042201, 2013.
  - N. Datta, M.-H. Hsieh, and M. M. Wilde, "Quantum Rate Distortion, Reverse Shannon Theorems, and Source-Channel Separation," *IEEE Transactions on Information Theory*, 59(1):615–630, 2013.
- Quantum Information Bottleneck Function
  - N. Datta, C. Hirche, and A. Winter, "Convexity and Operational Interpretation of the Quantum Information Bottleneck Function," 2019 *IEEE International Symposium on Information Theory* (ISIT).

Quantum Information and Computation
Classical Information over Quantum Channels

Hao-Chung Cheng (鄭皓中)
haochung@ntu.edu.tw

Department of Electrical Engineering
National Taiwan University

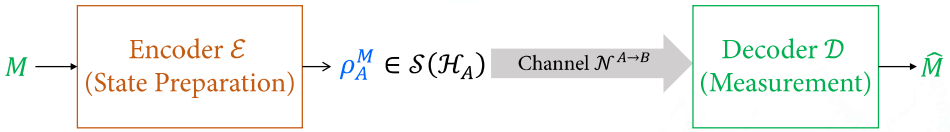
May 19, 2021

Outline

Motivation

Motivation

- How much *classical information* can Alice communicate to Bob by sending a *quantum state*? → Classical communication over a quantum channel.



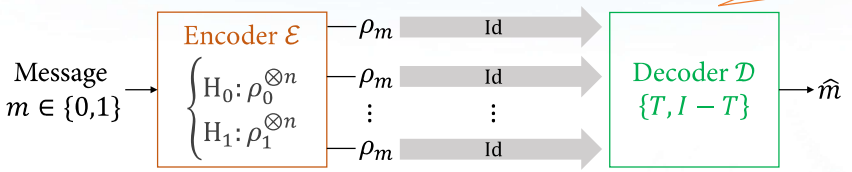
- Since each classical message can always be modeled by a bit string or an index in an alphabet, the goal of the decoder is hence to discriminate among several states.
  → This problem actually boils down to the *quantum state discrimination*.
- In the scenario of communications, one would like to utilize the channel many times (say n) to send $k = \lfloor nR \rfloor$ bit information. Then, how large R is so that we communicate the information *reliably* (i.e. with vanishing error probability)?

Quantum Hypothesis Testing

Binary Quantum Hypothesis Testing (IID)

- Communicating a classical bit through quantum states.

A test $0 \le T \le I$ forms a two-outcome POVM $\{T, I-T\}$.



AEP for quantum hypothesis testing

Quantum Stein's Lemma

An (n, R, ε)-code means $\exists$ a test T such that $\mathrm{Tr}\left[\rho_0^{\otimes n}(I-T)\right] \le \varepsilon$ and $\mathrm{Tr}\left[\rho_1^{\otimes n} T\right] \le \mathrm{e}^{-nR}$.

Then, for all $\varepsilon \in (0,1)$:

- (Achievability) If $R < D(\rho_0||\rho_1)$, $\exists N_0 \in \mathbb{N}$ such that $\exists$ an (n, R, ε)-code for all $n \ge N_0$.
- (Optimality) If $R > D(\rho_0||\rho_1)$, $\exists N_0 \in \mathbb{N}$ such that no (n, R, ε)-codes exist for all $n \ge N_0$.

Proof of Achievability (1/2)

- Definition (Neyman–Pearson Test). $T := \{\rho_0^{\otimes n} \ge \mu \rho_1^{\otimes n}\}$. ← Classical: $T := \left\{y : n \log \frac{\rho_0}{\rho_1} \ge \mu\right\}$

Quantum Chernoff Bound [Audenaert *et al.*'07, 08]

For all positive semi-definite operators $A, B \ge 0$, and $s \in [0,1]$,
$\mathrm{Tr}[A\{A < B\}] + \mathrm{Tr}[B\{A \ge B\}] = \frac{1}{2}\mathrm{Tr}[A + B - |A-B|] \le \mathrm{Tr}[A^{1-s}B^s]$.

1. Bounding Type-I error: $\mathrm{Tr}\left[\rho_0^{\otimes n}(I-T)\right] \le \mu^{1-\alpha}\mathrm{Tr}\left[(\rho_0^{\alpha})^{\otimes n}(\rho_1^{1-\alpha})^{\otimes n}\right]$

$$= \mu^{1-\alpha}(\mathrm{Tr}[\rho_0^{\alpha}\rho_1^{1-\alpha}])^n$$

Bounding Type-II error: $\mathrm{Tr}\left[\rho_1^{\otimes n} T\right] \le \mu^{-\alpha}\mathrm{Tr}\left[(\rho_0^{\alpha})^{\otimes n}(\rho_1^{1-\alpha})^{\otimes n}\right]$

$$= \mu^{-\alpha}(\mathrm{Tr}[\rho_0^{\alpha}\rho_1^{1-\alpha}])^n$$

Proof of Achievability (2/2)

Type-I error $\le \mu^{1-\alpha}(\mathrm{Tr}[\rho_0^{\alpha}\rho_1^{1-\alpha}])^n$
Type-II error $\le \mu^{-\alpha}(\mathrm{Tr}[\rho_0^{\alpha}\rho_1^{1-\alpha}])^n$

2. Choosing $\mu = \mathrm{e}^{\frac{n}{\alpha}(R + \log \mathrm{Tr}[\rho_0^{\alpha}\rho_1^{1-\alpha}])}$, we obtain Type-II error $\le \mathrm{e}^{-nR}$ as desired.

$\Rightarrow$ Type-I error $\le \mathrm{e}^{-n \sup_{\alpha \in [0,1]} \frac{1-\alpha}{\alpha}(D_\alpha(\rho_0||\rho_1) - R)}$. ← $\because$ the Chernoff bound holds $\forall \alpha \in [0,1]$

- Here, $D_\alpha(\rho_0||\rho_1) := \frac{1}{\alpha - 1}\log \mathrm{Tr}[\rho_0^{\alpha}\rho_1^{1-\alpha}]$ is called the *Petz Rényi divergence*.
  - $\lim_{\alpha \to 1} D_\alpha(\rho_0||\rho_1) = D(\rho_0||\rho_1)$; it is a one parameter generalization of $D(\rho_0||\rho_1)$.
  - The map $\alpha \mapsto D_\alpha(\rho_0||\rho_1)$ is non-negative increasing.
  - $\sup_{\alpha \in [0,1]} \frac{1-\alpha}{\alpha}(D_\alpha(\rho_0||\rho_1) - R)$ is called the error exponent.
    It is positive (exponential decaying) if and only if $R < D(\rho_0||\rho_1)$.
- This is called the *quantum Hoeffding bound* (which implies the quantum Stein's lemma). Conversely, it has been proved that the error exponent is optimal. (This also gives Petz's Rényi divergence an operational interpretation.)

□

Proof of Weak Converse (1/2)

1. For any test $0 \le T \le I$, we introduce a quantum-classical operation:
   $\Lambda(\rho^n) \coloneqq \mathrm{Tr}[\rho^n T] \otimes |0\rangle\langle 0| + \mathrm{Tr}[\rho^n (I-T)] \otimes |1\rangle\langle 1|$.
   Type-I error: $\alpha \coloneqq \mathrm{Tr}\big[\rho_0^{\otimes n}(I-T)\big]$; Type-II error: $\beta \coloneqq \mathrm{Tr}\big[\rho_1^{\otimes n} T\big] = e^{-nR}$. (Attaining constraint)
2. Using the data-processing inequality: $D\big(\Lambda(\cdot)||\Lambda(\cdot)\big) \le D(\cdot\,||\,\cdot)$ for any CPTP Λ.

$$\Rightarrow D\big(\rho_0^{\otimes n}||\rho_1^{\otimes n}\big) \ge D\left(\begin{pmatrix} \mathrm{Tr}[\rho_0^{\otimes n}T] & 0 \\ 0 & \mathrm{Tr}[\rho_0^{\otimes n}(I-T)] \end{pmatrix} \middle|\middle| \begin{pmatrix} \mathrm{Tr}[\rho_1^{\otimes n}T] & 0 \\ 0 & \mathrm{Tr}[\rho_1^{\otimes n}(I-T)] \end{pmatrix}\right)$$

($1-\alpha$, α, β, $1-\beta$)

$$= (1-\alpha)\log\frac{1-\alpha}{\beta} + \alpha\log\frac{\alpha}{1-\beta}$$

$$= (1-\alpha)\log(1-\alpha) + \alpha\log\alpha + (1-\alpha)\log\frac{1}{\beta} + \alpha\log\frac{1}{1-\beta}$$

Proof of Weak Converse (2/2)

$$D\big(\rho_0^{\otimes n}||\rho_1^{\otimes n}\big) \ge (1-\alpha)\log(1-\alpha) + \alpha\log\alpha + (1-\alpha)\log\frac{1}{\beta} + \alpha\log\frac{1}{1-\beta}$$

3. Denote the *binary entropy function*: $H(\alpha) \coloneqq -(1-\alpha)\log(1-\alpha) - \alpha\log\alpha$.

$$\Rightarrow D\big(\rho_0^{\otimes n}||\rho_1^{\otimes n}\big) \ge -H(\alpha) + (1-\alpha)\log\frac{1}{\beta} + \alpha\log\frac{1}{1-\beta} \quad (\leftarrow \text{Ignore})$$

$$\ge -H(\alpha) + (1-\alpha)\log\frac{1}{\beta} \quad (\leftarrow nR \text{ by assumption})$$

$$\Rightarrow \alpha \ge \frac{nR - D\big(\rho_0^{\otimes n}||\rho_1^{\otimes n}\big) - H(\alpha)}{nR}$$

(Additivity of $D(\cdot\,||\,\cdot)$)

$$= \frac{R - D(\rho_0||\rho_1)}{R} - \frac{H(\alpha)}{nR} \quad (\to 0)$$

$\Rightarrow \alpha > 0$ if and only if $R > D(\rho_0||\rho_1)$. □

Proof of Strong Converse (1/2)

(Classical: $Q_\alpha(p||q) = \sum_i p_i^\alpha q_i^{1-\alpha}$)

- Definition (Sandwiched Rényi quasi-divergence). $Q_\alpha^*(\rho||\sigma) \coloneqq \mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]$. (See e.g. [Mosonyi–Ogawa'17])

A data-processing inequality

$Q_\alpha^*\big(\Lambda(\rho)||\Lambda(\sigma)\big) \le Q_\alpha^*(\rho||\sigma)$, for all CPTP map Λ and $\alpha \ge 1$.

1. For any test $0 \le T \le I$, we introduce a quantum-classical operation as before:
   $\Lambda(\rho^n) \coloneqq \mathrm{Tr}[\rho^n T] \otimes |0\rangle\langle 0| + \mathrm{Tr}[\rho^n (I-T)] \otimes |1\rangle\langle 1|$.
2. Apply the data-processing inequality: for all $\alpha \ge 1$ and considerthe one-shot case,

$$\Rightarrow Q_\alpha^*(\rho_0||\rho_1) \ge Q_\alpha^*\left(\begin{pmatrix} \mathrm{Tr}[\rho_0 T] & 0 \\ 0 & \mathrm{Tr}[\rho_0(I-T)] \end{pmatrix} \middle|\middle| \begin{pmatrix} \mathrm{Tr}[\rho_1 T] & 0 \\ 0 & \mathrm{Tr}[\rho_1(I-T)] \end{pmatrix}\right) \quad (\downarrow \text{Ignore})$$

$$= (\mathrm{Tr}[\rho_0 T])^\alpha (\mathrm{Tr}[\rho_1 T])^{1-\alpha} + (\mathrm{Tr}[\rho_0(1-T)])^\alpha (\mathrm{Tr}[\rho_1(1-T)])^{1-\alpha}$$

Proof of Strong Converse (2/2)

$$\Rightarrow Q_\alpha^*(\rho_0||\rho_1) \ge (\mathrm{Tr}[\rho_0 T])^\alpha (\mathrm{Tr}[\rho_1 T])^{1-\alpha}$$

($1-$ Type-I error; Type-II error; By assumption Type-II error $\le e^{-R}$)

$$\Rightarrow \mathrm{Tr}[\rho_0 T] \le (\mathrm{Tr}[\rho_1 T])^{\frac{1-\alpha}{\alpha}} \big(Q_\alpha^*(\rho_0||\rho_1)\big)^{\frac{1}{\alpha}} \le e^{\frac{\alpha-1}{\alpha}R} \big(Q_\alpha^*(\rho_0||\rho_1)\big)^{\frac{1}{\alpha}}$$

$$\Rightarrow \text{Type-I success} \le e^{-\sup_{\alpha\ge 1}\frac{\alpha-1}{\alpha}(R - D_\alpha^*(\rho_0||\rho_1))} \quad (\leftarrow \because \text{the upper bound holds } \forall \alpha \ge 1)$$

- Here, $D_\alpha^*(\rho_0||\rho_1) \coloneqq \frac{1}{\alpha-1}\log Q_\alpha^*(\rho_0||\rho_1)$ is called the *sandwiched Rényi divergence*.
  - $\lim_{\alpha\to 1} D_\alpha^*(\rho_0||\rho_1) = D(\rho_0||\rho_1)$; it is a one parameter generalization of $D(\rho_0||\rho_1)$.
  - The map $\alpha \mapsto D_\alpha^*(\rho_0||\rho_1)$ is non-negative increasing.
  - For the n-shot case: Type-I success $\le e^{-n\sup_{\alpha\ge 1}\frac{\alpha-1}{\alpha}(R - D_\alpha^*(\rho_0||\rho_1))}$ $\because D_\alpha^*$ is additive.
  - The exponent is positive (exponential decaying) if and only if $R > D(\rho_0||\rho_1)$.

□

Summary for Binary Hypothesis Testing

Quantum Hoeffding Bound & Strong Converse

An (n, R, ε_n)-code means $\exists$ a test T s.t. $\mathrm{Tr}\left[\rho_0^{\otimes n}(I-T)\right] \le \varepsilon_n$ and $\mathrm{Tr}\left[\rho_1^{\otimes n} T\right] \le \mathrm{e}^{-nR}$.

- (Achievability) If $R < D(\rho_0||\rho_1)$, $\exists$ an (n, R, ε_n)-code s.t. ε_n exponentially decays in n.
- (Strong converse) If $R > D(\rho_0||\rho_1)$, $\forall$ (n, R, ε_n)-codes, $1-\varepsilon_n$ exponentially decays in n.

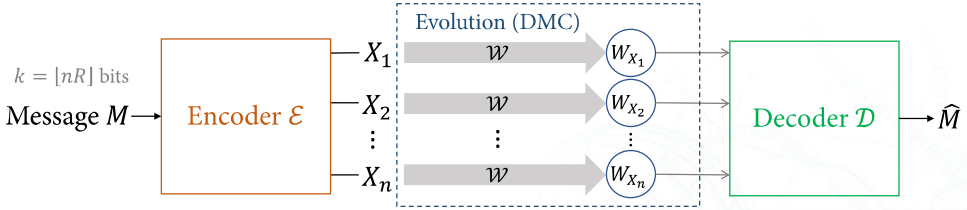
- The binary quantum hypothesis testing has been well studied.
  - The *optimal error exponent* is determined by the Petz's Rényi divergence.
  - The *optimal strong converse exponent* is determined by the sandwiched Rényi divergence.
- However, communicating only one bit of classical information is not satisfiable.
  → We would like to communicate at least nR-bit of classical information.

Classical Channel Coding

Classical Channel Coding (1/3)

$W_{x_1x_2}^{\otimes 2}(y_1, y_2) \coloneqq W_{x_1}(y_1) W_{x_2}(y_2)$

- A *discrete memoryless channel* (DMC) means that the channel $\mathcal{W}^{\otimes n}$ takes in a sequence of symbols $x^n \coloneqq x_1 x_2 \cdots x_n \in \mathcal{X}^n$ and outputs a product distribution $W_{x^n}^{\otimes n} = W_{x_1} \otimes W_{x_2} \otimes \cdots \otimes W_{x_n}$ for each $W_{x_i} \in \mathcal{P}(\mathcal{Y})$.



- The goal of channel coding is to communicate equiprobable messages, say k-bit strings, to the destination with *transmission rate* $R \coloneqq k/n$ and with vanishing error probability $\Pr(\widehat{M} = M) \ge 1-\varepsilon$.
  (For example, classical LDPC codes, and Polar codes are capacity-achieving codes.)

Classical Channel Coding (2/3)

- Definition (Channel code). An (n, R, ε)-code for the channel $\mathcal{W}: \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, is a pair of functions $\mathcal{E}: \mathcal{M}^n \coloneqq \{0,1\}^{\lfloor nR \rfloor} \to \mathcal{X}^n$ and $\mathcal{D}: \mathcal{Y}^n \to \{0,1\}^{\lfloor nR \rfloor}$ such that $\Pr(\mathcal{D} \circ \mathcal{E}(M) \ne M) = \sum_{m \in \mathcal{M}^n} \frac{1}{2^{\lfloor nR \rfloor}} \mathbf{1}_{\mathcal{D} \circ \mathcal{E}(M) \ne M} \le \varepsilon$, for equiprobable message M.

Average error probability

Shannon's (Noisy) Channel Coding Theorem

For a channel $\mathcal{W}: \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, define the channel capacity by $C_{\mathcal{W}} \coloneqq \sup_{P_X \in \mathcal{P}(\mathcal{X})} I(X;Y)$.

Then, for all $\varepsilon \in (0,1)$:

- (Achievability) If $R < C_{\mathcal{W}}$, $\exists N_0 \in \mathbb{N}$ such that $\exists$ an (n, R, ε)-code for all $n \ge N_0$.
- (Optimality) If $R > C_{\mathcal{W}}$, $\exists N_0 \in \mathbb{N}$ such that no (n, R, ε)-codes exist for all $n \ge N_0$.

Classical Channel Coding (3/3)

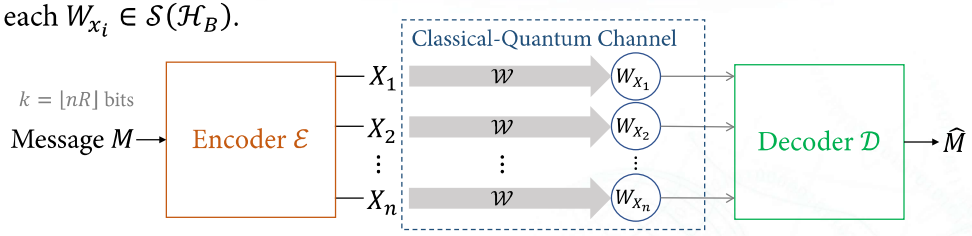
- Given a channel $\mathcal{W}: \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ (a conditional distribution $P_{Y|X}$) and an input distribution $P_X \in \mathcal{P}(\mathcal{X})$, we define the *mutual information* $I(X;Y) \coloneqq D(P_{XY}||P_X \otimes P_Y)$ with respect to the joint distribution P_{XY} induced by the input and channel $\mathcal{W}$. It quantifies the *amount of information* about X by observing Y.
- Since the encoder $\mathcal{E}$ has the freedom to the input distribution P_X on $\mathcal{P}(\mathcal{X})$, we maximize the all possible P_X to obtain the *channel capacity* $C_{\mathcal{W}} \coloneqq \sup_{P_X \in \mathcal{P}(\mathcal{X})} I(X;Y)$.
- Proofs of Shannon's Noisy Channel Coding Theorems:
  - Joint Typicality [Cover–Thomas'06]
  - Packing Lemma [Csiszár–Körner'11]
  - Channel Dispersion [Polyanskiy–Poor–Verdú'11]
  - Large deviation techniques [Gallager'65]
  - Etc…

  Joint Typicality, Packing Lemma → The HSW Approach & Winter's typical decoder
  Channel Dispersion → One-shot converse & quantum AEP
  Large deviation techniques → Hayashi–Nagaoka decoder

Classical-Quantum Channel Coding

Classical-Quantum Channel Coding (1/3)

- A *memoryless classical-quantum channel* $\mathcal{W}^{\otimes n}$ takes in a sequence of symbols $x^n \coloneqq x_1 x_2 \cdots x_n \in \mathcal{X}^n$ and outputs a product state $W_{x^n}^{\otimes n} = W_{x_1} \otimes W_{x_2} \otimes \cdots \otimes W_{x_n}$ for each $W_{x_i} \in \mathcal{S}(\mathcal{H}_B)$.



- The goal of a classical-quantum channel coding is exactly the same as that of a classical channel coding, except the fact that the each channel output is a *quantum state* instead of a *probability distribution*.
- The c-q channel is the simplest quantum channel whose output is not *entangled*.

Classical-Quantum Channel Coding (2/3)

- Definition (Channel code). An (n, R, ε)-code for the channel $\mathcal{W}: \mathcal{X} \to \mathcal{S}(\mathcal{H}_B)$, is a pair of functions $\mathcal{E}: \mathcal{M}^n \coloneqq \{0,1\}^{\lfloor nR \rfloor} \to \mathcal{X}^n$ and $\mathcal{D}: \mathcal{S}(\mathcal{H}_B^{\otimes n}) \to \{0,1\}^{\lfloor nR \rfloor}$ s.t. $\Pr(\mathcal{D} \circ \mathcal{E}(M) \neq M) = \sum_{m \in \mathcal{M}^n} \frac{1}{2^{\lfloor nR \rfloor}} \mathbf{1}_{\mathcal{D} \circ \mathcal{E}(M) \neq M} \leq \varepsilon$, for equally probable message M.

Average error probability

The HSW Theorem (c-q channels)

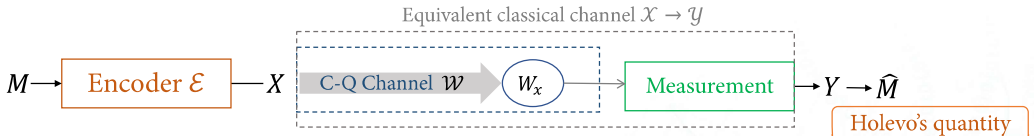
For a channel $\mathcal{W}: \mathcal{X} \to \mathcal{S}(\mathcal{H}_B)$, define the channel capacity by $C_{\mathcal{W}} \coloneqq \sup_{P_X \in \mathcal{P}(\mathcal{X})} I(X;B)_\rho$, where the joint c-q state is $\rho_{XB} \coloneqq \sum_{x \in \mathcal{X}} P_X(x) |x\rangle\langle x| \otimes W_x \in \mathcal{S}(\mathcal{H}_X \otimes \mathcal{H}_B)$.

Then, for all $\varepsilon \in (0,1)$:

- (Achievability) If $R < C_{\mathcal{W}}$, $\exists N_0 \in \mathbb{N}$ such that ∃ an (n, R, ε)-code for all $n \geq N_0$.
- (Optimality) If $R > C_{\mathcal{W}}$, $\exists N_0 \in \mathbb{N}$ such that no (n, R, ε)-codes exist for all $n \geq N_0$.

Classical-Quantum Channel Coding (3/3)

- Suppose that we first apply a *quantum measurement* on $\mathcal{S}(\mathcal{H}_B)$ to obtain classical information (modeled by a random variable Y). Then equivalently, we have a classical channel from input X to output Y, which is characterized by $I(X;Y)$.

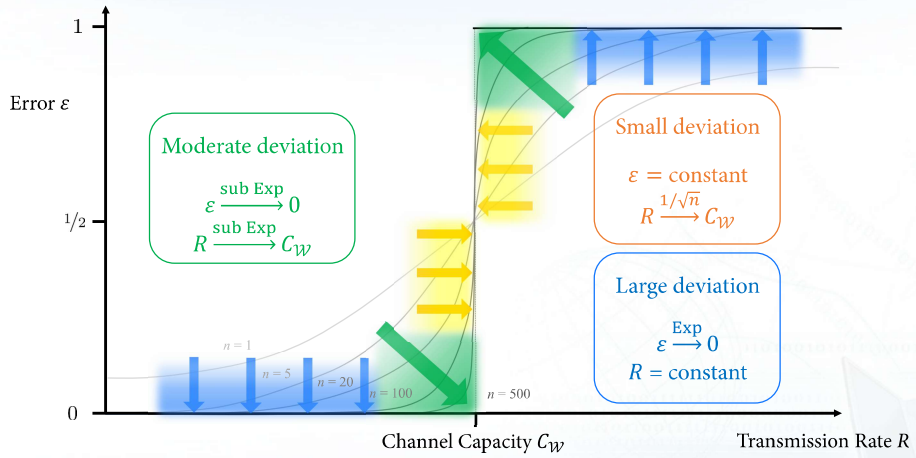


- Theorem (Holevo's bound). $I(X;Y) \le I(X;B)_\rho := D(\rho_{XB} || \rho_X \otimes \rho_B) \le \dim \mathcal{H}_B$.

  Recall that measurement $\Lambda^{B\to Y}(\rho_B) := \mathrm{Tr}[\rho_B \Pi_B^y] \otimes |y\rangle\langle y|$ is a quantum channel. Then Holevo's bound follows from the data-processing inequality of $D(\cdot\,||\,\cdot)$.

- One of the goals of the HSW Theorem is to show that $\sup_{P_X} I(X;B)_\rho$ is an *achievable rate* by means of classical encoders $\mathcal{E}$, and a quantum decoder $\mathcal{D}$.

Characterization



Proof of Achievability (1/4)

- We would like to communicate M messages by the following encoder and decoder.
- Encoder: Each message m is mapped to a codeword $x(m)$ independently with $P_X(x)$.
- Decoder: Apply a POVM $\{\Pi^m\}_m$ to decode each message m as follows: Random coding

  For each m, let $\Lambda_m := \{W_{x(m)} \ge \mu\sigma\}$ for some $\mu \ge 0$, and state $\sigma \ge 0$.
  This means that we decode to m if the log-likelihood ratio (w.r.t. σ) is large.

  However, it is not easy to check if $\sum_m \Lambda_m = I$ (i.e. they form a valid POVM).
  So, we choose $\Pi_m := (\sum_i \Lambda_i)^{-1/2} \Lambda_m (\sum_i \Lambda_i)^{-1/2}$.

  $\Rightarrow I - \Pi_m \le 2(I - \Lambda_m) + 4\sum_{i\neq m} \Lambda_i$.

A union bound lemma (Hayashi–Nagaoka)

$$I - (A+B)^{-1/2} A (A+B)^{-1/2} \le 2(I-A) + 4B, \forall A, B \ge 0.$$

Proof of Achievability (2/4)

For $A \ge 0$ and $B \le C$, $\mathrm{Tr}[AB] \le \mathrm{Tr}[AC]$.

1. Conditional error $P_{e,m} = \mathrm{Tr}[W_{x(m)}(I - \Pi_m)]$

   Chernoff bound

   $$\le 2\mathrm{Tr}[W_{x(m)}\{W_{x(m)} < \mu\sigma\}] + 4\sum_{i\neq m} \mathrm{Tr}[W_{x(m)}\{W_{x(i)} \ge \mu\sigma\}]$$
   $$\le 2\mathrm{Tr}\left[(W_{x(m)})^{1-s}(\mu\sigma)^s\right] + 4\sum_{i\neq m} \mathrm{Tr}[W_{x(m)}\{W_{x(i)} \ge \mu\sigma\}]$$

2. Invoking random coding to take average (w.r.t. to P_X) *independent* for each m.

   $\Rightarrow \mathbb{E}\left[\sum_{i\neq m} \mathrm{Tr}[W_{X(m)}\{W_{X(i)} \ge \mu\sigma\}]\right] = \mathbb{E}\left[\sum_{i\neq m} \mathrm{Tr}[\mathbb{E}[W_X]\{W_{X(i)} \ge \mu\sigma\}]\right]$

3. Choosing $\sigma = \mathbb{E}[W_X]$, and using the Chernoff bound on the second part:

   $$\Rightarrow \mathbb{E}\left[\sum_{i\neq m} \mathrm{Tr}[\sigma\{W_{x(i)} \ge \mu\sigma\}]\right] \le \mu^{-1}\mathbb{E}\left[\sum_{i\neq m} \mathrm{Tr}\left[(W_{X(i)})^{1-s}(\mu\sigma)^s\right]\right]$$
   $$= \mu^{s-1}(|\mathcal{M}| - 1)\mathrm{Tr}[\mathbb{E}[W_X^{1-s}]\sigma^s]$$

Proof of Achievability (3/4)

Message independent!

$$\mathbb{E}[P_{e,m}] \le 2\mu^s \mathrm{Tr}[\mathbb{E}[W_X^{1-s}]\sigma^s] + 4\mu^{s-1}(|\mathcal{M}|-1)\mathrm{Tr}[\mathbb{E}[W_X^{1-s}]\sigma^s]$$

4. Choosing $\mu = (|\mathcal{M}|-1)$, and taking average over all equiprobable messages m.

$$\Rightarrow \mathbb{E}[P_e] \le 2(|\mathcal{M}|-1)^s \mathrm{Tr}[\mathbb{E}[W_X^{1-s}]\sigma^s] + 4(|\mathcal{M}|-1)^s \mathrm{Tr}[\mathbb{E}[W_X^{1-s}]\sigma^s]$$
$$\le 6|\mathcal{M}|^s \mathrm{Tr}[\mathbb{E}[W_X^{1-s}]\sigma^s]$$

5. Choosing $\alpha = 1/(1+s) \in [0.5, 1]$, and write the formula in exponential.

↓ ∵ the upper bound holds $\forall \alpha \in [0.5,1]$

$$\Rightarrow \mathbb{E}[P_e] \le 6 \cdot 2^{-\sup_{\alpha\in[0.5,1]} \frac{1-\alpha}{\alpha}\left(I^{\downarrow}_{2-1/\alpha}(X;B)_\rho - \log|\mathcal{M}|\right)}$$

Exponential decay if and only if $\log|\mathcal{M}| < I_1^{\downarrow}(X;B)_\rho = I(X;B)_\rho$

$$I_\alpha^{\downarrow}(X;B)_\rho := D_\alpha(\rho_{XB}||\rho_X \otimes \rho_B) \equiv \frac{1}{\alpha-1}\log \mathrm{Tr}[\rho_{XB}^\alpha \rho_X^{1-\alpha} \otimes \rho_B^{1-\alpha}]$$
$$\rho_{XB} := \sum_x P_X(x)|x\rangle\langle x| \otimes W_x,\ \rho_X = \sum_x P_X(x)|x\rangle\langle x|,\ \rho_B = \mathbb{E}[W_X]$$

Proof of Achievability (4/4)

$$\mathbb{E}[P_e] \le 2^{-\sup_{\alpha\in[0.5,1]} \frac{1-\alpha}{\alpha}\left(I^{\downarrow}_{2-1/\alpha}(X;B)_\rho - \log|\mathcal{M}|\right)}$$

6. IID extension: encoding $\mathcal{E}: m \mapsto x^n(m)$ with probability $P_X(x_1)\cdots P_X(x_n)$.

Channel output: $W_{x^n(m)}^{\otimes n} = W_{x_1(m)} \otimes \cdots \otimes W_{x_2(m)}$

$$\Rightarrow \rho_{X^nB^n} = \sum_{x^n} P_{x^n}(x^n)|x^n\rangle\langle x^n| \otimes W_{x^n}^{\otimes n} = \rho_{XB}^{\otimes n}$$
$$\Rightarrow I^{\downarrow}_{2-1/\alpha}(X^n;B^n)_{\rho^{\otimes n}} = n\, I^{\downarrow}_{2-1/\alpha}(X;B)_\rho$$

By $M := 2^{\lceil nR \rceil}$, $\mathbb{E}[P_e] \le 6 \cdot 2^{-n\sup_{\alpha\in[0.5,1]} \frac{1-\alpha}{\alpha}\left(I^{\downarrow}_{2-1/\alpha}(X;B)_\rho - \frac{1}{n}\log|\mathcal{M}|\right)}$

7. Lastly, we are free to choose best $P_X \in \mathcal{P}(\mathcal{X})$.

Exponential decay if and only if $R := \frac{1}{n}\log|\mathcal{M}| < I(X;B)_\rho$

□

Proof of Weak Converse (1/2)

$H(q) := -q\log q - (1-q)\log(1-q)$

1. Fano's inequality

$$H(M|\widehat{M}) \le H(q) + q\log(|\mathcal{M}|-1), \text{ where } q = \Pr(\widehat{M} \ne M).$$

$$\Rightarrow \Pr(\widehat{M} \ne M) \ge \frac{H(M|\widehat{M}) - H(q)}{\log(|\mathcal{M}|-1)} \ge \frac{H(M|\widehat{M}) - H(q)}{\log|\mathcal{M}|}$$

≈ constant. ← nR.

Equiprobable messages

2. Evaluating the mutual information: $H(M|\widehat{M}) = H(M) - I(M;\widehat{M}) = nR - I(M;\widehat{M})$.

Holevo's bound: $I(M;\widehat{M}) \le I(X^n;B^n)_\rho$ from the Markov chain $M \to X^n \to B^n \to \widehat{M}$.

$$\Rightarrow \rho_{X^nB^n} = \sum_{x^n} P_{x^n}(x^n)|x^n\rangle\langle x^n| \otimes W_{x^n}^{\otimes n} = \rho_{XB}^{\otimes n} \Rightarrow I(X^n;B^n)_\rho = n\, I(X;B)_\rho$$

$$\Rightarrow H(M|\widehat{M}) \ge nR - nI(X;B)_\rho \ge nR - n\max_{P_X} I(X;B)_\rho$$

Capacity C_W

Proof of Weak Converse (2/2)

$H(q) := -q\log q - (1-q)\log(1-q)$

$$\Rightarrow \Pr(\widehat{M} \ne M) \ge \frac{n(R - C_W) - H(q)}{nR} = \frac{R - C_W}{R} - \frac{H(q)}{nR}$$

$\to 0$

$$\Rightarrow \Pr(\widehat{M} \ne M) > 0 \text{ if and only if } R > C_W.$$

- We term this a *weak converse* because we only show the error probability is > 0. As we will show later, the *strong converse* means that actually error probability $\to 1$.
- This proof relies on a key fact – the Holevo quantity $\sup_{P_X} I(X;B)_\rho$ is additive under IID extension: $\sup_{P_{X^n}} I(X^n;B^n)_\rho = n \max_{P_X} I(X;B)_\rho$, which is because the channel output is always *separable* (i.e. classical correlation). There are quantum channels whose Holevo quantity is *superadditive* – the channel output could be *entangled*!

□

Proof of Strong Converse (1/3)

Classical: $Q_\alpha(p||q) = \sum_i p_i^\alpha q_i^{1-\alpha}$

- Definition (Sandwiched Rényi quasi-divergence). $Q_\alpha^*(\rho||\sigma) \coloneqq \mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]$.

A data-processing inequality

$Q_\alpha^*\big(\Lambda(\rho)||\Lambda(\sigma)\big) \le Q_\alpha^*(\rho||\sigma)$, for all CPTP map Λ and $\alpha \ge 1$.

See e.g. [Mosonyi–Ogawa'17]

Any codes $m \mapsto x(m)$

1. Introduce: (i) a classical-quantum state $\rho \coloneqq \frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}|m\rangle\langle m| \otimes W_{x(m)}$;

   (ii) an auxiliary comparison state $\tau \coloneqq \frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}|m\rangle\langle m| \otimes \sigma_B$;

   Any POVM $\{\Pi^m\}_m$

   (iii) any test (in a direct sum form) $T \coloneqq \sum_{m\in\mathcal{M}}|m\rangle\langle m| \otimes \Pi^m \ge 0$.

   $\Rightarrow \Lambda(\rho) \coloneqq \mathrm{Tr}[\rho T] \otimes |0\rangle\langle 0| + \mathrm{Tr}[\rho(I-T)] \otimes |1\rangle\langle 1|$

   $\{T, I-T\}$ forms a two-outcome POVM (i.e. a quantum-classical channel Λ)

Proof of Strong Converse (2/3)

$\rho \coloneqq \frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}|m\rangle\langle m| \otimes W_{x(m)}$
$\tau \coloneqq \frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}|m\rangle\langle m| \otimes \sigma_B$

2. Data-processing inequality: for all $\alpha \ge 1$,

$$\Rightarrow Q_\alpha^*(\rho||\tau) \ge Q_\alpha^*\big(\Lambda(\rho)||\Lambda(\tau)\big) = Q_\alpha^*\left(\begin{pmatrix}\mathrm{Tr}[\rho T] & 0\\ 0 & \mathrm{Tr}[\rho(I-T)]\end{pmatrix}||\begin{pmatrix}\mathrm{Tr}[\tau T] & 0\\ 0 & \mathrm{Tr}[\tau(I-T)]\end{pmatrix}\right)$$

$$= (\mathrm{Tr}[\rho T])^\alpha(\mathrm{Tr}[\tau T])^{1-\alpha} + (\mathrm{Tr}[\rho(1-T)])^\alpha(\mathrm{Tr}[\tau(1-T)])^{1-\alpha} \quad \leftarrow \text{Ignore}$$

$$\ge (\mathrm{Tr}[\rho T])^\alpha(\mathrm{Tr}[\tau T])^{1-\alpha} = P_{\mathrm{s}}^\alpha \frac{1}{|\mathcal{M}|^{1-\alpha}}$$

$$\Rightarrow P_{\mathrm{s}} \le |\mathcal{M}|^{\frac{1-\alpha}{\alpha}}\big(Q_\alpha^*(\rho||\tau)\big)^{\frac{1}{\alpha}}, \text{ where } Q_\alpha^*(\rho||\tau) = \frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}\mathrm{Tr}\left[W_{x(m)}^\alpha\sigma_B^{1-\alpha}\right] \le \sup_{x\in\mathcal{X}}\mathrm{Tr}\left[W_x^\alpha\sigma_B^{1-\alpha}\right]$$

3. This holds for all $\alpha \ge 1$ and σ_B. Hence, we take $\inf_{\alpha\ge1}\inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)}$ on RHS.

Proof of Strong Converse (3/3)

- Definition (Sandwiched Rényi divergence radius).
  $C_{\alpha,\mathcal{W}}^* \coloneqq \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)}\sup_{x\in\mathcal{X}} D_\alpha^*(W_x||\sigma_B)$, where $D_\alpha^*(\rho||\sigma) \coloneqq \frac{1}{\alpha-1}\log Q_\alpha^*(W_x||\sigma_B)$.

$\rho_{XB} = \sum_x P_X(x)|x\rangle\langle x| \otimes W_x$

A minimax identity [Mosonyi–Ogawa'17]

$C_{\alpha,\mathcal{W}}^* = \sup_{P_X\in\mathcal{P}(\mathcal{X})} I_\alpha^*(X;B)_\rho$, where $I_\alpha^*(X;B)_\rho \coloneqq \inf_{\sigma_B\in\mathcal{S}(\mathcal{H}_B)} D_\alpha^*(\rho_{XB}||\rho_X\otimes\sigma_B)$.

$\Rightarrow P_{\mathrm{s}} \le 2^{-\sup_{\alpha\ge1}\frac{\alpha-1}{\alpha}(\log|\mathcal{M}|-C_{\alpha,\mathcal{W}}^*)}$

Exponential decay if and only if $\log|\mathcal{M}| > \lim_{\alpha\to1} C_{\alpha,\mathcal{W}}^* = C_{\mathcal{W}}$

4. IID extension: $I_\alpha^*(X^n;B^n)_\rho = nI_\alpha^*(X;B)_\rho$,

$\Rightarrow P_{\mathrm{s}} \le 2^{-n\sup_{\alpha\ge1}\frac{\alpha-1}{\alpha}(R-C_{\alpha,\mathcal{W}}^*)}$

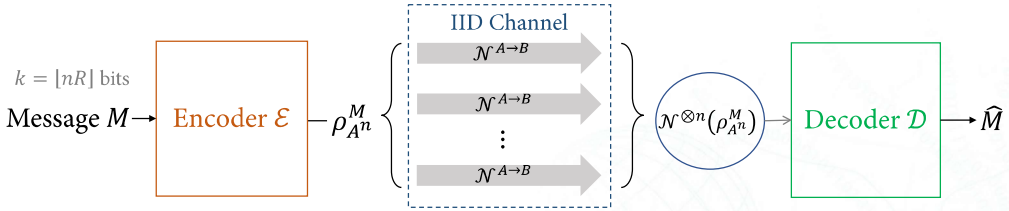
Exponential decay if and only if $R \coloneqq \frac{1}{n}\log|\mathcal{M}| > \lim_{\alpha\to1} C_{\alpha,\mathcal{W}}^* = C_{\mathcal{W}}$

□

Beyond Classical-Quantum Channels

Classical Information over Quantum Channels

- Consider *memoryless quantum channel* $(\mathcal{N}^{A\to B})^{\otimes n}: \mathcal{S}(\mathcal{H}_A^{\otimes n}) \to \mathcal{S}(\mathcal{H}_B^{\otimes n})$. The message m is mapped to a state $\rho_{A^n}^m \in \mathcal{S}(\mathcal{H}_A^{\otimes n})$ and sent to the channel.



- For $n = 1$, one can take $W_{x(m)} := \mathcal{N}(\rho_A^m)$ as in the c-q channel coding.
- Since we are free to choose the c-q encoding $m \mapsto \rho_A^m$, the (one-shot) channel capacity is given by $C_{\mathcal{N}} = \sup_{\rho_{XA}} I(X;B)_{\mathrm{Id}\otimes\mathcal{N}^{A\to B}(\rho_{XA})}$. ← $\rho_{XA} := \sum_x P_X(x)|x\rangle\langle x| \otimes \rho_A^x$

The Holevo Quantity (1/2)

- For the n–shot case, one can send an *entangled state* (among A^n). We view it as an one-shot c-q channel case: $m \mapsto \rho_{A^n}^m \mapsto \mathcal{N}^{\otimes n}(\rho_{A^n}^m) =: W_{x(m)}$.
- Again, from the previous analysis, the averaged n-shot channel capacity (Holevo quantity) is given by $\frac{1}{n} C_{\mathcal{N}^{\otimes n}} := \frac{1}{n} \sup_{\rho_{XA^n}} I(X;B^n)_{\mathrm{Id}\otimes\mathcal{N}^{\otimes n}(\rho_{XA^n})}$.

$\Rightarrow C_{\mathcal{N}}^{\infty} := \lim_{n\to\infty} \frac{1}{n} C_{\mathcal{N}^{\otimes n}}$ called the *regularized Holevo quantity* for channel $\mathcal{N}$.

called *classical capacity* of $\mathcal{N}$

The HSW Theorem (general channels)

For all $\varepsilon \in (0,1)$:

- (Achievability) If $R < C_{\mathcal{N}}^{\infty}$, $\exists N_0 \in \mathbb{N}$ such that ∃ an (n, R, ε)-code for all $n \geq N_0$.
- (Optimality) If $R > C_{\mathcal{N}}^{\infty}$, $\exists N_0 \in \mathbb{N}$ such that no (n, R, ε)-codes exist for all $n \geq N_0$.

The Holevo Quantity (2/2)

- A crucial question: Is the Holevo quantity *additive*? E.g. $C_{\mathcal{N}^{\otimes 2}} = 2C_{\mathcal{N}}$?
- There are some channels whose Holevo quantities are additive [§20.4, Wilde'17].
  - Classical-quantum (c-q) channels: $x \mapsto W_x$.
  - Entanglement-breaking (EB) channels: $\rho_A \mapsto \sum_\omega \mathrm{Tr}[\rho_A \Pi_A^\omega] \rho_B^\omega$.
  - Quantum Hadamard channels (whose complementary channel is EB).
  - Quantum erasure channel: $\rho_A \mapsto (1-p)\rho_A + p|\mathrm{e}\rangle\langle\mathrm{e}|_{A^\perp}$.
  - Depolarizing channel: $\rho_A \mapsto (1-p)\rho_A + p\, I/\dim \mathcal{H}_A$.
  - Quantum covariant channel $\mathcal{N}$: $\mathcal{N}(U\rho_A U^\dagger) = U\mathcal{N}(\rho_A)U^\dagger$ for all unitaries U.

The output state is EB.

- In 2009, Hasting proves that, in general, the Holevo quantity is super-additive, i.e. $\exists \mathcal{N}$ such that $C_{\mathcal{N}^{\otimes 2}} < 2C_{\mathcal{N}}$. This means that *entangled inputs* indeed increase the transmission rate. However, this makes computing $C_{\mathcal{N}}^{\infty}$ highly intractable! This is a large notorious and open research field (and also the strong converse).

Concluding Remarks

Concluding Remarks

- How much classical information can Alice communicate to Bob by sending a quantum state? Holevo's bound tells us that, unfortunately, it is upper bounded by the dimension of the output Hilbert space; → 1 qubit $\not\geq$ N cbits $\forall N > 1$.
- However, if every physical medium is quantum-mechanical, by the data-processing inequality, we can achieve higher transmission rate than the purely classical one.
- This also motivates the study of *entanglement-assisted classical communication*.
- We studied the trade-off between the (i) transmission rate R, (ii) block-length n, and (iii) performance – error probability P_e of communicating classical information over a quantum channel $\mathcal{N}$. The, the HSW Theorem shows that the (regularized) Holevo quantity $C_{\mathcal{N}}^{\infty}$ is a fundamental limit on R, for which $P_e \to 0$ or > 0 (as $n \to \infty$).
  - Fix $P_e \leq \varepsilon$, study $R = C_{\mathcal{N}}^{\infty} + \frac{1}{\sqrt{n}} V_{\mathcal{N}}^{\infty} + o\left(\frac{1}{\sqrt{n}}\right)$: second-order analysis (generally open). ↑ reliable
  - Fix $R < C_{\mathcal{N}}^{\infty}$, study how fast $P_e \to 0$: error-exponent analysis (still open c-q channels).
  - Fix $R > C_{\mathcal{N}}^{\infty}$, study how fast $P_e \to 1$: strong converse exponent (known for additive $C_{\mathcal{N}}^{\infty}$).

Appendix: Quantum Chernoff Bound

Proof of The Quantum Chernoff Bound (1/2)

1. Using the decomposition $A - B = (A - B)_+ - (A - B)_-$, we have
   $\frac{1}{2}(A + B - |A - B|) = A - (A - B)_+ = B - (B - A)_+ = A\{A < B\} + B\{A \geq B\}$.
   ⇒ It is equivalently to prove $\mathrm{Tr}[A] - \mathrm{Tr}[A^{1-s}B^s] = \mathrm{Tr}[A^{1-s}(A^s - B^s)] \leq \mathrm{Tr}(A - B)_+$.

2. The key technique is the following matrix monotone function:

Matrix Monotone Function

For all positive semi-definite operators $0 \leq A \leq B$, then $A^s \leq B^s$ for all $s \in [0,1]$.

∗ Note that not all power functions are matrix monotone, e.g. $\exists 0 \leq A \leq B$, $e^A \not\leq e^B$.

Proof of The Quantum Chernoff Bound (2/2)

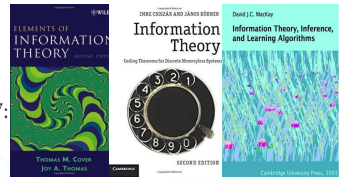
3. Use matrix monotonicity and the following facts:
   - $A \leq A + (A - B)_- = B + (A - B)_+$.
   - $B \leq B + (A - B)_+$.

$$
\begin{aligned}
\Rightarrow \mathrm{Tr}[A^{1-s}(A^s - B^s)] &\leq \mathrm{Tr}\left[\left((B + (A - B)_+)^s - B^s\right)A^{1-s}\right] \\
&\leq \mathrm{Tr}\left[\left((B + (A - B)_+)^s - B^s\right)(B + (A - B)_+)^{1-s}\right] \\
&= \mathrm{Tr}[B + (A - B)_+] - \mathrm{Tr}[B^s(B + (A - B)_+)^{1-s}] \\
&\leq \mathrm{Tr}[B + (A - B)_+] - \mathrm{Tr}[B^s B^{1-s}] \\
&= \mathrm{Tr}[(A - B)_+].
\end{aligned}
$$

□

References (1/5)

- Books on Classical Information Theory
  - T. M. Cover and J. A. Thomas, *Elements of Information Theory* (2nd Edition), Wiley-Interscience, 2006.
  - I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems* (2nd Edition), Cambridge University Press; 2nd edition, 2015.
  - D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 2003.
  - Y. Polyanskiy and Y. Wu, *Lecture Notes on Information Theory*; (v5), 2017: http://people.lids.mit.edu/yp/homepage/data/itlectures_v5.pdf



- Classical Second-Order Analysis (Channel Dispersion)
  - Y. Polyanskiy, H. V. Poor and S. Verdú, "Channel Coding Rate in the Finite Blocklength Regime," *IEEE Transactions on Information Theory*, 56(5), 2307–2359, 2010.
  - V. Y. F. Tan, *Asymptotic Estimates in Information Theory with Non-Vanishing Error Probabilities*, Foundations and Trends® in Communications and Information Theory, 11(1-2):1–184, 2014.
- Classical Error Exponent Analysis (Large Deviation)
  - R. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE IT*, 11(1):3–18, 1965.

References (2/5)

- Quantum Typicality & Packing Lemma & The HSW Theorem
  - A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Transaction on Information Theory*, 45(7):2481-2485, 1999; *Coding Theorems of Quantum Information Theory*, (Ph.D Thesis, Universitat Bielefeld), arXiv:quant-ph/9907077.
  - A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problems of Information Transmission*, 9:177–183, 1973.
  - A. S. Holevo, "Statistical problems in quantum physics," In Second Japan-USSR Symposium on Probability Theory, volume 330 of Lecture Notes in Mathematics: 104–119, 1973.
  - A. S. Holevo, "The capacity of the quantum channel with general signal states,"
  - *IEEE Transactions on Information Theory*, 44:269–273, 1998.
  - B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Physical Review A*, 56(1):131 – 138, 1997.
- Strong Converse for Classical-Quantum Channel Coding
  - M. Mosonyi and T. Ogawa, "Strong Converse Exponent for Classical-Quantum Channel Coding," *Communications in Mathematical Physics*, 355(1):373–426, 2017.

References (3/5)

- Books and Monographs on Quantum Information Theory
  - M. Tomamichel, *Quantum Information Processing with Finite Resources*, Springer Publisher, 2015.
  - M. M. Wilde, *Quantum Information Theory* (2nd Edition), Cambridge University Press, 2017.
  - S. Khatri, and M. M. Wilde, "Principles of Quantum Communication Theory: A Modern Approach," arXiv:2011.04672 [quant-ph].
  - J. Watrous, *The Theory of Quantum Information*, Cambridge University Press, 2019.
  - M. Hayashi , S. Ishizaka, A. Kawachi, G. Kimura, and T. Ogawa, *Introduction to Quantum Information Science*, Springer, 2015.
  - A. S. Holevo, *Quantum Systems, Channels, Information*, de Gruyter, 2013.
  - A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, Springer, 2011.
  - A. S. Holevo, *Statistical Structure of Quantum Theory*, Springer, 2001.

References (4/5)

- Quantum Hypothesis Testing
  - K. M. R. Audenaert, M. Nussbaum, A. Szkoła, F. Verstraete, "Asymptotic Error Rates in Quantum Hypothesis Testing," *Communications in Mathematical Physics*, 279, 251–283, 2008.
  - M. Nussbaum, A. Szkoła, "The Chernoff lower bound for symmetric quantum hypothesis testing," *The Annals of Statistics*, 37(2), 2009.
  - K. Li, "Second-order asymptotics for quantum hypothesis testing," *The Annals of Statistics*, 42(1):171–189, 2014.
  - K. Li, "Discriminating quantum states: The multiple Chernoff distance," *The Annals of Statistics*, 44(4), 2016.
- On the additivity of the Holevo quantity
  - M. B. Hastings, "Superadditivity of communication capacity using entangled inputs," *Nature Physics*, 5(255), 2009.
  - G. Aubrnn, S. Szarek, and E. Werner, "Hastings's Additivity Counterexample via Dvoretzky's Theorem," *Communications in Mathematical Physics*, 305, 85–97, 2011.

References (5/5)

- Classical-Quantum Channel Coding
  - M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Transaction on Information Theory*, 49(7):1753–1768, 2003.
  - M. Hayashi, "Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding," *Physical Review A*, 76(6), 2007.
  - H.-C. Cheng, M.-H. Hsieh, and M. Tomamichel, "Quantum Sphere-Packing Bounds with Polynomial Prefactors," *IEEE Transaction on Information Theory*, 65(5):2872–2898, 2019.
  - H.-C. Cheng, and M.-H. Hsieh, "Moderate Deviation Analysis for Classical-Quantum Channels and Quantum Hypothesis Testing," *IEEE Transaction on Information Theory*, 64(2):1385–1403, 2018.
- Classical-Quantum Channel Coding (Second-Order Analysis)
  - L. Wang and R. Renner, "One-Shot Classical-Quantum Capacity and Hypothesis Testing," *Physical Review Letters,* 108(20):200501, 2012.
  - M. Tomamichel and M. Hayashi, "A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks," *IEEE Transaction on Information Theory*, 59(11):7693–7710, 2013.
  - W. Matthews and S. Wehner, "Finite Blocklength Converse Bounds for Quantum Channels," *IEEE Transaction on Information Theory*, 60(11):7317–7329, 2014.

Quantum Information and Computation
Quantum Information over Quantum Channels

Hao-Chung Cheng (鄭皓中)

haochung@ntu.edu.tw

Department of Electrical Engineering

National Taiwan University

May 26, 2021

Outline

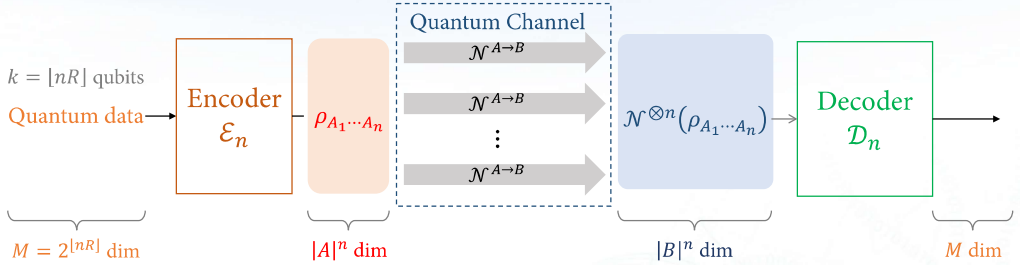
Motivation

Motivation

- How much *quantum information* can be *reliably* transmitted over a *noisy* quantum channel? → quantum communication over a quantum channel.
- Can we *reliably* communicate *classical information* over a quantum channel in a *private* way? → private classical communication over a quantum channel.
- Using additional *resources* to *assist* the communication.
  - The given resources should not trivialized the communication task.
    E.g. the entanglement should be independent of the message
  - Free resource may be easy to obtain in real life (perspective from *Resource Theories*).
  - Assisted-capacities may be easier to calculate (yielding useful bounds).
  - More insight and understanding to communication problems.
  - Leading to other quantum information-theoretic tasks such as the *reverse Shannon theorem*, *quantum error correction codes*, and *entanglement purification*.

Quantum Communication

Quantum Communication



- **Definition** (Quantum channel code). An (n, R, ε)-code for the channel $\mathcal{N}^{A\to B}$, is a pair of functions $\mathcal{E}_n^{2^{\lfloor nR \rfloor}\to A^n}$ and $\mathcal{D}_n^{B^n\to M}$ such that ***error*** $\leq \varepsilon$.
- **Definition** (Quantum capacity). $Q(\mathcal{N}) := \sup\{R: \forall \varepsilon \in (0,1), \exists\ (n, R, \varepsilon)\text{−code}\}$.
  ↑ *achievable rate*
- *Question*: How to define the error?

Plausible Error Criteria (1/2)

1. Evaluating on joint system M and the reference system R.

(a) $\forall |\psi\rangle, \langle\psi|\hat{\sigma}|\psi\rangle \geq 1-\varepsilon \Leftrightarrow$ *Channel Fidelity* $\inf_{\rho_M} F_c(\rho_M, \mathcal{D}_n \circ \mathcal{N}^{\otimes n} \circ \mathcal{E}_n) \geq 1-\varepsilon$

$\Leftrightarrow$ *Diamond norm* $\|\mathcal{D}_n \circ \mathcal{N}^{\otimes n} \circ \mathcal{E}_n - \mathrm{Id}\|_\diamond \leq \sqrt{\varepsilon}$

↑ *Diamond norm* $\|\Lambda\|_\diamond := \sup_\rho \|\Lambda \otimes \mathrm{Id}(\rho)\|_1$

← the code simulates Id channel in $\|\cdot\|_\diamond$

(b) For MES $|\Phi\rangle$, $\langle\Phi|\hat{\sigma}|\Phi\rangle \geq 1-\varepsilon$

← Choi-state for the code simulates MES

Plausible Error Criteria (2/2)

2. Evaluating only on single system M.

$|\phi\rangle_M$ → Encoder $\mathcal{E}_n$ → Channel $\mathcal{N}^{\otimes n}$ → Decoder $\mathcal{D}$ → $\hat{\rho}_M$

(a) $\forall|\phi\rangle, \langle\phi|\hat{\rho}|\phi\rangle \geq 1-\varepsilon$ ← *worst case* pure-input / output overlap

(b) $\mathbb{E}_{|\phi\rangle}\langle\phi|\hat{\rho}|\phi\rangle \geq 1-\varepsilon$ ← *average case* pure-input / output overlap

↑ uniform (Haar) distribution

- The four criteria are not equivalent, e.g. 1. (a) is the strongest.
- However, the four criteria all give the same *quantum capacity* $Q(\mathcal{N})$. (Idea similar to the expurgation method in classical channel coding.)

[Barnum–Knill–Nielsen'00], [Kretschmann–Werner'04]

An Important Information Quantity

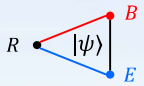
R and B not symmetric

- Definition (Coherent information). Let ρ_{RB} be a quantum state. The *coherent information* (from R to B) is defined as $I_c(R\rangle B)_\rho := H(B)_\rho - H(RB)_\rho \equiv -H(R|B)_\rho$.
- Let $|\psi\rangle_{RBE}$ be a *purification* of ρ_{RB} (it is *unique* up to unitary on E).

$$I_c(R\rangle B)_\rho := H(B)_\rho - H(RB)_\rho$$
$$= H(B)_\rho - H(E)_{|\psi\rangle\langle\psi|}$$ ← from the Schmidt decomposition: $H(E) = H(RB)$
$$= \frac{1}{2}\left(I(B:R)_{|\psi\rangle\langle\psi|} - I(E:R)_{|\psi\rangle\langle\psi|}\right)$$ ← from the Def. of QMI: $I(B:R) = H(B) + H(E) - H(BE)$

Meaning: $\frac{1}{2}$ (how much more R is correlated with B than with E in QMI)

Examples



$|\Phi\rangle = 1$ ebit

	$\|\psi\rangle_{RBE}$	$H(B)$	$H(E)$	$I(R;B)$	$I(R;E)$	$I_c(R\rangle B)$
	$\|\Phi\rangle_{RB}\|0\rangle_E$	1	0	2	0	1
	$\|0\rangle_B\|\Phi\rangle_{RE}$	0	1	0	2	−1
•	$\|0\rangle_R\|0\rangle_B\|0\rangle_E$	0	0	0	0	0
•	$\|0\rangle_R\|\Phi\rangle_{BE}$	1	1	0	0	0
•	$\|0\rangle_R\|0\rangle_B\|0\rangle_E + \|1\rangle_R\|1\rangle_B\|1\rangle_E$	1	1	1	1	0

GHZ

Total correlation between RB ↑

Quantum correlation between RB ↑

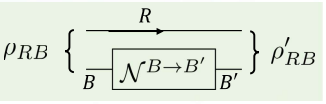
Coherent Information

$I_c(R\rangle B)_\rho := H(B)_\rho - H(RB)_\rho$

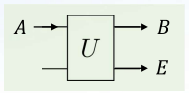
↑ suggests $H(R|B)_\rho$ could be negative!

[Horodecki–Oppenheim–Winter'05]

- Properties:
  - Invariant under local unitaries on (R, B, or E separately).
  - Invariant under appending or discarding *local pure states*.
  - Non-increasing under local operations (CPTP maps) on B: $I_c(R\rangle B)_\rho \geq I_c(R\rangle B')_{\mathrm{Id}\otimes\mathcal{N}(\rho)}$ for any CPTP $\mathcal{N}^{B\to B'}$.
  - Can be increasing or decreasing under local operation of R. (E.g. Discarding anything in R goes to the environment E.)
  - Continuity: For $\|\rho-\sigma\|_1 \leq \delta$, $\left|I_c(R\rangle B)_\rho - I_c(R\rangle B)_\sigma\right| \leq 4\delta\log|R| + 2H(\delta)$. (Follows from the continuity of the conditional entropy.)
    → Leads to the continuity of the quantum channel capacity $Q(\mathcal{N})$.
  - If $B = B_1X$, where X is a classical system, i.e. $\rho_{RB} = \sum_x P_X(x)|x\rangle\langle x| \otimes \sigma^x_{RB_1}$. then $I_c(R\rangle B)_\rho = \sum_x P_X(x)\left[H\left(\sigma^x_{B_1}\right) - H\left(\sigma^x_{RB_1}\right)\right]$ ← Average over the classical labels

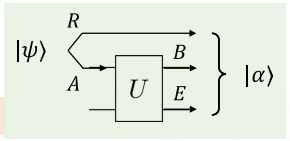


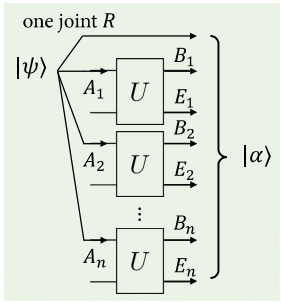
Coherent Information of a Quantum Channel

- For channel: $A \to \boxed{\mathcal{N}} \to B$ with Stinespring's extension 

- **Definition** (1-shot quantum capacity of $\mathcal{N}^{A\to B}$).
$Q^{(1)}(\mathcal{N}) \coloneqq \sup_{|\psi\rangle_{RA}} I_c(R\rangle B)_{(\mathrm{Id}\otimes U)|\psi\rangle_{RA}}$.

1-shot →





The LSD Theorem [Lloyd'97–Shor'02–Devetak'05]

For a channel $\mathcal{N}$, the *quantum channel capacity* is given by $Q(\mathcal{N}) = \sup_{n\in\mathbb{N}} \frac{1}{n} Q^{(1)}(\mathcal{N}^{\otimes n}) = \lim_{n\to\infty} \frac{1}{n} Q^{(1)}(\mathcal{N}^{\otimes n})$.

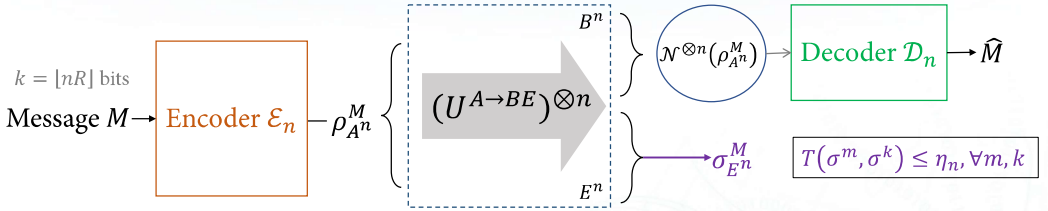
On The Quantum Channel Capacity

- The LSD Theorem
  - Achievability: Typicality and a *decoupling approach*. [Hayden–Horodecki–Winter–Yard'08]
  - Optimality: Continuity and the data-processing inequality (monotonicity of I_c).
- Just like the *Holevo quantity* (for classical communication over quantum channels), the *quantum capacity* is generally not additive → *regularization* is needed.
- *Superactiviation*. $\exists \mathcal{N}_1, \mathcal{N}_2$ with $Q(\mathcal{N}_1) = Q(\mathcal{N}_2) = 0$, but $Q(\mathcal{N}_1 \otimes \mathcal{N}_2) > 0$. [Smith–Yard'08]
- Research problems:
  - When the quantum capacity is zero? → *Degradability*
  - Tight upper and lower bounds.
    (It's even largely unknown for the depolarizing channel)
  - The so-called *positive partial transpose* (PPT) conjecture.
  - Does the strong converse rate equal quantum capacity?

Private Classical Communication

Private Information over Quantum Channels

- Given a *quantum channel* $\mathcal{N}^{A\to B}$, let $U^{A\to BE}$ be its Stinespring's dilation.



- The rate R is *achievable* if $\Pr(\widehat{M} \neq M) \le \varepsilon_n \to 0$, $\eta_n \to 0$ as $n \to \infty$.
- **Definition** (Private capacity). $P(\mathcal{N}) \coloneqq \sup\{\text{achievable } R\}$.
- We aim to transmit *private* classical data. The private capacity is the maximum rate for the classical communication such that the *complementary channel* has *vanishing* information about the transmitted classical data.

On The Private Capacity

- **Definition** (1-shot private capacity of $\mathcal{N}^{A\to B}$).
  $P^{(1)}(\mathcal{N}) \coloneqq \sup_{\rho_{XA}=\sum_x P_X(x)|x\rangle\langle x|\otimes\rho_A^x} I(X;B)_{\mathrm{Id}\otimes\mathcal{N}(\rho)} - I(X;E)_{\mathrm{Id}\otimes\mathcal{N}(\rho)}.$

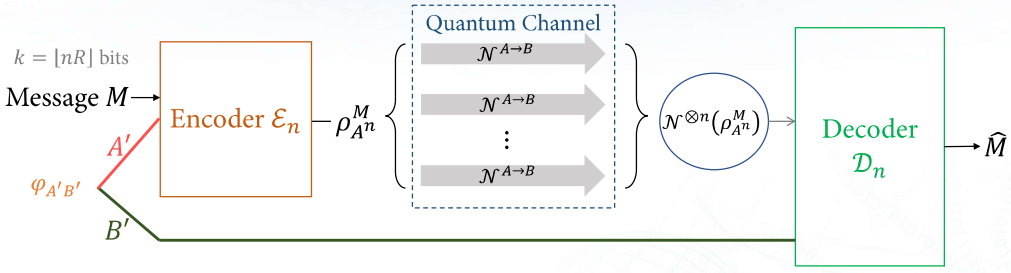
The Private Capacity [Devetak'05]

For a channel $\mathcal{N}$, the *private capacity* is given by
$P(\mathcal{N}) = \sup_{n\in\mathbb{N}} \frac{1}{n} P^{(1)}(\mathcal{N}^{\otimes n}) = \lim_{n\to\infty} \frac{1}{n} P^{(1)}(\mathcal{N}^{\otimes n}).$

- The private capacity is generally not *additive* → *regularization* is required. [Li et al.'09]
- For any channel $\mathcal{N}$, one has the relation $Q(\mathcal{N}) \le P(\mathcal{N}) \le C(\mathcal{N})$.
  (There exists channel $\mathcal{N}$ such that $Q(\mathcal{N}) = 0$, while $P(\mathcal{N}) > 0$.

Entanglement-Assisted Communication

Entanglement-Assisted Classical Communication (1/2)



- An entanglement-assisted classical (EAC) code consists of a quadruple:
  $\left\{\mathcal{M} = \{1,\cdots 2^{\lfloor nR\rfloor}\}, \varphi_{A'B'}, \left\{\mathcal{E}_m^{A'\to A^n}\right\}_{m\in\mathcal{M}}, \left\{\Pi_{B^nB'}^m\right\}_{m\in\mathcal{M}}\right\}.$
- **Definition** (EAC capacity). $C_{\mathrm{E}}(\mathcal{N}) \coloneqq \sup\{\text{achievable } R\}$.

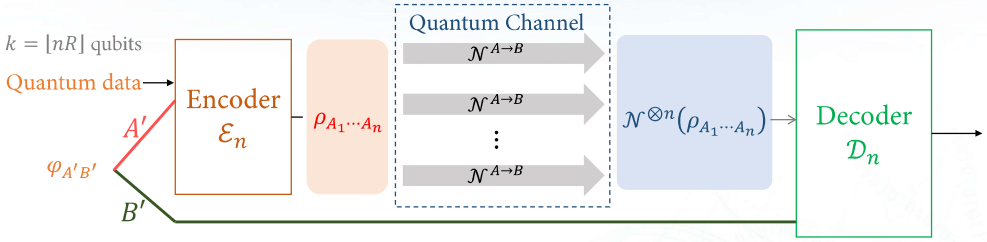
Entanglement-Assisted Classical Communication (2/2)

The EAC Capacity (BSST) [Bennet *et al.*'99]

For a channel $\mathcal{N}$, the *EAC capacity* is given by
$C_{\mathrm{E}}(\mathcal{N}) = \sup_{|\phi\rangle_{RA}} I(R;B)_{\mathrm{Id}\otimes\mathcal{N}(|\phi\rangle\langle\phi|)}.$

- The EAC capacity is *additive* → *regularization* is not needed.
- As the capacity expression in the classical case: $\sup_{\sum_x P_X(x)|x\rangle\langle x|_R\otimes|x\rangle\langle x|_X} I(R;Y)_{\mathrm{Id}\otimes\mathcal{N}(P_{RX})}.$
- If $\mathcal{N}$ is a classical channel, then $C_{\mathrm{E}}(\mathcal{N}) = C(\mathcal{N})$ → Nothing assisted.
- For near noiseless channel, $C_E/C \approx 2$ → Dense coding.
- For some very noisy channels, both vanish but $C_E/C \to \infty$ (depolarizing channels).

Entanglement-Assisted Quantum Communication (1/2)



- One can also study the entanglement-assisted quantum (EAQ) communication.
- **Theorem** (EAQ capacity). $Q_{\mathrm{E}}(\mathcal{N}) = C_E(\mathcal{N})/2$.

Entanglement-Assisted Quantum Communication (2/2)

1. $(2Q_{\mathrm{E}} \geq C_{\mathrm{E}})$:
   Suppose $\approx nC_{\mathrm{E}}$ cbits can be communicated with vanishing error $\varepsilon_n \to 0$ via EAC.

   Then, use these cbits to teleport $nC_{\mathrm{E}}/2$ qubits
   (also with vanishing error $\varepsilon_n \to 0$ in the usual sense).

2. $(2Q_{\mathrm{E}} \leq C_{\mathrm{E}})$:
   Suppose $\approx nQ_{\mathrm{E}}$ qubits can be communicated with vanishing error $\varepsilon_n \to 0$ via EAQ.

   Then, use these qubits to dense code $2nQ_{\mathrm{E}}$ cbits (with vanishing error $\varepsilon_n \to 0$).

□

Concluding Remarks

Concluding Remarks

- Other resources for assisting communication, e.g. two-way classical communication used in the quantum communication.
- Other notions of capacities, e.g. the zero-error capacities.
- We only consider the point-to-point scenario; in contrast, *network setting* include:
  - Multiple-access channels (multiple inputs – single output)
  - Broadcast channels (single input – multiple outputs)
  - Interference channels (multiple inputs – multiple outputs)
- Standard toolbox – purifying everything; the reference system servers as an "anchor".
- Other quantum information-theoretic tasks:
  - Quantum state splitting; quantum state merging; and quantum state redistribution.
  - Entanglement purification/distillation; and various resource theories.
- What is *quantum information*?

References (1/3)

- Quantum Channel Distances
  - D. Aharonov, A. Kitaev, N. Nisan, "Quantum Circuits with Mixed States," in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computation (STOC)*, 20–30, 1997.
  - H. Barnum, E. Knill, M. A. Nielsen, "On Quantum Fidelities and Channel Capacities," *IEEE Transactions on Information Theory*, 46(4):1317–1329, 2000.
  - D. Kretschmann, and R. F. Werner, "*Tema con variazioni*: quantum channel capacity," *New Journal of Physics*, 6(26), 2004.
  - L. Gyonhyosi, S. Imre, H. V. Nguyen, "A Survey on Quantum Channel Capacities," *IEEE Communications Surveys & Tutorials*, 20(2), 2018.
- Quantum Channel Capacities
  - S. Lloyd, "Capacity of the noisy quantum channel," *Physical Review A*, 55(3):1613–1622, 1997.
  - P. Shor, "The quantum channel capacity and coherent information," *Lecture Notes, MSRI Workshop on Quantum Computation*, 2002. (Unpublished)
  - I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Transactions on Information Theory*, 51(1): 44–55, 2005.

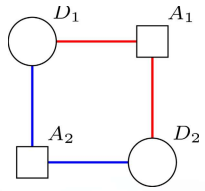
References (2/3)

- The Decoupling Approach to the LSD Theorem
  - P. Hayden, M. Horodecki, A. Winter, J. Yard, "A decoupling approach to the quantum capacity," *Open Systems & Information Dynamics*, 15(01):7–19, 2008.
  - F. Dupuis, *The decoupling approach to quantum information theory*, Ph.D. Thesis (Université de Montréal), 2009.
  - F. Dupuis, M. Berta, J. Wullschleger, R. Renner, "One-shot decoupling," *Communications in Mathematical Physics*, 328(251), 2014.
- Superactivation of quantum channels
  - G. Smith, J. Yard, "Quantum Communication With Zero-Capacity Channels," *Science*, 321, 1812–1815, 2008.
- Quantum conditional entropy can be negative
  - M. Horodecki, J. Oppenheim, A. Winter, "Partial quantum information," *Nature*, 436, 673–676, 2005.

References (3/3)

- Private Communication
  - I. Devetak, "The Private Classical Capacity and Quantum Capacity of a Quantum Channel" *IEEE Transactions on Information Theory*, 51(1):4–55, 2005.
  - K. Li, A. Winter, X. Zou, and G. Guo, "Private Capacity of Quantum Channels is Not Additive," *Physical Review Letters*, 103, 120501, 2009.
- Entanglement-Assisted Capacities
  - A. S. Holevo, "On entanglement-assisted classical capacity," *Journal of Mathematical Physics*, 43(9):4326–4333, 2002.
  - C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels," *Physical Review Letters*, 83(15):3081–3084, 1999.
  - C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Transactions on Information Theory*, 48:2637–2655, 2002.

Quantum Information and Computation
Quantum Error Correction



Hao-Chung Cheng (鄭皓中)

haochung@ntu.edu.tw

Department of Electrical Engineering

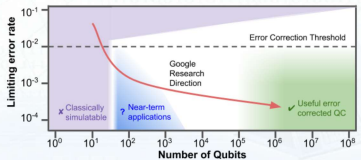
National Taiwan University

June 2, 2021

Motivation

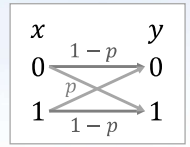
Motivation

- As we have seen in classical/quantum communications, channel coding (which can be implemented by error-correcting codes) is crucial to achieving channel capacity.
- Other than that, error-correcting codes are vital in quantum computation as well.
  - Quantum states/gates are fragile; → prone to interact with environment.
  - Currently the best known fidelity among physical implementations are single-qubit gate: error rate $\sim 0.1\%$; two-qubit gate: error rate $\sim 5\%$.
- Typically, a practical quantum algorithm using n qubits and M elementary steps has $nM \gg 10^{10}$, which means that the *logical error* for each logical operation should be much less than 10^{-10}.
- Near term: advancing physical implementations; Mid-to-long term: Quantum error correction.



Classical and Quantum Error Correction

Classical Error-Correcting Codes

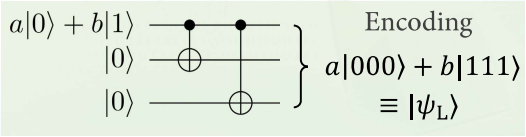


- Consider a bit-flip (binary symmetric) channel with probability p.
- Probability of error $\varepsilon := \Pr(Y \neq X) = p$.
- Classical repetition code encoding: $\begin{cases} 0 \mapsto 000 \\ 1 \mapsto 111 \end{cases}$ ← codewords
- Decoding via majority vote: $000,001,010,100 \mapsto 0$; $011,101,110,111 \mapsto 1$. This corrects **single** bit-error.
- The decoding error happens when two or three bit-error occur: $\varepsilon = 3p^2(1-p) + p^3 = 3p^2 - 2p^3 < p$ for $p \in (0, 1/2)$.
- Reason: *Hamming distance* $d_{\mathrm{H}}(000,111) = 3$.
- A classical binary code can correct up to t single bit errors $\Leftrightarrow \min d_{\mathrm{H}} \geq 2t + 1$.
- Notation: an $[n, k, d]$ code means k-bit to n-bit with $\min d_{\mathrm{H}} \geq d$.

Satisfies triangle inequality

3-Qubit Code Protecting from One Bit Flip (1/3)

- Quantum repetition code $|\psi\rangle \coloneqq a|0\rangle + b|1\rangle \mapsto |\psi\rangle \otimes |\psi\rangle \otimes |\psi\rangle$? ✖
  - No-cloning theorem.
  - Errors are continuous – there are infinitely many noisy quantum evolutions.
  - Measurement destroys quantum information – once the state collapses, no recovery.
- *Quantum bit-flip channel:* $\mathcal{N}(\rho) = (1-p)\rho + pX\rho X$; e.g. $|b\rangle \mapsto |b \oplus 1\rangle, b \in \{0,1\}$. with probability p
- 3-qubit encoding:
  - $\begin{cases} |0\rangle \mapsto |0_L\rangle \coloneqq |000\rangle \\ |1\rangle \mapsto |1_L\rangle \coloneqq |111\rangle \end{cases}$ ← Each qubit within the 3-qubit codewords are called the *physical qubit*, while the whole 3-qubit state means protecting the *logical qubit* $|0/1\rangle$
  - Not violating the no-cloning theorem: $|\psi\rangle \mapsto a|0_L\rangle + b|1_L\rangle \neq (a|0\rangle + b|1\rangle)^{\otimes 3}$.
- *Idea*: Measurement of parities (mod 2) without measurement of qubits.



3-Qubit Code Protecting from One Bit Flip (2/3)

- To measure with operators $Z_1Z_0 \equiv I_2 \otimes Z_1 \otimes Z_0$ and Z_2Z_1 (numbering $q_2q_1q_0$).
  - An Hermitian operator corresponds to an *observable*, with measurement results being the eigenvalues of this operator.
  - Consider Z_1Z_0: $(Z_1Z_0)^2 = I$, so the eigenvalues are ± 1.
  - $+1$: even parity → the same qubit states (either $|00\rangle$ or $|11\rangle$ or their superposition).
  - -1: odd parity → different qubit states (either $|01\rangle$ or $|10\rangle$ or their superposition).
- Assume a bit flip X_0 occurs:

$$X_0|\psi_L\rangle = a|001\rangle + b|110\rangle$$

$Z_1Z_0: -$

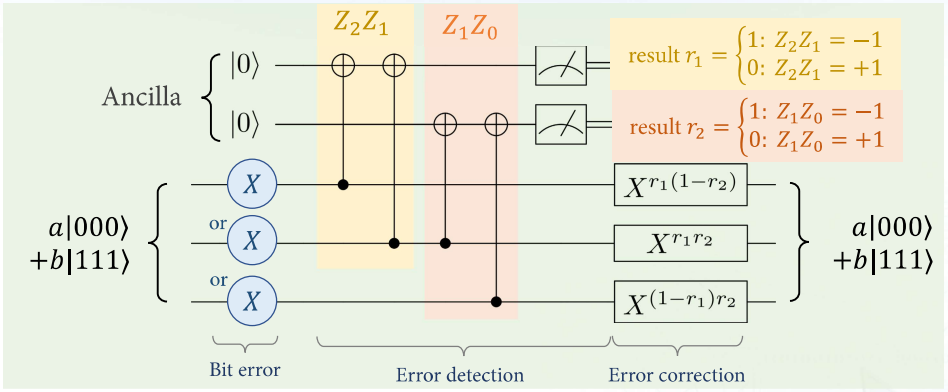
$Z_2Z_1: +$

Parity check	X_0	X_1	X_2	I
Z_2Z_1	+	−	−	+
Z_1Z_0	−	−	+	+

⇒ Able to distinguish the four scenarios.

⇒ Don't know a & b, hence not destroy the state.

3-Qubit Code Protecting from One Bit Flip (3/3)



- Error in the upper qubit ⇒ upper ancilla 1 ⇒ upper qubit corrected;
- Error in the middle qubit ⇒ both ancillas 1 ⇒ corrected qubit corrected;
- Error in the lower qubit ⇒ lower ancilla 1 ⇒ lower qubit corrected; no error ⇒ no correction.

The error probability is $3p^2 - 2p^3$

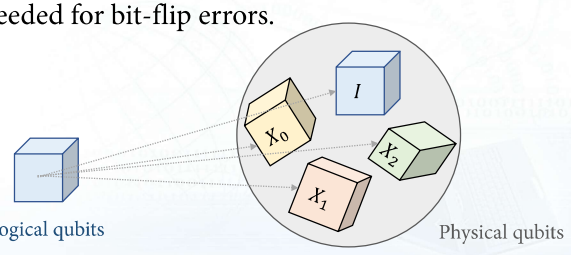
Analysis of Space Dimensions for 3-Qubit Code

↓ code space: span{$|000\rangle, |111\rangle$}

- Valid (uncorrupted) codewords live in a 2-dim subspace of 2^3-dim Hilbert space.
- After corruption of X_0, the states move to a different (orthogonal) 2-dim subspace. Similarly, the state moves to different subspaces after corruption due to X_1 and X_2.
- There are 4 orthogonal 2-dim subspace (correct and 3 bit-errors), which all fit into the 2^3-dim Hilbert space. This is why we can distinguish errors and correct them.
- In general, an n-qubit code, protecting against 1-qubit bit-flip requires $2^n \geq 2\underbrace{(1+n)}$, which means $n \geq 3$ is needed for bit-flip errors.

$(1+n)$ 2-dim subspaces for n possible errors

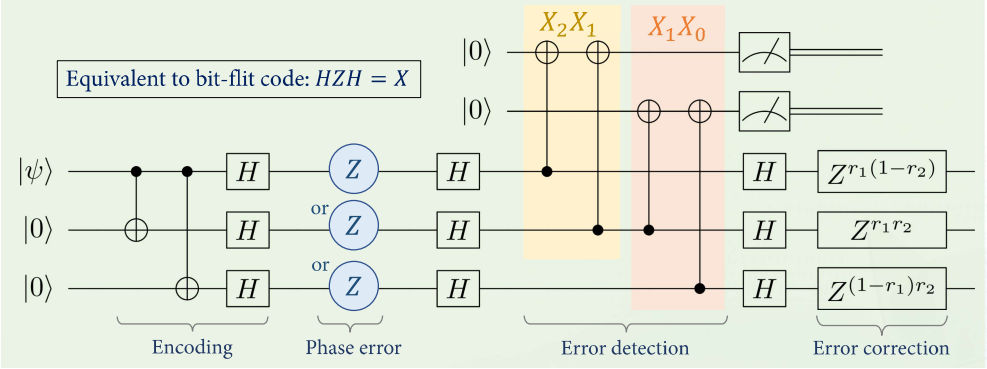
- For general errors (X, Z, and Y): $2^n \geq 2(1+3n) \rightarrow n \geq 5$.



3-Qubit Code Protecting from One Phase Flip

- Encoding: $\begin{cases} |0\rangle \mapsto |0_L\rangle \coloneqq \frac{1}{\sqrt{8}}(|0\rangle + |1\rangle)(|0\rangle + |1\rangle)(|0\rangle + |1\rangle) \\ |1\rangle \mapsto |1_L\rangle \coloneqq \frac{1}{\sqrt{8}}(|0\rangle - |1\rangle)(|0\rangle - |1\rangle)(|0\rangle - |1\rangle) \end{cases}$

If one qubit flips the phase, $(|0\rangle + |1\rangle) \leftrightarrow (|0\rangle - |1\rangle)$, possible to find and correct it.



9-Qubit Shor's [9,1,3] Code (1/2)

Concatenation of the two previous codes

- Encoding: $\begin{cases} |0\rangle \mapsto |0_L\rangle \coloneqq \frac{1}{\sqrt{8}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \\ |1\rangle \mapsto |1_L\rangle \coloneqq \frac{1}{\sqrt{8}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \end{cases}$

One encoding deals with X-error, the other one with Z-errors, while Y-errors are taken care of automatically since $Y = -iZX \equiv ZX$.

- For X_7 error, then this changes parities within the first of the 3-qubit block:
  $(|010\rangle + |101\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)$
  $(|010\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)$
- For Z_6, Z_7 or Z_8 error, then this changes parities within the sign $(+ \leftrightarrow -)$ in the block:
  $(|000\rangle - |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)$
  $(|000\rangle + |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)$
- For Y error (the combination of the above errors), then both.

9-Qubit Shor's [9,1,3] Code (2/2)

- Error detection and correction:
  - Measure parities within the three 3-qubit blocks: Z_8Z_7, Z_7Z_6, Z_5Z_4, Z_4Z_3, Z_2Z_1, Z_1Z_0.
    If X-error is detected in the i-th qubit, correct by X_i.
    → If bit flips happened in the different blocks, they all can be corrected.
  - Measure parities of phases the 3-qubit blocks: $X_8X_7X_6X_5X_4X_3$, $X_5X_4X_3X_2X_1X_0$.
    If Z-error is detected in the j-th 3-qubit block, correct by Z to **any** qubit in this block.
    → $X_2X_1X_0(|000\rangle \pm |111\rangle) = \pm(|000\rangle \pm |111\rangle)$ → $X_2X_1X_0(|001\rangle \pm |110\rangle) = \pm(|001\rangle \pm |110\rangle)$
  - The above two steps are *interchangeable* ∵ those operators (called *stabilizers* ***commute***).

E.g. for Y_7 error: $\frac{a}{\sqrt{8}}(|010\rangle - |101\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)$
$+\frac{b}{\sqrt{8}}(|010\rangle + |101\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)$

$X_8X_7X_6X_5X_4X_3: -$

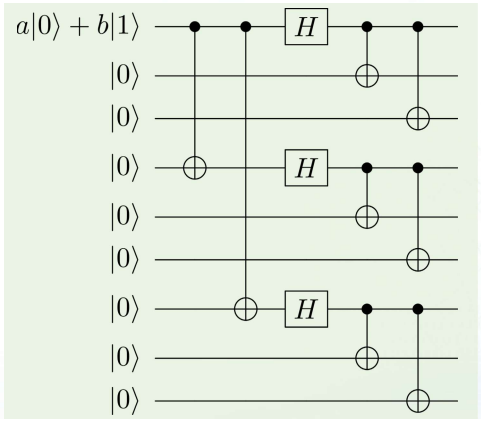
$X_5X_4X_3X_2X_1X_0: +$

Is Shor's [9,1,3] code optimal?

- The 8 measured operators: Z_8Z_7, Z_7Z_6, Z_5Z_4, Z_4Z_3, Z_2Z_1, Z_1Z_0,
  $X_8X_7X_6X_5X_4X_3$, $X_5X_4X_3X_2X_1X_0$.
  → There are $2^8 = 256$ possible results.
- A 9-qubit Hilbert space (512-dim) can hold 256 many 2-dim (qubit) spaces.
- However, there are only $1 + 3 \times 9 = 28$ scenario of errors.
  (Even less: $28 - 6 = 22$, because Z-errors may lead to the same result.)
  ← *degenerate quantum code*
- Shor's [9,1,3] code is not optimal. → There is a [5,1,3] code.

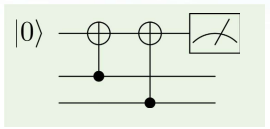
Shor's [9,1,3] Code: Encoding

- Encoding: $a|0\rangle + b|1\rangle \mapsto \frac{a}{\sqrt{8}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)$
  $+\frac{b}{\sqrt{8}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)$



Shor's [9,1,3] Code: Syndrome Extraction (1/3)

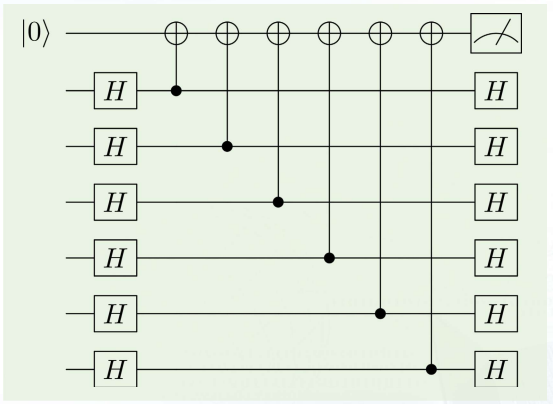
- Z_iZ_j as before:



↑ need 6 ancilla qubits for Z_iZ_j

→ Overall need 6 + 2 = 8 ancillas

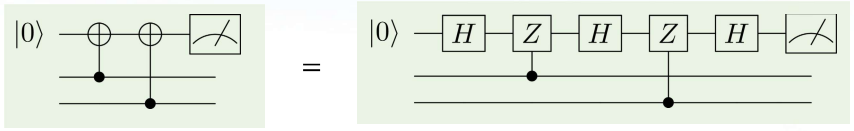
- Realization of $X_5X_4X_3X_2X_1X_0$:

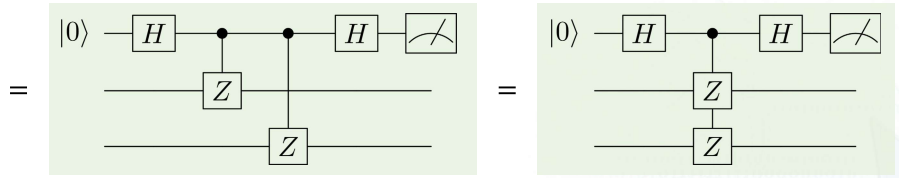


↑ need 2 ancillas

Shor's [9,1,3] Code: Syndrome Extraction (2/3)

- Z_iZ_j:

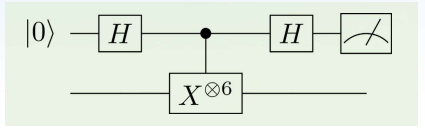




In this way it is clear which operator is measured.

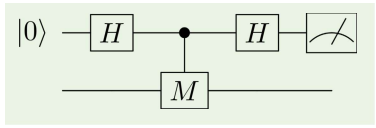
Shor's [9,1,3] Code: Syndrome Extraction (3/3)

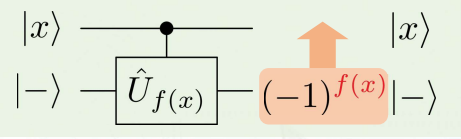
- Similarly, for $X_5X_4X_3X_2X_1X_0$:



Phase kickback via U_f: $|x\rangle|-\rangle \mapsto (-1)^{f(x)}|x\rangle|-\rangle, \forall x \in \{0,1\}^n$

- General way for multi-qubit operator M has eigenvalues ± 1.





$$|0\rangle \mapsto \frac{|0\rangle + |1\rangle}{\sqrt{2}} \mapsto \begin{cases} \frac{|0\rangle+|1\rangle}{\sqrt{2}} \text{ if eigenvalue of } M \text{ is } +1 \\ \frac{|0\rangle-|1\rangle}{\sqrt{2}} \text{ if eigenvalue of } M \text{ is } -1 \end{cases} \mapsto \begin{cases} |0\rangle \\ |1\rangle \end{cases}$$

Discretization of Errors (1/2)

- Why protecting from only 3 errors (X, Y, Z) is sufficient?
- *Intuition*: for a small perturbation of a unitary evolution, measurement converts small continuous errors into big errors (X, Y, Z), otherwise restores initial state.
- 1-qubit Stinespring dilation: $\begin{cases} |e\rangle|0\rangle \mapsto |e_0\rangle|0\rangle + |e_1\rangle|1\rangle \\ |e\rangle|1\rangle \mapsto |e_2\rangle|0\rangle + |e_3\rangle|1\rangle \end{cases}$

  $|e\rangle$ is initial state of environment
  $|e_i\rangle$ are not normalized

$$\Rightarrow |e\rangle|\psi\rangle \mapsto (|e_0\rangle + |e_1\rangle X)\underbrace{|0\rangle\langle 0|}_{\frac{I+Z}{2}}|\psi\rangle + (|e_2\rangle X + |e_3\rangle)\underbrace{|1\rangle\langle 1|}_{\frac{I-Z}{2}}|\psi\rangle$$

$$= \frac{|e_0\rangle + |e_3\rangle}{2}|\psi\rangle + \frac{|e_1\rangle + |e_2\rangle}{2}X|\psi\rangle + \frac{|e_0\rangle - |e_3\rangle}{2}Z|\psi\rangle + \frac{|e_1\rangle - |e_2\rangle}{2}XZ|\psi\rangle$$

Discretization of Errors (2/2)

$\begin{cases} |e\rangle|0\rangle \mapsto |e_0\rangle|0\rangle + |e_1\rangle|1\rangle \\ |e\rangle|1\rangle \mapsto |e_2\rangle|0\rangle + |e_3\rangle|1\rangle \end{cases}$

- For *gradual process*, $\||e_0\rangle\| \approx \||e_3\rangle\| \approx 1$, $\||e_1\rangle\|, \||e_2\rangle\| \ll 1$.

$$|e\rangle|\psi\rangle \mapsto \underbrace{\frac{|e_0\rangle + |e_3\rangle}{2}}_{\text{large}}|\psi\rangle + \underbrace{\frac{|e_1\rangle + |e_2\rangle}{2}}_{\text{small}}X|\psi\rangle + \underbrace{\frac{|e_0\rangle - |e_3\rangle}{2}}_{\text{small}}Z|\psi\rangle + \underbrace{\frac{|e_1\rangle - |e_2\rangle}{2}}_{\text{small}}XZ|\psi\rangle$$

- E.g. depolarizing channel: $\mathcal{N}(\rho) \coloneqq (1-p)\rho + \frac{p}{3}X\rho X + \frac{p}{3}Y\rho Y + \frac{p}{3}Z\rho Z$.
- More challenging in the multiple qubit case → usually consider independent errors.

Discretization of Errors [§10.3.1, N&C]

If a quantum code can correct noise with Krauss operators $\{E_i\}_i$. Then it can correct noise with linear combinations of $\{E_i\}_i$, i.e. $\{\sum_i m_{ij}E_i\}_j$.

Stabilizer Codes

- To distinguish subspaces, we measure a set of *commuting* operators M_i.
  → They are constructed as tensor products of Pauli operators.
  → Each of them $M_i^2 = I$, and hence eigenvalues are ± 1.
- Subspaces corresponding to different measurement results are orthogonal.
  → Measured operators (and the subspaces) should be able to diagnose errors.
- Since M_i are *commuting*, we can measure them simultaneously or in any sequence.
- Measurement of any such operator M_i projects a state in the Hilbert space into one of two subspaces, corresponding eigenvalues $\lambda = +1$ and $\lambda = -1$.
  - $\lambda = +1$ (usually measurement result 0) ↔ projector $\Pi_0^{M_i} = (I + M_i)/2$.
  - $\lambda = -1$ (usually measurement result 1) ↔ projector $\Pi_1^{M_i} = (I - M_i)/2$.
- E.g. If $M_i|\psi\rangle = +1|\psi\rangle$, then $\Pi_0^{M_i}|\psi\rangle = |\psi\rangle$, while $\Pi_1^{M_i}|\psi\rangle = 0$.

[5,1,3] Code (1/2)

- Setup: $2^5 = 32$-dim Hilbert space; $3 \times 5 = 15$ possible errors.
  All dimensions should be used since $2 \times (1 + 15) = 32$.
- Need to distinguish $16 = 2^4$ scenarios. Hence would require 4 operators M_i.

$$\begin{aligned} M_0 &= Z_4X_3X_2Z_1 \\ M_1 &= X_4X_3Z_2Z_0 \\ M_2 &= X_4Z_3Z_1X_0 \\ M_3 &= Z_4Z_2X_1X_0 \end{aligned}$$

- $M_i^2 = I$ for all of them because $X_k^2 = Z_k^2 = I$
- M_i commute because $X_kZ_k = -Z_kX_k$ (anticommute)

- Encoding:

$$\begin{cases} |0\rangle \mapsto |0_L\rangle \coloneqq \frac{1}{4}(I + M_0)(I + M_1)(I + M_2)(I + M_3)|00000\rangle \\ |1\rangle \mapsto |1_L\rangle \coloneqq \frac{1}{4}(I + M_0)(I + M_1)(I + M_2)(I + M_3)|11111\rangle \end{cases}.$$

[5,1,3] Code (2/2)

- Important properties
  - $|0_L\rangle$ and $|1_L\rangle$ are orthogonal to each other because $|0_L\rangle$ is a superposition of terms with odd number of 0s and even numbers of 1s; opposite for $|1_L\rangle$.
  - Can check $|0_L\rangle$ and $|1_L\rangle$ are normalized.
  - $|0_L\rangle$ and $|1_L\rangle$ are eigenstates of all M_i with eigenvalues $+1$; ∵ M_i commute and → measurement of M_i does not disturb $|\psi\rangle$ (its stabilizer).

	$X_0Y_0Z_0$	$X_1Y_1Z_1$	$X_2Y_2Z_2$	$X_3Y_3Z_3$	$X_4Y_4Z_4$	I
$M_0 = Z_4X_3X_2Z_1$	$++-$	$--+$	$+--$	$+--$	$--+$	$+$
$M_1 = X_4X_3Z_2Z_0$	$-++$	$++-$	$--+$	$+--$	$+--$	$+$
$M_2 = X_4Z_3Z_1X_0$	$+--$	$--+$	$++-$	$--+$	$+--$	$+$
$M_3 = Z_4Z_2X_1X_0$	$+--$	$+--$	$--+$	$++-$	$--+$	$+$

Steane's [7,1,3] Code

← came from *classical Hamming code*

- Although [7,1,3] code is longer than [5,1,3], but it is *fault-tolerant* for quantum computing – several important logic operations can be done *without decoding*.
- Need to distinguish $1 + 3 \times 7 = 22$ scenarios need at least 5 operators M_i ($2^5 = 32$).

$$M_0 = X_6X_5X_4X_0$$
$$M_1 = X_6X_5X_3X_1$$
$$M_2 = X_6X_4X_3X_2$$

$$N_0 = Z_6Z_5Z_4Z_0$$
$$N_1 = Z_6Z_5Z_3Z_1$$
$$N_2 = Z_6Z_4Z_3Z_2$$

6	5	4	3	2	1	0
*	*	*				*
*	*		*		*	
*		*	*	*		

- Encoding:

$$\begin{cases} |0\rangle \mapsto |0_L\rangle := \frac{1}{\sqrt{8}}(I + M_0)(I + M_1)(I + M_2)\,|0000000\rangle \\ |1\rangle \mapsto |1_L\rangle := \frac{1}{\sqrt{8}}(I + M_0)(I + M_1)(I + M_2)\,|1111111\rangle \end{cases}.$$

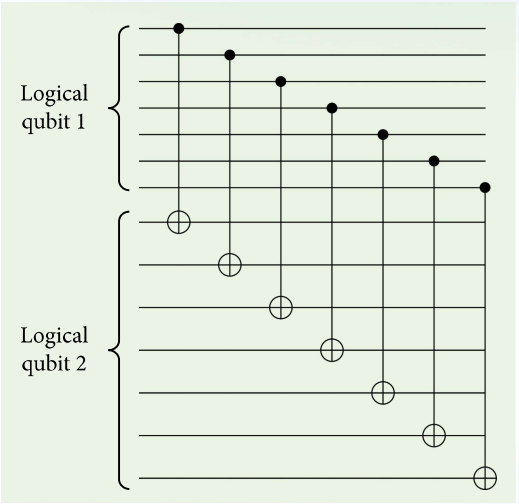
Steane's [7,1,3] Code: Transversal Gates (Bitwise)

- X-operation on logical qubit ↔ $X^{\otimes 7}$ on physical qubits.
  - ∵ all X_k commute with M_i (use only X); hence $X^{\otimes 7}|0_L\rangle = |1_L\rangle$, $X^{\otimes 7}|1_L\rangle = |0_L\rangle$.
- Z-operation on logical qubit ↔ $Z^{\otimes 7}$ on physical qubits.
  - $Z^{\otimes 7}$ commute with M_i (4 anticommuting pairs).
  - $Z^{\otimes 7}|0^{\otimes 7}\rangle = |0^{\otimes 7}\rangle$, $Z^{\otimes 7}|1^{\otimes 7}\rangle = -|1^{\otimes 7}\rangle$. Hence $Z^{\otimes 7}|0_L\rangle = |0_L\rangle$, $Z^{\otimes 7}|1_L\rangle = -|1_L\rangle$.
- Y-operation on logical qubit ↔ $Y^{\otimes 7}$ on physical qubits (∵ $Y \equiv ZX$).

↑ Actually, the same for the [5,1,3] code: logic X, Y, and Z are easy to implement.

- The same for Hadamard: $H_L = H^{\otimes 7}$. ← Not possible in [5,1,3]
- Most importantly, CNOT can be implemented qubit-by-qubit.

Steane's [7,1,3] Code: Transversal CNOT



$$|0_L\rangle = \frac{1}{\sqrt{8}}(I + M_0)(I + M_1)(I + M_2)\,|0000000\rangle$$
$$|1_L\rangle = \frac{1}{\sqrt{8}}(I + M_0)(I + M_1)(I + M_2)\,|1111111\rangle$$

$$M_0 = X_6X_5X_4X_0$$
$$M_1 = X_6X_5X_3X_1$$
$$M_2 = X_6X_4X_3X_2$$

Key observation:
$|0_L\rangle$ and $|1_L\rangle$ are invariant under M_i

Fault-Tolerant Quantum Computation

- *Fault tolerant*:
  Failure of a component leads to at most one error in each encoded block.
- Spirit:
  - Do logic operations without decoding.
  - Correct faulty gates (if only one works incorrectly).
  - Singe-qubit errors do not become multi-qubit errors.
- Method:
  - Production of ancillas, measurements, gates, “wires”, etc.
- Recently, much attention has turned to topological codes
  - Toric code (Kitaev)
  - Surface codes
  - Color codes, etc.

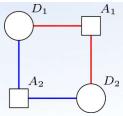
Surface Codes

Milestone

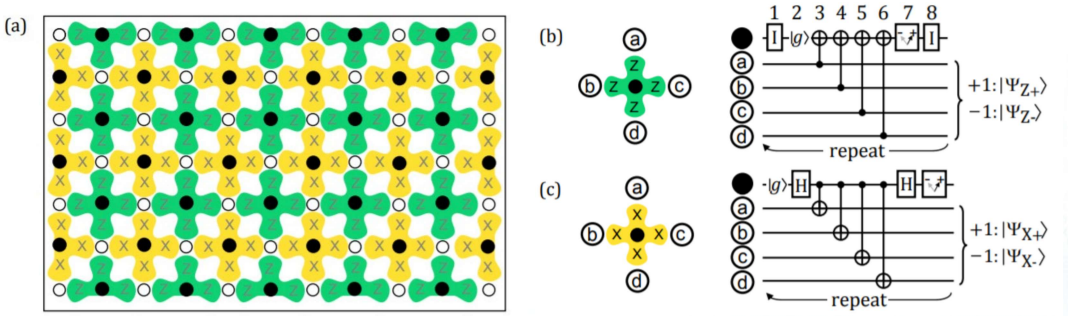
Surface codes: Towards practical large-scale quantum computation

Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland
Phys. Rev. A **86**, 032324 – Published 18 September 2012

An article within the collection: *Physical Review A 50th Anniversary Milestones*



- General design principle behind topological codes is that the code is built up by ‘patching’ together repeated elements. Such a modular approach ensures that the surface code can be straight-forwardly scaled in size whilst ensuring stabilizer commutativity.



Threshold Theorem

Threshold Theorem [§10.6.1, N&C]

A quantum circuit containing $M(n)$ may be simulated with probability of error at most ε using $O\big(\text{poly log}(M(n)/\varepsilon)M(n)\big)$ gates on hardware whose components fail with probability at most p, provided that $p < p_{\text{th}}$, and given reasonable assumptions about the noise in the underlying hardware.

- Problem: The threshold p_{th} is usually low because we need many qubits for error correction, and this increases the error probability.
- Currently, $p_{\text{th}} \sim (10^{-6}, 10^{-2})$ for different codes ($p_{\text{th}} \sim 10^{-2}$ for surface codes). Often, it is safe to say $p_{\text{th}} \sim 10^{-4}$.
- This makes quantum computation potentially possible.

Concluding Remarks

Concluding Remarks

- Inspired from classical linear code → the Calderbank–Shor–Steane quantum codes.
- Method behind the scene: the *stabilizer formalism* (via Daniel Gottesman).
- Note that both encoding and decoding may introduce new errors.
  → threshold theorem, fault tolerant, and topological codes.
- Currently, one might require huge amounts of physical qubits (say ~ 1000) for protecting a single logical qubit (might depend on the physical platform).
- Other topics: quantum LDPC codes (iterative decoding).
- Modeling the noise is a practical and theoretical problem.
- Quantum error correction is widely considered as a crucial part in realizing practical quantum computation → a very active research area in the near term.

References (1/3)

- Early works of the CSS codes
  - P. Shor, "Scheme for reducing decoherence in quantum computer memory," *Physical Review A*, 52:2496, 1995.
  - A. Steane, "Error correcting codes in quantum theory," *Physical Review Letters*, 77:793, 1996.
  - A. R. Calderbank, and P. Shor, Good quantum error-correcting codes exist, *Physical Review A*, 54:1098, 1996.
  - A. Steane, "Simple quantum error-correcting codes," *Physical Review A*, 54:4741, 1996.
  - D. Gottesman, "A class of quantum error-correcting codes saturating the quantum Hamming bound," *Physical Review A*, 54:1862.
- Early works of the topological codes
  - A. Y. Kitaev, "Quantum computations: algorithms and error correction," Russian Mathematical Surveys, 52(1191), 1997.
  - S. Bravyi, A. Kitaev, "Quantum codes on a lattice with boundary," arXiv:quant-ph/9811052, 1998.
  - E. Dennis, A. Kitaev, A. Landahl, J. Preskill, Topological quantum memory, *Journal of Mathematical Physics,* 43:4452–4505, 2002.

References (2/3)

- Review papers
  - D. Gottesman, *Stabilizer codes and quantum error correction*, Ph.D. Thesis (Caltech), arXiv:quant-ph/9705052.
  - D. Gottesman, "An introduction to quantum error correction," arXiv:quant-ph/0004072.
  - D. Gottesman, "An introduction to quantum error correction and fault-tolerant quantum computation," arXiv:0904.2557.
  - J. Preskill, Reliable quantum computers, *Proceedings of Royal Society Lond A*, 454(1969):385–410,1998.
  - Simon J. Devitt, Kae Nemoto, William J. Munro, "Quantum Error Correction for Beginners," *Reports on Progress in Physics,* 76(076001), 2013.
  - D. A. Lidar, and T. A. Brun, *Quantum Error Correction*, Cambridge University Press, 2013.
  - B. M. Terhal, "Quantum Error Correction for Quantum Memories," Review of Modern Physics, 87(307), 2015.
  - J. Roffe, "Quantum Error Correction: An Introductory Guide," *Contemporary Physics*, 226–245, 2019.
- The threshold theorem
  - E. Knill, R. Laflamme, and W. H. Zurek, Resilient quantum computation. *Science*, 279(5349):342–345, 1998.

References (3/3)

- Surface codes
  - A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, "Surface codes: Towards practical large-scale quantum computation," *Physical Review A*, 86:032324, 2012.
- Fault-tolerant quantum computation
  - P. Shor, "*Fault-tolerant quantum computation*" in *Annual Symposium on Fundamentals of Computer Science (FOCS)*, 56–65, 1996.
  - D. Aharonov, and M. Ben-Or, "Fault-tolerant quantum computation with constant error," in Proceedings of 29th Annual ACM Symposium on Theory of Computing, p. 46, 1997.
  - A. Y. Kitaev, Fault-tolerant quantum computation by Anyons, *Annals of Physics*, 303(1):2–30, 2003.
  - D. Gottesman, "A theory of fault-tolerant quantum computation," *Physical Review A*, 57:127, 1998.
  - J. Preskill, "Fault-tolerant quantum computation," arXiv:quant-ph/9712048.
  - D. Gottesman, "Fault-tolerant quantum computation with local gates," *Journal of Modern Optics*, 47(333), 2000.
  - E. T. Campbell, B. M. Terhal, C. Vuillot, "Roads towards fault-tolerant universal quantum computation," *Nature*, 549:172–179, 2017.