Quantum Information Processing Theory

A collection of notes on major definitions, results, and commentary based on the corresponding course at Illinois, as instructed by Chitambar

Lecture Notes By

Dheeran E. Wiggins

Department of Electrical and Computer Engineering
University of Illinois Urbana-Champaign

Disclaimer
The lecture notes in this document were based on Quantum Information Processing Theory [498], as instructed by Eric Chitambar [Department of Electrical and Computer Engineering] in the Fall semester of 2024 [FA24] at the University of Illinois Urbana-Champaign. All non-textbook approaches, exercises, and comments are adapted from Chitambar's lectures.

Textbook
Many of the exercises and presentations were selected from *Quantum Computation and Quantum Information*, by Michael A. Nielsen and Isaac L. Chuang.

Author Information
Dheeran E. Wiggins is, at the time of writing [Fall, 2024], a second-year student at Illinois studying mathematics. All typesetting and verbiage are his own.

dheeran2@illinois.edu

Just trace backwards in time along the wire.

– Eric Chitambar

Contents

On the Axioms of Quantum Information

State Space Axiom 1

At the root of quantum mechanics lay philosophical questions about the nature of science–whether *realism* or *instrumentalism* is adopted plays a substantial role in how we *interpret* the theories, and the perceived importance, of quantum information.

1.1 Aside on the Philosophy of Information

Quantum mechanics arose to explain experimental phenomena discovered in the early twentieth century.

Nature of Quantum Mechanics

Every quantum experiment consists of three stages:

(i) *System Preparation*: setting the initial state of our system.
(ii) *System Evolution*: dynamically evolving the system.
(iii) *System Measurement*: coupling with some measurement device to *observe* an outcome.[1]

Still, quantum mechanics poses some major challenges to scientific realism. Quantum observables are altered by measurement, either of itself or of an "associated" observable. We often represent quantum states via mathematical tools like "pure state" vectors $|\psi\rangle$ and "mixed state" density matrices ρ. Whether these representations are *real* is, once again, a question of scientific philosophy. In either case, the density matrix ρ allows for calculated measurement outcomes.

Classical Versus Quantum Information

Generally speaking, classical information processing handles the storage and manipulation of long strings of bits. The fundamental piece of technology here is the *transistor*.[2] No matter the device, in principle, any calculation can be performed on *any* classical machine, simply mapping the sets of *bits* between the systems. Quantum information diverges, introducing the notion of a *qubit*. What is a qubit? Well, a qubit, in analog with a bit, is a two-level quantum system, using the quantum properties of our system to superimpose states.[3]

Example 1.1.1 There are a few common qubits used, in practice.

(a) *Photon Polarization*: The quantum particles associated with the electromagnetic field are called photons. Each photon has a property known as *polarization*, which is its direction of oscillation in space. We could describe orthogonal sates via $|0\rangle, |1\rangle$, and diagonal polarizations via $|+\rangle, |-\rangle$.

[1] Notably, quantum mechanics tells us how to mathematically compute probabilites of our experimental system.

[2] These transistors can be miniaturized, to an extent, and *billions* can be put in a processor.

[3] Given states $|0\rangle, |1\rangle$, we could take the $\mathbb{C}$-linear combination

$$|+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$
$$|-\rangle := \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

(b) *Spin Systems*: Quantum systems have a physical property known as *spin*, and an associated *spin number* which is a half-integer $n/2 \in \mathbb{Z}_+$. The "spin up" and "spin down" states can encode $|0\rangle$ and $|1\rangle$.

(c) *Atomic Systems*: We can construct qubits using any "gapped" quantum system in which there are two energy levels $|0\rangle, |1\rangle$ which are sufficiently separated.

(d) *Superconducting Systems*: Built at low temperature, superconducting elements called Josephson junctions can encode $|0\rangle, |1\rangle$.

The bit and qubit values are also called *logical values*. These are abstractions of their physical counterparts, where there is a correspondence between logical and physical transformations.

1.2 State Space Axiom

The starting point in formulating quantum mechanics is the *state space axiom*.[4]

Definition 1.2.1 (State Space Axiom) *Given a quantum system,*

(i) the system is represented by a complex Hilbert space $\mathscr{H}$, known as the state space.

(ii) states of the system are represented by trace-one, positive (semi-definite) operators acting on $\mathscr{H}$ called density operators.

That is to say, we have a correspondence,

$$\left\{\begin{matrix}\text{states in the}\\ \text{physical system}\end{matrix}\right\} \xleftrightarrow[\sim]{\text{bijection}} \left\{\begin{matrix}\text{density operators}\\ \text{in } \mathscr{H}\end{matrix}\right\}.$$

Remark 1.2.1 The set of all density operators is denoted $\mathscr{D}(\mathscr{H})$.

Proposition 1.2.1 *The dimension of a quantum system corresponds to the number of distinct observable outcomes it can generate.*[5]

Example 1.2.1 Let $\mathscr{H} := \mathbb{C}^3$. For what values of a, b, c is ρ a valid density matrix?[6]

$$\rho := \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

Solution. Well, the eigenvalues are a, b, c, so $a, b, c \geq 0$. We also need $\operatorname{tr}\rho = a + b + c = 1$. □

Remark 1.2.2 Notably, a, b, c, as above, form a probability distribution.

[4] This gives us a rigorous mathematical framework to work in.

[5] With d distinct outcomes observed, we model the system with a Hilbert space of $\dim \mathscr{H} = d$. This is determined by experiment.

[6] Recall that a positive operator $\rho \geq \mathbf{0}$ is when the eigenvalues are nonnegative and $\rho^\dagger = \rho$, Hermitian.

Example 1.2.2 Let $\mathcal{H} := \mathbb{C}^3$. For what values of a, b, c is ρ a valid density matrix?

$$\rho := \begin{pmatrix} a & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

Solution. One eigenvalue is $c \geq 0$, and we can look at the top left block in $\mathbb{M}_2(\mathbb{C})$ for the other two:

$$\begin{aligned} \begin{vmatrix} a-\lambda & 1/\sqrt{2} \\ 1/\sqrt{2} & b-\lambda \end{vmatrix} &= 0 \\ (a-\lambda)(b-\lambda) - \frac{1}{2} &= 0 \\ ab - \frac{1}{2} + \lambda^2 - \lambda(a+b) &= 0 \\ \lambda &= \frac{1}{2}\big((a+b) \pm \sqrt{(a-b)^2+2}\big). \end{aligned}$$

We can then take

$$(a+b)^2 \geq (a-b)^2 + 2 \Rightarrow ab \geq \frac{1}{2},$$

along with $a \geq 0$, ensuring that all eigenvalues are nonnegative.[7] □

Example 1.2.3 If ρ is a valid density matrix on some space $\mathcal{H}$, explain why $\langle\psi|\rho|\psi\rangle \geq 0$ for any state $|\psi\rangle \in \mathcal{H}$.

Proof. We use spectral decomposition to write[8]

$$\rho = \sum_k \lambda_k |e_k\rangle\langle e_k| .$$

By assumption, $\lambda_k \geq 0$ for all k, so let us compute:

$$\begin{aligned} \langle\psi|\rho|\psi\rangle = \langle\psi| \sum_k \lambda_k |e_k\rangle\langle e_k| \, |\psi\rangle &= \sum_k \lambda_k \langle\psi|e_k\rangle \langle e_k|\psi\rangle \\ &= \sum_k \lambda_k |c_k|^2 \geq 0, \quad \text{where } c_k := \langle\psi|e_k\rangle . \end{aligned}$$

□

So, we can do computations, but what is the motivation for these properties of the state space axiom? We will give a preview of the *measurement axiom* to understand.

Definition 1.2.2 (Measurement Axiom) *Every orthonormal basis* $\{|e_x\rangle\}_{x=1}^d$ *for* $\mathcal{H} := \mathbb{C}^d$ *represents a physically realizable measurement on any* d*-dimensional quantum system. The probability of obtaining outcome* x *when*

[7] 7: For unit trace, again we just need

$$a + b + c = 1.$$

[8] 8: The λ_k are eigenvalues and $\{|e_k\rangle\}$ are orthonormal eigenvectors.

measuring prepared state ρ is "Born's rule:"[9]

$$p(x) = \langle e_x|\rho|e_x\rangle \geq 0.$$

We also need[10]

$$1 = \sum_{x=1}^{d} p(x) = \sum_{x=1}^{d} \langle e_x|\rho|e_x\rangle = \operatorname{tr}\rho.$$

Now, there are two types of states that we need to distinguish from one another.

Definition 1.2.3 (Pure State) *Given $\rho \in \mathscr{D}(\mathscr{H})$, we say ρ is a pure state if* $\operatorname{rk}\rho = 1$.[11]

Remark 1.2.3 The choice of $|\psi\rangle$ is not unique! For every $\theta \in [0, 2\pi)$, the ket $e^{i\theta}|\psi\rangle$ yields the same density matrix.

Definition 1.2.4 (Global Phase) *The angle θ is called a global phase of $|\psi\rangle\langle\psi|$ if it is part of scaling factor $e^{i\theta}|\psi\rangle$.*[12]

Definition 1.2.5 (Relative Phase) *If θ occurs in a a scaling in superposition between two kets forming ψ, then it is called a relative phase of state $|\psi\rangle\langle\psi|$.*[13]

Why do we care about pure sates? Well, one reason is they allow for "deterministic" measurements, via Born's rule. That is, there is some outcome with probability 1 for $|\psi\rangle\langle\psi|$.

Example 1.2.4 Consider the pure state $\rho := |+\rangle\langle+|$ with

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

Suppose we measure ρ in the basis $\{|e_1\rangle, |e_2\rangle\}$ with

$$|e_1\rangle := \frac{1}{\sqrt{2}}\big(|0\rangle + e^{i\varphi}|1\rangle\big)$$

and

$$|e_2\rangle := \frac{1}{\sqrt{2}}\big(|0\rangle - e^{i\varphi}|1\rangle\big).$$

What is the probability of outcome $|e_1\rangle$?

Solution. Via Born's rule, we know

$$p(e_1) = \langle e_1|\rho|e_1\rangle = \langle e_1|+\rangle\,|+\rangle\langle e_1| = |\langle+|e_1\rangle|^2.$$

Computing through gives us that[14]

[9] Born's rule is precisely why we need the positivty of the state space axiom.

[10] This is why we need unit trace.

[11] Via spectral decomposition, if $\operatorname{rk}\rho = 1$, then $\rho = \lambda_1 |e_1\rangle\langle e_1|$, so

$$\rho = |\psi\rangle\langle\psi|.$$

Note that it is convention to write a pure state as $\rho = |\psi\rangle$ rather than the outer product $\rho = |\psi\rangle\langle\psi|$.

[12] Since it is equivalent, these are physically inconsequential.

[13] These are extremely physically consequential.

[14] When $\varphi = 0$, $p(e_1) = 1$.

$$p(e_1) = \frac{1}{2}(1 + \cos\varphi).$$

We get this result solely because we have a pure state. □

Definition 1.2.6 (Mixed State) *A density matrix $\rho \in \mathfrak{D}(\mathcal{H})$ is called mixed if it is not pure.*[15]

Remark 1.2.4 Note that the eigenvalues of a mixed state ρ form a probability distribution, via the unit trace axiom.

Definition 1.2.7 (Ensemble) *The set $\{p_i, |\psi_i\rangle\}$*[16] *is known as an ensemble of pure sates, and the density matrix*

$$\rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i|$$

is called an ensemble average.[17]

Remark 1.2.5 You may hear people saying that the mixed state

$$\rho = \sum_{i=1}^{r} p_i \, |e_i\rangle\langle e_i|$$

is typically interpreted as the system *being in state $|e_i\rangle$ with probability p_i*. This is not a good interpretation.[18]

In this case, how do we characterize all the ensembles $\{p_i, |\psi_i\rangle\}_{i=1}^{r}$ that have the same ensemble average?

Theorem 1.2.2 *Two pure state ensembles $\{p_i, |\psi_i\rangle\}_{i=1}^{r}$ and $\{q_j, |\varphi_j\rangle\}_{j=1}^{s}$ have the same ensemble average if and only if*

$$\sqrt{q_i}\,|\varphi_i\rangle = \sum_{j=1}^{s} u_{ij}\sqrt{p_j}\,|\psi_j\rangle$$

for all $i \in [r]$, where the $[U]_{ij} := u_{ij}$ is a unitary matrix.

Proof. The proof is omitted for brevity.[19] □

$$[\text{RNG}] \xrightarrow{\;i\;} \begin{bmatrix}\text{Particle}\\ \text{Emitter}\end{bmatrix} \longrightarrow |\varphi_i\rangle$$

Figure 1.1: The preparation of a system in state $|\varphi_i\rangle$ with probability p_i

Our struggles with describing a system in some state $|\varphi_i\rangle$ with probability q_i suggest that we need a multi-part system, incorporating a "random number generator" on number i to trigger a "particle emitter" to release a state in state i.[20] We now take some time to discuss *qubits* and the *Bloch*

[15] That is, $\operatorname{rk}\rho > 1$.

[16] The p_i are the eigenvalues.

[17] Alternatively, you will hear *convex combination*. This gives us a way to average over the pure states.

[18] The reason for this is the fact that two different pure state ensembles can have the same ensemble average.

[19] It follows a rather simple argument via the polar decomposition.

[20] That is, we need a way to mathematically describe *multiple systems*.

sphere.

Definition 1.2.8 (Qubit) *A qubit is a generic name for any two-dimensional quantum system $\mathscr{H} \simeq \mathbb{C}^2$.*

In the computational basis, we have a pure state in the form $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. We only really have two parameters (θ, φ) on $\mathbb{S}^2 \subseteq \mathbb{R}^3$, called the Bloch sphere. Note that we have

$$\hat{n} = (\cos\varphi \sin\theta, \sin\varphi \sin\theta, \cos\theta)$$

if and only if

$$|\hat{n}\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}e^{i\varphi}|1\rangle .$$

The vector $\hat{n} \in \mathbb{R}^3$ is known as the *Bloch vector* of $|\hat{n}\rangle \in \mathbb{C}^2$. Precisely, $\theta \in [0, \pi]$ is the polar angle, and $\varphi \in [0, 2\pi)$ is the azimuthal angle.

Remark 1.2.6 We have that $(\theta, \varphi) = (0, 0)$ yields $|0\rangle$, whereas $(\pi, 0)$ yields $|1\rangle$. Similarly, $(\pi/2, 0)$ yields $|+\rangle$, and $(\pi/2, \pi)$ yields $|-\rangle$. Finally, on the y-axis, if we have $(\pi/2, \pi/2)$ and $(\pi/2, 3\pi/2)$, we get

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

respectively.

Example 1.2.5 The H and T states are important in quantum computing. The Bloch vector of the H state lies along the line $(x, 0, x)$, and the Bloch vector of T lies along the line (x, x, x). Write $|H\rangle$ and $|T\rangle$.[21]

Solution. For H, we have $(\theta, \varphi) = (\pi/4, 0)$ and for T we have $(\theta, \varphi) = (\arccos 1/\sqrt{3}, \pi/4)$. Thus,

$$|H\rangle = \cos\frac{\pi}{8}|0\rangle + \sin\frac{\pi}{8}|1\rangle$$
$$|T\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\pi/4}\sin\frac{\theta}{2}|1\rangle .$$

□

If we write out the density matrix $|\hat{n}\rangle\langle\hat{n}|$, we get[22]

$$|\hat{n}\rangle\langle\hat{n}| = \frac{1}{2}(I_2 + \sin\theta\cos\varphi\sigma_x + \sin\theta\sin\varphi\sigma_y + \cos\theta\sigma_z).$$

That is, we can decompose the density matrix via the Pauli matrices, using the components of the cartesian representation. Note that the Pauli matrices are hermitian and unitary. That is, $\sigma_j^\dagger = \sigma_j$ and σ_j^2 for all $j \in \{x, y, z\}$. They also anti-commute.[23] Finally, they satisfy $\sigma_j\sigma_k = i\epsilon_{jkl}\sigma_l$, where ϵ_{ijk} is the *Levi-Civita* symbol.

[21] Use the computational basis, and assume the Bloch vectors lie in the first quadrant of the Bloch sphere embedded in $\mathbb{R}^3$.

[22] Use the half-angle formulae, after expanding the outer product.

[23] That is, $\{\sigma_j, \sigma_k\} = 2\delta_{jk}\mathbb{1}$.

Remark 1.2.7 Via these observations, we can write

$$|\hat{n}\rangle\langle\hat{n}| = \frac{1}{2}(\mathbb{1} + \hat{n}\cdot\vec{\sigma}) = \frac{1}{2}\left(\mathbb{1} + \sum_{i=1}^{3} n_i\sigma_i\right).$$

where $\vec{\sigma} = \sigma_x\hat{x} + \sigma_y\hat{y} + \sigma_z\hat{z}$ and $\hat{n} = (n_x, n_y, n_z)$.

As such,[24]

$$|\langle\hat{m}|\hat{n}\rangle|^2 = \frac{1}{2}(1 + \hat{m}\cdot\hat{n}).$$

This means that $|\hat{m}\rangle$ is orthogonal to $|\hat{n}\rangle$ if and only if $\hat{m}\cdot\hat{n} = -1$. Thus, orthogonal states on the Bloch sphere correspond to *antipodal* points on the sphere; i.e., $|\hat{m}\rangle = |-\hat{n}\rangle$. Now, considering mixed states, we have that a qubit density matrix ρ has the spectral decomposition

$$\begin{aligned}\rho &= \frac{1+\lambda}{2}\,|\hat{n}\rangle\langle\hat{n}| + \frac{1-\lambda}{2}\,|-\hat{n}\rangle\langle-\hat{n}| \\ &= \frac{1}{2}(\mathbb{1} + \lambda\hat{n}\cdot\vec{\sigma}).\end{aligned}$$

Then, every ρ can be written as a Bloch vector $\hat{n}$ with shrunken length λ. That is, $\mathbf{r} := \lambda\hat{n}$ is the Bloch vector:[25]

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r}\cdot\vec{\sigma}).$$

Example 1.2.6 Find values a, b such that $\rho = a\mathbb{1} + b\sigma_y$ is a valid density matrix. What is the Bloch vector of ρ when it is a valid density matrix?

Solution. We need $\operatorname{tr}\rho = 1$ and $\rho \geq 0$. Thus, $\operatorname{tr}\rho = 1 = 2a$, meaning $a = 1/2$, so rewriting gives us

$$\rho = \frac{1}{2}\mathbb{1} + b\sigma_y.$$

We know that σ_y has eigenvectors $\left|\widetilde{\pm}\right\rangle$. Spectral decomposition tells us that

$$\sigma_y = \left|\widetilde{+}\middle\rangle\middle\langle\widetilde{+}\right| - \left|\widetilde{-}\middle\rangle\middle\langle\widetilde{-}\right|,$$

and since

$$\mathbb{1} = \left|\widetilde{+}\middle\rangle\middle\langle\widetilde{+}\right| + \left|\widetilde{-}\middle\rangle\middle\langle\widetilde{-}\right|,$$

so substituting gives us

$$\rho = \left(\frac{1}{2} \pm b\right)\left|\widetilde{\pm}\middle\rangle\middle\langle\widetilde{\pm}\right|.$$

As such, $b \in [-1/2, 1/2]$. The eigenvalues are $1/2 \pm b$, so $\lambda = 2b$, meaning our Bloch vector is

$$\mathbf{r} = \lambda\hat{y} = 2b\hat{y}.$$

□

[24] Once again, we get a nice geometrical picture which takes an inner product in $\mathbb{C}^2$ to a dot product in $\mathbb{R}^3$.

[25] When we have $\lambda = 0$, we get the *totally mixed state* $\rho = \mathbb{1}/2$.

Example 1.2.7 What are Bloch vectors of states

$$\rho = p\frac{1}{2}\mathbb{1} + (1-p)\,|H\rangle\langle H|$$

and

$$\sigma = q\frac{1}{2}\mathbb{1} + (1-q)\,|T\rangle\langle T|?$$

Solution. We have that the Bloch vector of $|H\rangle$:

$$\hat{n} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right),$$

so

$$\begin{aligned}\rho &= p\frac{1}{2}\mathbb{1} + (1-p)\,|H\rangle\langle H| \\ &= \frac{1}{2}(\mathbb{1} + (1-p)\hat{n}\cdot\sigma),\end{aligned}$$

so the Bloch vector of ρ is[26]

$$\mathbf{r} = \frac{1-p}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}.$$

Now, for $|T\rangle$, we have the Bloch vector

$$\hat{n} = \frac{1}{\sqrt{3}}(1,1,1).$$

Then, writing out the outer product, we get

$$|T\rangle\langle T| = \frac{1}{2}\left(\mathbb{1} + \frac{1}{\sqrt{3}}(\sigma_x + \sigma_y + \sigma_z)\right).$$

Thus,

$$\begin{aligned}\sigma &= q\frac{1}{2}\mathbb{1} + (1-q)\frac{1}{2}\left(\mathbb{1} + \frac{1}{\sqrt{3}}(\sigma_x + \sigma_y + \sigma_z)\right) \\ &= \frac{1}{2}\left(\mathbb{1} + \frac{1-q}{\sqrt{3}}(\sigma_x + \sigma_y + \sigma_z).\right)\end{aligned}$$

Thus, the Bloch vector of σ is

$$\mathbf{r} = \frac{1-q}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

□

Proposition 1.2.3 *For $\rho \in \mathcal{D}(\mathscr{H})$, show that* $\mathrm{tr}(\rho^2) = 1$ *if and only if ρ is a pure state.*[27]

[26] The Bloch vector is what we dot with σ, so the goal is to get ρ into a form such that we can identify $\mathbf{r}$.

[27] Show this using both the spectral decomposition for arbitrary d and the Bloch sphere for $d = 2$.

Proof. If ρ is pure, $\rho = |\psi\rangle\langle\psi|$. Then, $\rho^2 = |\psi\rangle\langle\psi|$, so $\mathrm{tr}(\rho^2) = \mathrm{tr}(\rho) = 1$. Now, for the forward direction, assume $\mathrm{tr}(\rho^2) = 1$. Spectral decomposition tells us that we can write

$$\rho = \sum_k \lambda_k |e_k\rangle\langle e_k|,$$

where $0 \le \lambda_k \le 1$[28] and $|e_k\rangle$ are orthonormal. Then,

$$\rho^2 = \sum_k \lambda_k^2 |e_k\rangle\langle e_k|,$$

so taking the trace, we get

$$\mathrm{tr}(\rho^2) = \sum_k \lambda_k^2 \le \sum_k \lambda_k = 1.$$

Under which case do we get equality? Well, we need each $\lambda_k \in \{0, 1\}$. Thus, there is only one nonzero eigenvalue equaling 1, so $\rho = |\psi\rangle\langle\psi|$ is pure.[29]

Now, consider the qubit case. Then, we can write

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r}\cdot\vec{\sigma}).$$

Then, the trace of the square is

$$\mathrm{tr}(\rho^2) = \frac{1}{4}\mathrm{tr}\big((\mathbb{1} + \mathbf{r}\cdot\vec{\sigma})(\mathbb{1} + \mathbf{r}\cdot\vec{\sigma})\big) = \frac{1}{2}(1 + \mathbf{r}\cdot\mathbf{r}),$$

and these equals 1 if and only if $\mathbf{r}\cdot\mathbf{r} = 1$. Thus, ρ is pure. □

Example 1.2.8 Describe the motion of the Bloch vector on the Bloch sphere of the time-dependent state

$$|\psi(t)\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi t}\sin\frac{\theta}{2}|1\rangle.$$

Solution. Well, only φ is scaled by t, and φ is the azimuthal angle, so on the Bloch sphere, so the resultant motion of the time-dependent state would be a rotation around the z-axis, keeping the polar angle θ fixed.[30] □

Example 1.2.9 Two orthonormal bases $\{|\hat{n}\rangle, |-\hat{n}\rangle\}$ and $\{|\hat{m}\rangle, |-\hat{m}\rangle\}$ are called *mutually unbiased* if

$$\frac{1}{2} = |\langle\pm\hat{n}|\pm\hat{m}\rangle|^2 = |\langle\pm\hat{n}|\mp\hat{m}\rangle|^2.$$

Describe the relationship, geometrically, of the Bloch vectors of any two mutually unbiased bases.

Solution. In either case, we need $\hat{m}\cdot\hat{n}$ to vanish, so any vectors are mutually unbiased with $\hat{n}$ if and only if they lie in the plane $\perp$ to $\hat{n}$. □

[28] We know ρ is positive, and we know the eigenvalues must be bounded by 1, as $\mathrm{tr}\,\rho = 1$.

[29] Rank-1 implies pure.

[30] That is, we are "stuck" on a line of latitude of the sphere.

2 Multiple System Axiom

Recall that the state space axiom tells us how we mathematically represent a quantum system. Yet, what if we have one large system consisting of multiple subsystems, say A and B?

Definition 2.0.1 (Multiple System Axiom) *The joint system of* A *and* B *is a Hilbert space*

$$\mathscr{H}^{\mathsf{AB}} := \mathscr{H}^{\mathsf{A}} \otimes \mathscr{H}^{\mathsf{B}}.$$

We will give some computational background on what *tensor spaces* are. Well, let $\mathscr{H}^{\mathsf{A}}$ and $\mathscr{H}^{\mathsf{B}}$ be Hilbert spaces with bases $|i\rangle_i^{\mathsf{A}}$ and $|j\rangle_j^{\mathsf{B}}$. Then, the basis of the tensor product basis is precisely the set $|i\rangle^{\mathsf{A}} \otimes |j\rangle_{ij}^{\mathsf{B}}$. Then, for any element $|\psi\rangle^{\mathsf{AB}} \in \mathscr{H}^{\mathsf{AB}}$ can be written as[1]

$$|\psi\rangle^{\mathsf{AB}} = \sum_{i=1}^{\dim \mathscr{H}^{\mathsf{A}}} \sum_{j=1}^{\dim \mathscr{H}^{\mathsf{B}}} c_{ij} \, |i\rangle^{\mathsf{A}} \otimes |j\rangle^{\mathsf{B}}.$$

Definition 2.0.2 (Joint State) *Given a state* $|\alpha\rangle^{\mathsf{A}}$ *of Alice's system and* $|\beta\rangle^{\mathsf{B}}$ *of Bob's system, their joint state is* $|\alpha\rangle^{\mathsf{A}} \otimes |\beta\rangle^{\mathsf{B}}$.[2]

Example 2.0.1 Consider the two qubit state space $\mathscr{H}^{\mathsf{AB}} \simeq \mathbb{C}^2 \otimes \mathbb{C}^2$. Express

$$|\Psi\rangle^{\mathsf{AB}} = \frac{1}{\sqrt{3}}(|0+\rangle + |-1\rangle).$$

Solution. We can rewrite

$$\begin{aligned} |\Psi\rangle^{\mathsf{AB}} &= \frac{1}{\sqrt{3}}\left(|0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) + \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \otimes |1\rangle\right) \\ &= \frac{1}{\sqrt{6}}|0\rangle \otimes |0\rangle + \frac{2}{\sqrt{6}}|0\rangle \otimes |1\rangle - \frac{1}{\sqrt{6}}|1\rangle \otimes |1\rangle \\ &= \frac{1}{\sqrt{6}}(|00\rangle + 2|01\rangle - |11\rangle). \end{aligned}$$

□

This axiom leads us naturally to the property of *quantum entanglement*. Joint states are a very special type of state known as a *product state*. In particular, it is very rare for us to be able to decompose a pure state into a product state.

Definition 2.0.3 (Entangled State) *If a pure bipartite state cannot be written as a product state, then it is called an entangled state.*[3]

[1] Note that $\otimes$, is bilinear.

[2] Sometimes, we will simply use juxtaposition for denoting the joint state.

[3] Quantum entanglement is the absolute *essential* feature of multipartite quantum systems that enables "stronger-than-classical" processing.

Recall that for any state $|\alpha\rangle \in \mathcal{H}^{\mathsf{A}}$ in a finite Hilbert space, we can associate it to a column vector in $\mathbb{C}^{d\mathsf{A}}$.

Definition 2.0.4 (Kronecker Product) *Given $|\alpha\rangle$, $|\beta\rangle$, we write the Kronecker product of the corresponding column vectors as*[4]

$$|\alpha\rangle \otimes |\beta\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{d\mathsf{A}} \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{d\mathsf{B}} \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ \vdots \\ a_1 b_{d\mathsf{B}} \\ a_2 b_1 \\ \vdots \\ a_{d\mathsf{A}} b_{d\mathsf{B}} \end{pmatrix}$$

Example 2.0.2 What is the matrix representation of[5]

$$\left|\Phi^+\right\rangle := \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)?$$

Solution. We have that

$$|00\rangle = |0\rangle \otimes |0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and

$$|11\rangle = |1\rangle \otimes |1\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

so

$$\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

□

Now, how do linear operators work on a Hilbertian tensor product? Well, every map[6]

$$K : \mathcal{H}^{\mathsf{A}} \otimes \mathcal{H}^{\mathsf{B}} \to \mathcal{H}^{\mathsf{A}'} \otimes \mathcal{H}^{\mathsf{B}'}$$

can be written as

$$K = \sum_{i=1}^{d\mathsf{A}'} \sum_{k=1}^{d\mathsf{A}} \sum_{j=1}^{d\mathsf{B}'} \sum_{\ell=1}^{d\mathsf{B}} c_{ijk\ell} \, |i\rangle\langle k| \otimes |j\rangle\langle \ell| \, ,$$

so the action of $K \, |\psi\rangle^{\mathsf{AB}}$ is

$$\sum_{i=1}^{d\mathsf{A}'} \sum_{k=1}^{d\mathsf{A}} \sum_{j=1}^{d\mathsf{B}'} \sum_{\ell=1}^{d\mathsf{B}} c_{ijk\ell} b_{k\ell} \, |i\rangle^{\mathsf{A}'} \otimes |j\rangle^{\mathsf{B}'} .$$

[4] It is clear this gives us a vector in $\mathbb{C}^{d\mathsf{A}d\mathsf{B}}$, as desired.

[5] Note that the notation Φ^+ is because of the addition of the tensored states. We can also define a Φ^-, in the way you would expect.

[6] These sort of gymnastics are usually referred to as "tensor contractions," which you may see in a variety of related fields.

Example 2.0.3 The two-qubit unitary operator

$$U_{\mathsf{CNOT}}^{\mathsf{AB}} := |0\rangle\langle 0|^{\mathsf{A}} \otimes \mathbb{1}^{\mathsf{B}} + |1\rangle\langle 1|^{\mathsf{A}} \otimes \sigma_x^{\mathsf{B}}$$

is called the *controlled-not* (CNOT) gate. Compute the action of $U_{\mathsf{CNOT}}^{\mathsf{AB}}$ on the computational basis states. What is its action on the state $|+\rangle\,|0\rangle$?

Solution. Recall that $\mathbb{1} = |0\rangle\langle 0| + |1\rangle\langle 1|$ and $\sigma_x = |1\rangle\langle 0| + |0\rangle\langle 1|$.[7] Thus, we can write

$$U_{\mathsf{CNOT}} = |0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 0| \otimes |1\rangle\langle 1| + |1\rangle\langle 1| \otimes |1\rangle\langle 0| + |1\rangle\langle 1| \otimes |0\rangle\langle 1|\,,$$

so we have the action[8]

$$U_{\mathsf{CNOT}} : \begin{cases} |00\rangle \mapsto |00\rangle \\ |01\rangle \mapsto |01\rangle \\ |10\rangle \mapsto |11\rangle \\ |11\rangle \mapsto |10\rangle\,. \end{cases}$$

We get that

$$U_{\mathsf{CNOT}}\,|{+}0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \left|\Phi^+\right\rangle.$$

Similarly, $U_{\mathsf{CNOT}}\,|{-}{+}\rangle = |{-}{+}\rangle$. □

Remark 2.0.1 As with vectors, we can represent the tensor product of $T_A \otimes T_B$ by the Kronecker product $A \otimes B$.[9]

Example 2.0.4 Consider a generic two-qubit state

$$|\psi\rangle := a\,|00\rangle + b\,|01\rangle + c\,|10\rangle + d\,|11\rangle\,.$$

The SWAP operator on $(\mathbb{C}^2)^{\otimes 2}$ is defined as

$$\mathbb{F} := |00\rangle\langle 00| + |10\rangle\langle 01| + |01\rangle\langle 10| + |11\rangle\langle 11|\,.$$

Compute $\mathbb{F}\,|\psi\rangle$.

Solution. We compute[10]

$$\mathbb{F}\,|\psi\rangle = a\,|00\rangle + b\,|10\rangle + c\,|01\rangle + d\,|11\rangle\,.$$

□

Remark 2.0.2 Let $|\alpha\rangle = a_0\,|0\rangle + a_1\,|1\rangle$ and $|\beta\rangle = b_0\,|0\rangle + b_1\,|1\rangle$. Then,

$$\mathbb{F}\,|\alpha\rangle^{\mathsf{A}}\,|\beta\rangle^{\mathsf{B}} = a_0 b_0\,|00\rangle + a_0 b_1\,|10\rangle + a_1 b_0\,|01\rangle + a_1 b_1\,|11\rangle = |\beta\rangle^{\mathsf{A}}\,|\alpha\rangle^{\beta}\,.$$

[7] Remember, σ_x is just the Pauli operator $X \in \mathscr{P}_1$.

[8] The CNOT gate gets its name from the fact that the "control" checks whether the first qubit is on, and if so, it performs a "not" on the second qubit.

[9] We just multiply each component of A by all of B.

[10] The computation makes it pretty clear why we call $\mathbb{F}$ the SWAP operator.

Note that we have the matrix representation[11]

$$\mathbb{F} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Combining the two axioms we have seen thus far, we have that for two systems $\mathscr{H}^{\mathsf{A}}, \mathscr{H}^{\mathsf{B}}$, their joint state is a bipartite density operator

$$\rho^{\mathsf{AB}} = \sum_{i,k=1}^{d\mathsf{A}} \sum_{j,\ell}^{d\mathsf{B}} c_{ijk\ell} \, |i\rangle\langle k|^{\mathsf{A}} \otimes |j\rangle\langle \ell|^{\mathsf{B}},$$

where

(i) $\rho^{\mathsf{AB}} \geq 0$,
(ii) $\operatorname{tr} \rho^{\mathsf{AB}} = 1$.

Note that we can then take a spectral decomposition[12]

$$\rho^{\mathsf{AB}} = \sum_{i=1}^{r} p_i \, |e_i\rangle\langle e_i| \, .$$

Remark 2.0.3 Recall our visualization of random preparations via an RNG and a particle emitter. Now, we need to include the state of the RNG in our description. However, an RNG is a classical system.[13]

Definition 2.0.5 (Classical Register) *A classical register* X *is a system whose allowed states are always density matrices diagonal in the computational basis. Then,*[14]

$$\rho^{\mathsf{X}} = \sum_{x \in \mathcal{X}} p_x \, |x\rangle\langle x| \, .$$

Then, we can form a *quantum-classical* state:

$$\rho^{\mathsf{SX}} = \sum_{x \in \mathcal{X}} p_x \, |\varphi_x\rangle\langle \varphi_x|^{\mathsf{S}} \otimes |x\rangle\langle x|^{\mathsf{X}} \, .$$

Example 2.0.5 Suppose two fair dice are rolled. Let $s \in [12] \setminus \{1\}$ denote their sum. Let e be a variable such that

$$\begin{cases} e = 0, & \text{dice show same number} \\ e = 1, & \text{dice show different numbers.} \end{cases}$$

Compute the joint probabilities $p(s, e)$ for all outcomes. Compute the marginal probabilities $p(s)$, $p(e)$. Compute the conditional probabilities, as well.[15]

[11] Use the Kroncker product to determine this.

[12] Note that the eigenvectors $|e_i\rangle$ could be entangled vectors in $\mathscr{H}^{\mathsf{A}} \otimes \mathscr{H}^{\mathsf{B}}$.

[13] That is, it outputs bits.

[14] You can think of classical registers as probability distributions.

[15] We can get $p(s, e)$ by writing out all 36 outcomes in a table. Then, we can compute

$$p(s) = \sum_{e} p(s, e),$$

and likewise for $p(e)$. Finally, recall that

$$p(e \mid s) = \frac{p(s, e)}{p(s)}.$$

Remark 2.0.4 Recall Born's rule:

$$p(x) = \langle e_x|\rho\rangle\, e_x = \operatorname{tr}\left(\rho\, |e_x\rangle\!\langle e_x|\right).$$

If Alice and Bob share the bipartite state ρ^{AB}, then

$$p(x, y) = \operatorname{tr}\left(\rho^{\mathsf{AB}}\, |a_x\rangle\!\langle a_x| \otimes |b_y\rangle\!\langle b_y|\right).$$

Then, as we know, the marginal distribution for Alice is[16]

$$p(x) = \sum_{y=1}^{d\mathsf{B}} \operatorname{tr}\left(\rho^{\mathsf{AB}}\, |a_x\rangle\!\langle a_x| \otimes |b_y\rangle\!\langle b_y|\right) = \operatorname{tr}\left(\rho^{\mathsf{AB}}\, |a_x\rangle\!\langle a_x| \otimes \mathbb{1}\right).$$

Remark 2.0.5 This is a manifestation of "no-signaling."

We can then rewrite

$$\rho^{\mathsf{AB}} = \sum_{j,\ell=1}^{d\mathsf{B}} R^{\mathsf{A}}_{j,\ell} \otimes |j\rangle\!\langle \ell|^{\mathsf{B}},$$

where

$$R_{j,\ell} := \sum_{i,k=1}^{d\mathsf{A}} c_{ijk\ell}\, |i\rangle\!\langle k|^{\mathsf{A}} \in \mathbb{B}(\mathcal{H}^{\mathsf{A}}).$$

Thus,[17]

$$p(x) = \langle a_x| \left(\sum_{j=1}^{d\mathsf{B}} R^{\mathsf{A}}_{jj} \right) |a_x\rangle .$$

Definition 2.0.6 (Partial Trace I) *For a bipartite operator*

$$\rho^{\mathsf{AB}} = \sum_{j,\ell}^{d\mathsf{A}} |i\rangle\!\langle k|^{\mathsf{A}} \otimes S^{\mathsf{B}}_{ik},$$

its partial trace over A *is the operator*

$$\rho^{\mathsf{B}} := \operatorname{tr}_{\mathsf{A}} \rho^{\mathsf{AB}} \in \mathbb{B}(\mathcal{H}^{\mathsf{B}}) \textit{ given by } \rho^{\mathsf{B}} = \sum_{i=1}^{d\mathsf{A}} S^{\mathsf{B}}_{ii}.$$

If $K = R \otimes S$ is a product operator, then

$$\operatorname{tr}_{\mathsf{A}}(K) = \operatorname{tr}_{\mathsf{A}}(R \otimes S) = \operatorname{tr}(R)S.$$

We get that if ρ^{AB} is a density matrix, then so is $\operatorname{tr}_{\mathsf{A}} \rho^{\mathsf{AB}} = \rho^{\mathsf{B}}$.[18] In this case, $\rho^{\mathsf{A}}, \rho^{\mathsf{B}}$ are *reduced density matrices.*

Definition 2.0.7 (Partial Trace II) *The partial trace* $\operatorname{tr}_{\mathsf{B}}$ *is a map*

$$\mathbb{B}(\mathcal{H}^{\mathsf{A}} \otimes \mathcal{H}^{\mathsf{B}}) \xrightarrow{\operatorname{tr}_{\mathsf{B}} := \mathrm{id}_{\mathsf{A}} \otimes \operatorname{tr}} \mathbb{B}(\mathcal{H}^{\mathsf{A}}).$$

[16] We use the completion relation.

[17] That is, we can express Alice's marginal in terms of the new operator, which depends on the joint density operator.

[18] That is, the partial trace is CPTP.

Note that in the *Pauli basis*, every two-qubit state has a convenient form:[19]

$$\rho^{\mathsf{AB}} := \frac{1}{4}\left(\mathbb{1}^{\otimes 2} + \mathbf{r}\cdot\vec{\sigma}\otimes\mathbb{1} + \mathbb{1}\otimes\mathbf{s}\cdot\vec{\sigma} + \sum_{i,j=1}^{3} t_{ij}\sigma_i\otimes\sigma_j\right).$$

Remark 2.0.6 Together, the last nine terms, $\sum t_{ij}\sigma_i \otimes \sigma_j$, is called the correlation matrix.

Let us now look at a particularly useful basis for our entangled qubits.

Definition 2.0.8 (Bell Basis) *There is an entangled basis for* $(\mathbb{C}^2)^{\otimes 2}$*:*

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

Remark 2.0.7 Note that

$$\left|\Phi_{b_0 b_1}\right\rangle = \sigma_z^{b_0}\sigma_x^{b_1}\otimes\mathbb{1}\,|\Phi_{00}\rangle\,,$$

where $b_0, b_1 \in \{0, 1\}$.[20] The Bell states can be interpreted as errors.

Example 2.0.6 Write the computational basis in the Bell basis.

Solution. By observation,

$$\begin{aligned}
|00\rangle &= \frac{1}{\sqrt{2}}(\left|\Phi^+\right\rangle + |\Phi^-\rangle)\\
|01\rangle &= \frac{1}{\sqrt{2}}(\left|\Psi^+\right\rangle + |\Psi^-\rangle),
\end{aligned}$$

and so forth. □

Definition 2.0.9 (Bell-Diagonal) *We call a density matrix Bell-diagonal state if its eigenvectors are the Bell states:*

$$\rho^{\mathsf{AB}} = \sum_{i,j=0}^{1} p_{ij}\left|\Phi_{ij}\middle\rangle\!\middle\langle\Phi_{ij}\right|.$$

Example 2.0.7 Compute the local Bloch vectors and the correlation matrix of a generic Bell-diagonal state.

[19] We have that **r** is Alice's Bloch vector and **s** is Bob's Bloch vector. *Nice!* Note further that this expression has 16 terms, as we would expect.

[20] Note that if we take $b_0 = b_1 = 1$, then

$$\sigma_z\sigma_x = i\sigma_y.$$

We call b_0 the "phase" bit and b_1 the "amplitude" bit.

Solution. Remember,

$$r_i = \operatorname{tr}\left(\rho^{\mathsf{AB}}\sigma_i \otimes \mathbb{1}\right) = \operatorname{tr}\left(\sigma_i \operatorname{tr}_{\mathsf{B}}(\rho^{\mathsf{AB}}(\mathbb{1}\otimes\mathbb{1})).\right)$$

Of course,[21]

$$\operatorname{tr}_{\mathsf{B}}(\rho^{\mathsf{AB}}) = \sum_{ij} p_{ij} \operatorname{tr}_{\mathsf{B}}(\left|\Phi_{ij}\middle\rangle\!\middle\langle\Phi_{ij}\right|).$$

Now, we just saw that

$$\left|\Phi_{ij}\right\rangle = \sigma_x^i \sigma_x^j \otimes \mathbb{1}\left|\Phi_{00}\right\rangle.$$

Thus, we can rewrite the partial trace as

$$\frac{1}{2}\operatorname{tr}_{\mathsf{B}}(\left|\Phi_{ij}\middle\rangle\!\middle\langle\Phi_{ij}\right|) = \sigma_z^i \sigma_x^j (|0\rangle\!\langle 0| + |1\rangle\!\langle 1|)\sigma_x^j \sigma_z^i = \mathbb{1}/2.$$

Finally, we get

$$\operatorname{tr}(\sigma_x \mathbb{1}/2) = 0.$$

Then, the correlation matrix is given by[22]

$$t_{ij} = \operatorname{tr}\left(\rho^{\mathsf{AB}}\sigma_i \otimes \sigma_j\right).$$

□

Theorem 2.0.1 (First Canonical Isomorphism) *We have an isomorphism*

$$\mathscr{H}^{\mathsf{AB}} = \mathscr{H}^{\mathsf{A}} \otimes \mathscr{H}^{\mathsf{B}} \xrightarrow{\sim} \mathbb{B}(\mathscr{H}^{\mathsf{B}} : \mathscr{H}^{\mathsf{A}}).$$

The map is given by[23]

$$|\psi\rangle^{\mathsf{AB}} \mapsto M_\psi,$$

where

$$|\psi\rangle^{\mathsf{AB}} = \sum_{i=1}^{d\mathsf{A}} \sum_{j=1}^{d\mathsf{B}} m_{ij}\,|i\rangle^{\mathsf{A}} \otimes |j\rangle^{\mathsf{B}}$$

and

$$M_\psi = \sum_{i=1}^{d\mathsf{A}} \sum_{j=1}^{d\mathsf{B}} m_{ij}\,|i\rangle^{\mathsf{A}} \langle j|^{\mathsf{B}}.$$

Corollary 2.0.2 *The state $|\psi\rangle$ is a product state if and only if M_ψ is rank one.*

Proof. The proof is clear from observation,. □

Theorem 2.0.3 (Ricochet Property) *An arbitrary $A \in \mathbb{B}(\mathscr{H}^{\mathsf{B}} : \mathscr{H}^{\mathsf{A}})$ satisfies*

$$A \otimes \mathbb{1}\left|\varphi^+_{d\mathsf{B}}\right\rangle = \mathbb{1} \otimes A^t \left|\varphi^+_{d\mathsf{A}}\right\rangle.$$

Let $|\psi\rangle^{\mathsf{AB}} \in \mathscr{H}^{\mathsf{AB}}$ be arbitrary, writing

$$|\psi\rangle^{\mathsf{AB}} = M_\psi \otimes \mathbb{1}\left|\varphi^+_{d\mathsf{B}}\right\rangle.$$

[21] In general, if we have

$$\operatorname{tr}(\rho^{\mathsf{AB}}) = \operatorname{tr}(\operatorname{tr}_{\mathsf{A}} \rho^{\mathsf{AB}}).$$

We can always write

$$\operatorname{tr}_{\mathsf{B}}((A \otimes \mathbb{1})\rho(A^\dagger \otimes \mathbb{1}))$$

as $A \operatorname{tr}_{\mathsf{B}} \rho A^\dagger$.

[22] Complete this as an exercise.

[23] That is, $|i\rangle\,|j\rangle$ corresponds to $|i\rangle\!\langle j|$ and $|\alpha\rangle\,|\beta\rangle$ corresponds to $|\alpha\rangle\!\langle\beta^*|$, where

$$|\beta\rangle = \sum_{i=1}^{d\mathsf{B}} b_i\,|i\rangle.$$

Taking the partial trace over B of $|\psi\rangle\langle\psi|$ gives us

$$\mathrm{tr}_{\mathsf{B}}\left(|\psi\rangle\langle\psi|\right) = M_\psi\, \mathrm{tr}_{\mathsf{B}}\left(\left|\varphi^+_{d\mathsf{B}}\right\rangle\!\left\langle\varphi^+_{d\mathsf{B}}\right|\right) M_\psi^\dagger = M_\psi M_\psi^\dagger .$$

Likewise, we could write[24]

$$|\psi\rangle^{\mathsf{AB}} = \mathbb{1} \otimes M_\psi^t \left|\varphi^+_{d\mathsf{A}}\right\rangle ,$$

so

$$\mathrm{tr}_{\mathsf{A}}\left(|\psi\rangle\langle\psi|\right) = M_\psi^\dagger M_\psi^* .$$

Then, for any bipartite density matrix

$$\rho^{\mathsf{AB}} = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| ,$$

we have the reduced density matrices[25]

$$\rho^{\mathsf{A}} = \sum_i p_i \, M_{\psi_i} M_{\psi_i}^\dagger$$

and

$$\rho^{\mathsf{B}} = \sum_i p_i \, M_{\psi_i}^t M_{\psi_i}^* .$$

Remark 2.0.8 Recall that the SVD of an operators tells us that if $M \in \mathbb{B}(\mathscr{H}^{\mathsf{B}}, \mathscr{H}^{\mathsf{A}})$, then there exist unitaries $U \in \mathbb{B}(\mathscr{H}^{\mathsf{A}})$ and $V \in \mathbb{B}(\mathscr{H}^{\mathsf{B}})$ such that

$$M = U \Lambda V^\dagger ,$$

where $\Lambda \in \mathbb{B}(\mathscr{H}^{\mathsf{B}}, \mathscr{H}^{\mathsf{A}})$.[26] Note that

$$MM^\dagger = U\Lambda^2 U^\dagger \text{ and } M^\dagger M = V \Lambda^2 V^\dagger .$$

Thus, the singular values of M are the square roots of the eigenvalues of $M^\dagger M$ and $MM^\dagger$.

Theorem 2.0.4 (Schmidt Decomposition) *Every bipartite state $|\psi\rangle^{\mathsf{AB}} \in \mathscr{H}^{\mathsf{AB}}$ can be written as*

$$|\psi\rangle^{\mathsf{AB}} = \sum_{j=1}^{r} \sigma_j \left|\alpha_j\right\rangle^{\mathsf{A}} \otimes \left|\beta_j\right\rangle^{\mathsf{B}} ,$$

where the $\left|\alpha_j\right\rangle$ and $\left|\beta_j\right\rangle$ are orthonormal bases called the Schmidt bases, the σ_i are the Schmidt coefficients, and r is the Schmidt rank.

Proof. Let us apply SVD to a bipartite state $|\psi\rangle^{\mathsf{AB}}$. We get

$$|\psi\rangle = M_\psi \otimes \mathbb{1} \left|\varphi^+_{d\mathsf{B}}\right\rangle = U\Lambda V^\dagger \otimes \mathbb{1} \left|\varphi^+_{d\mathsf{B}}\right\rangle ,$$

and via Ricochet we get

$$U\Lambda \otimes V^* \left|\varphi^+_{d\mathsf{B}}\right\rangle = (U\Lambda \otimes V^*) \sum_{i=1}^{d\mathsf{B}} |i\rangle \otimes |i\rangle .$$

[24] Here, we use the Ricochet property.

[25] Thus, isntead of worrying about partial traces, we have a nice formula for the reduced density matrices of an arbitrary bipartite density.

[26] Note that Λ is nonnegative and diagonal. The nonzero diagonal elements are the *singular values*. The number of nonzero singular values is the rank.

After some computation, we get[27]

$$|\psi\rangle^{\mathsf{AB}} = \sum_{j=1}^{r} \sigma_j \left|\alpha_j\right\rangle^{\mathsf{A}} \otimes \left|\beta_j\right\rangle^{\mathsf{B}}.$$

□

Remark 2.0.9 Look at why this is valuable. The Schmidt decomposition gives us a way to perfectly correlate our A and B subsystems. If we wrote $|\psi\rangle$ in our standard form, we would have two summations.[28]

Corollary 2.0.5 *Taking the Schmidt decomposition, we get*

$$\rho^{\mathsf{A}} = \sum_{i=1}^{r} \sigma_i^2 \, |\alpha_i\rangle\langle\alpha_i|$$

and

$$\rho^{\mathsf{B}} = \sum_{i=1}^{r} \sigma_i^2 \, |\beta_i\rangle\langle\beta_i|\,.$$

Thus, the spectra of ρ^{A}, ρ^{B} are identical.[29]

We now give a brief discussion/application of *entanglement measures*. Such measures quantify "how much" entanglement is in a state.

Example 2.0.8 (Min Schmidt Coefficient) If

$$|\psi\rangle^{\mathsf{AB}} = \sum_{j=1}^{r} \sigma_j \left|\alpha_j\right\rangle^{\mathsf{A}} \otimes \left|\beta_j\right\rangle^{\mathsf{B}},$$

define[30]

$$E_1(|\psi\rangle) := -\min_j \ln \sigma_j^2 = -\ln \sigma_{\min}^2.$$

Note that $E_1(|\psi\rangle) = \ln d\mathsf{A}$ if and only if $\sigma_j^2 = 1/d\mathsf{A}$ for all j, so

$$|\psi\rangle = \frac{1}{\sqrt{r}} \sum_{i=1}^{r} |\alpha_i\rangle \otimes |\beta_i\rangle\,,$$

so $|\psi\rangle$ is *maximally entangled.*

Definition 2.0.10 (Shannon Entropy) *For a probability distribution $\{p(x)\}_{x=1}^{r}$, its Shannon entropy is defined as*[31]

$$H(\{p(x)\}) = -\sum_{x=1}^{r} p(x) \ln p(x).$$

Well,

$$0 \le H(\{p(x)\}) \le \ln r.$$

[27] Let

$$|\alpha_j\rangle := U\,|j\rangle$$

and

$$|\beta_j\rangle := V^*\,|j\rangle\,.$$

[28] Note that the proof of Schmidt is *exactly* applying the SVD.

[29] As a result, the Schmidt rank

$$r \le \min\{d\mathsf{A}, d\mathsf{B}\}.$$

[30] We can show that

$$0 \le E_1(|\psi\rangle) \le \ln d\mathsf{A}.$$

When is it zero? Well, consider a product state $|\psi\rangle$. Then, the Schmidt coefficient is 1, and $\ln 1 = 0$. Thus, the zero case precisely corresponds to product states.

[31] This entropy is a fundamental quantity in informaion theory.

Where $p(x) = 1$ if and only if we have zero, and $p(x) = 1/r$ for all x if and only if we have $\ln r$.

Example 2.0.9 (Entanglement Entropy) We define the entanglement entropy as the Shannon entropy of the squared Schmidt coefficients:

$$E(|\psi\rangle) = -\sum_{j=1}^{r} \sigma_j^2 \ln \sigma_j^2 .$$

Definition 2.0.11 (Product States of Mixed States) *A bipartite density state ρ^{AB} is a product state if it is the tensor product of two density matrices.*

Definition 2.0.12 (Entanglement of Mixed States) *A bipartite state ρ^{AB} is entangled if it is not a convex combination of product states*[32]

$$\rho^{\mathsf{AB}} \neq \sum_i p_i \rho_i^{\mathsf{A}} \otimes \omega_i^{\mathsf{B}} = \sum_\lambda q_\lambda |\alpha_\lambda\rangle\langle\alpha_\lambda| \otimes |\beta_\lambda\rangle\langle\beta_\lambda| .$$

Definition 2.0.13 (Separable State) *A bipartite state ρ^{AB} is called separable if it is a convex combination of product states.*[33]

Remark 2.0.10 It is an NP-Hard question to decide if a given density matrix ρ^{AB} is separable,

[32] In the second equality, we decomposed each ρ_i and ω_i into a convex combination of pure states.

[33] That is, for mixed states, entanglement means "not separable," and the product states are contained within the separable states.

System Evolution Axiom 3

Thus far, we have described the "static" structure of quantum systems. Now, we turn to "dynamical" properties of quantum systems. Now, for time in $t \in [t_1, t_2]$, we say system Q undergoes *closed* evolution if it does not exchange energy with any other system. In contrast, system Q undergoes *open* evolution if it exchanges some energy with another system.

3.1 Closed Evolution

Definition 3.1.1 (System Evolution Axiom) *A quantum system Q undergoing closed evolution is described by a unitary transformation on the state space.*

That is, for closed evolution $t \in [t_1, t_2]$,

$$\rho^Q(t_2) = U\rho^Q(t_1)U^\dagger,$$

where $U \in \mathbb{B}(\mathscr{H}^Q)$.[1]

Remark 3.1.1 Remember, every unitary takes time. Still, we often omit denoting the times t_1 and t_2.

Note that closed evolution is reversible:

$$\rho \mapsto U\rho U^\dagger \mapsto U^\dagger U\rho U^\dagger U = \rho.$$

We often refer to multi-qubit unitaries as *gates*. Let us begin by looking at qubit gates in detail. Suppose a 2×2 unitary U is applied on a qubit system. How does U transform an arbitrary density matrix

$$\rho_0 = \frac{1}{2}(\mathbb{1} + \mathbf{r}_0 \cdot \vec{\sigma}).$$

Well,[2]

$$\rho_1 = U\rho_0 U^\dagger = \frac{1}{2}(\mathbb{1} + \mathbf{r}_1 \cdot \vec{\sigma}).$$

Definition 3.1.2 (Special Unitary Group) *The set* $\mathrm{SU}(2) \leq U(2)$ *is the group of unitary operators in* $\mathbb{B}(\mathbb{C}^2)$ *with determinant* $+1$.

Lemma 3.1.1 *Any unitary* $U \in \mathbb{B}(\mathbb{C}^2)$ *can be written as* $U = e^{i\alpha/2}V$ *with* $V \in \mathrm{SU}(2)$.[3]

[1] Recall that if we have a pure state,

$$|\psi(t_1)\rangle \mapsto |\psi(t_2)\rangle\,,$$

we are really moving from outer product to outer product, meaning the time evolution is precisely conjugating by $U, U^\dagger$.

[2] If we can deduce the relationship between $\mathbf{r}_0$ and $\mathbf{r}_1$, we get a nice geometrical Bloch understanding of unitary evolution.

[3] We have a determinant $\det U = e^{i\alpha/2}$, since eigenvalues of unitaries are always roots of unity.

Proposition 3.1.2 *Any* $U \in \mathrm{SU}(2)$ *can be written as*

$$U_{\hat{n}}(\theta) = e^{-i\frac{\theta}{2}\hat{n}\cdot\vec{\sigma}},$$

where $\hat{n} \in \mathbb{R}^3$ *is a unit vector, and* θ *is an angle.*

Note that $\hat{n}\cdot\vec{\sigma}$ has eigenvalues ± 1 with eigenvectors $|\pm\hat{n}\rangle$. Then,[4]

$$f(\hat{n}\cdot\sigma) = \cos\frac{\theta}{2}\mathbb{1} + i\,\sin\frac{\theta}{2}(\hat{n}\cdot\vec{\sigma}).$$

We can thus think of $\hat{n}$ as an axis of rotation and θ as the angle of rotation.

Definition 3.1.3 (Special Orthogonal Group) *The set* SO(3) *is the group of real orthogonal operators in* $\mathbb{B}(\mathbb{R}^3)$ *with determinant* $+1$.[5]

Theorem 3.1.3 *There exists an isomorphism* $\mathrm{SO}(3) \xrightarrow{\sim} \mathrm{SU}(2)/C_2$. *That is, for any given direction* $\hat{n}$ *and angle* θ, *there is a one-to-two correspondence* $O_{\hat{n}}(\theta) \leftrightarrow \pm U_{\hat{n}}(\theta)$ *between* SO(3) *and* SU(2) *such that*

$$\mathbf{r}_1 = O_{\hat{n}}(\theta)\mathbf{r}_0 \iff \mathbf{r}_1\cdot\vec{\sigma} = U_{\hat{n}}(\theta)(\mathbf{r}_0\cdot\vec{\sigma})U^\dagger_{\hat{n}}(\theta).$$

For instance, let $\hat{n} = \hat{x}$ and $\theta = \pi$. Then,

$$U_{\hat{x}}(\pi) = -i\sigma_x \in \mathrm{SU}(2),$$

so the Pauli X is a rotation of angle π around the x-axis.[6]

Example 3.1.1 Suppose we want to rotation from $|0\rangle \mapsto |+\rangle$. One way to do that is to rotate around the y-axis by $\pi/2$:

$$U_{\hat{y}}(\pi/2) = \cos(\pi/4)\mathbb{1} - i\,\sin(\pi/4)\sigma_y,$$

which is just

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Recall that Hadamard H sends $|0\rangle \mapsto |+\rangle$ and $|1\rangle \mapsto |-\rangle$, so we need another rotation[7]

$$U_{\hat{x}}(\pi)U_{\hat{y}}(\pi/2) = -i\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = -iH.$$

Thus, in practice, the gate sequence is

$$U_{\hat{x}}(\pi)U_{\hat{y}}(\pi/2)\,|0\rangle = i\,|+\rangle\,.$$

Proposition 3.1.4 *Let* $|\hat{m}_1\rangle$ *and* $|\hat{m}_2\rangle$ *be arbitrary qubit states. There exists a unitary* $U \in \mathrm{SU}(2)$ *such that*

$$U\,|\hat{m}_1\rangle = |-\hat{m}_1\rangle$$

[4] Let $f : x \mapsto e^{ix\theta/2}$.

[5] This is the group of rotations on $\mathbb{S}^2$.

[6] Likewise for Y and Z.

[7] We *need* the $-i$ there, as $H \notin \mathrm{SU}(2)$, since $\det H = -1$.

and

$$U\,|\hat{m}_2\rangle = |-\hat{m}_2\rangle\,.$$

Proof. Take the cross product $\hat{n} := (\hat{m}_1 \times \hat{m}_2)\|\hat{m}_1 \times \hat{m}_2\|$. Then, the unitary $U_{\hat{n}}(\pi)$ does what we want. □

Remark 3.1.2 If we are given three non-coplanar vectors, then no such unitary exists in SU(2). Such an operation is called a *universal spin flip*. There does exist an anti-unitary operator A such that $A\,|\hat{n}\rangle = |-\hat{n}\rangle$.[8]

We can now consider a universal gate set for quantum computing. Unlike for classical logic circuits, there are infinitely many unitary gates. However, it is possible to decompose a circuit into a finite set of elementary building blocks.

Theorem 3.1.5 (Solovay-Kitaev Theoren) *For* $U \in \mathrm{SU}(2)$ *and* $\varepsilon > 0$, *there exists a sequence of*

$$n = O\left(\ln^{3+\varepsilon}\frac{1}{\varepsilon}\right)$$

gates chosen from the set $\{\sigma_x, \sigma_y, \sigma_z, H, T\}$ *that approximates* U *within* ε *error*

Proposition 3.1.6 *The Pauli matrices can be generated by* H *and* T.

Proof. Well, $\sigma_z = T^4$, $\sigma_x = H\sigma_z H$, and $\sigma_y = -i\sigma_z\sigma_x$. □

[add notes from tuesday, oct 8]

Example 3.1.2 Let

$$\hat{n} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}\right).$$

Show how to build $U_{\hat{n}}(\theta)$ exactly from

$$\{\sigma_x, \sigma_y, U_{\hat{z}}(\theta), H, T\}_{\theta\in[0,2\pi)}$$

gates.

Proof. Transform $\hat{n} \mapsto \hat{z}$ by rotation $U_{\hat{z}}(-\pi/4)$ and then $U_{\hat{y}}(-\pi/4)$. Now, rotate about the $\hat{z}$ axis with $U_{\hat{z}}(\theta)$. Now, rotate back to the $\hat{n}$ axis by $U_{\hat{y}}(\pi/4)$ and then $U_{\hat{z}}(\pi/4)$.[9] □

Remark 3.1.3 The quantum circuit model describes a standard approach to computing some function $f : (\mathbb{Z}/2)^n \to (\mathbb{Z}/2)^m$ using a quantum computer. The input $b \in (\mathbb{Z}/2)^n$ is encoded in an n-qubit computational basis state:

$$b \mapsto |b\rangle = \bigotimes_{i=1}^{n} |b_i\rangle\,.$$

The function f is encoded into a unitary U_f that reversibly maps $|b\rangle$ to

[8] An anti-unitary is when $A^\dagger A = -\mathbb{1}$. This is not physical.

[9] We will come back to showing that $U_{\hat{y}}(-\pi/4)$ is constructible from our gate set.

$|f(b)\rangle$. Thus, we have an issue when f is not injective. A standard (but not always optimal) form of unitary computation: for all $b \in (\mathbb{Z}/2)^n$ and for all $x \in (\mathbb{Z}/2)^m$,[10]

$$U_f(|b\rangle \otimes |x\rangle) = |b\rangle \otimes |x \oplus f(b)\rangle .$$

Proposition 3.1.7 *We have the equality*[11]

$$|+\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{b \in (\mathbb{Z}/2)^n} |b\rangle .$$

Suppose a user has access to a "black box" that can compute a function $f : (\mathbb{Z}/2)^n \to (\mathbb{Z}/2)$ on a given input b. The *query complexity* of f describes the number of calls an agent must make to the black box to compute $f(b)$ for an arbitrary b. The black box is often dubbed the *oracle*. Is the classical complexity $C(f) \gg Q(f)$, the quantum complexity. This comes down to the *Deutsch-Jozsa* problem. Consider a Boolean function $f : (\mathbb{Z}/2)^n \to (\mathbb{Z}/2)$ that is either constant:

$$f(b) = c \in \{0, 1\},$$

or balanced:[12]

$$f(b) = \begin{cases} 0, & b \in S \\ 1, & b \notin S. \end{cases}$$

The goal is to decide whether f is constant or balanced by making queries to the oracle. Let $N := 2^n$. Then, $C(f) = O(N)$. Using the *eigenstate trick*,

$$O\,|\Phi\rangle = \sqrt{\frac{1}{N}} \sum_{b \in (\mathbb{Z}/2)^n} (-1)^{f(b)}\,|b\rangle ,$$

where

$$|\Phi\rangle = \left(\sqrt{\frac{1}{2}}(|0\rangle + |1\rangle)\right)^{\otimes n} .$$

Note that in general,[13]

$$H^{\otimes n}\,|b\rangle = \sum_{x \in (\mathbb{Z}/2)^n} (-1)^{b \cdot x}\,|x\rangle .$$

This allows us to compute

$$H^{\otimes n} O\,|\Phi\rangle = \frac{1}{N} \sum_{x \in (\mathbb{Z}/2)^n} \sum_{b \in (\mathbb{Z}/2)^n} (-1)^{f(b) \oplus b \cdot x}\,|x\rangle .$$

If $f(b) = c$ for all b, then we get $|0\rangle^{\otimes n}$.[14] On the other hand, if f is balanced, then with probability 0, the outcome $x = 0$ will be measured. The upshot here is that if it is constant, we always produce all zeroes, and if it is balanced, we never do. That is, f is constant if and only if $x = 0$ is measured.

Corollary 3.1.8 $Q(f) = 1 < C(f) = O(N) = O(2^n)$.[15]

[10] Note that $\oplus$ is XOR.

[11] This is trivial by induction.

[12] We say there exists $S \subset (\mathbb{Z}/2)^n$ such that $|S| = 2^{n-1}$.

[13] Define

$$b \cdot x \equiv \sum_{i=1}^{n} b_i x_i \pmod 2.$$

[14] This is computation. Interestingly, this does not depend on c.

[15] However, there exist randomized classical algorithms that can solve this problem with small error. Can we obtain a separation between $C(f)$ and $Q(f)$ even with a bounded error? It turns out, the answer is yes!

3.2 Quantum Channels

What if we instead allow our system Q to exchange energy with another system R? It is open relative to R between $t_1 \le t \le t_2$. The idea is to simply treat the joint system QR as being closed with respect to a larger environment E. Then, QR evolves by some unitary $U \in \mathbb{B}(\mathscr{H}^{QR})$. We now have to consider the *reduced dynamics* of Q by itself:

$$\rho^Q(t_i) = \operatorname{tr}_R(\rho^{QR}(t_i))$$

for $i \in \{1, 2\}$. The key point is that at time t_2, we have

$$\rho^Q(t_2) = \operatorname{tr}_R(U\rho^{QR}(t_1)U^\dagger).$$

This is known as the reduced dynamics picture. Assume further that Q and R are in a tensor product state:[16]

$$\rho^{QR}(t_1) = \rho^Q(t_1) \otimes \omega^R(t_1).$$

Then, reduced dynamics looks like

$$\rho^Q(t_1) \mapsto \rho^Q(t_2) = \operatorname{tr}_R\left(U(\rho^Q(t_1) \otimes \omega^R(t_1))U^\dagger\right),$$

where $\omega^R(t_1) \in \mathscr{D}(\mathscr{H}^R)$.

Example 3.2.1 Suppose the primary system and the environment are both qubits. The environment is initially in $|+\rangle$ and the interaction is described by controlled-σ_z.[17] How does the primary system evolve under the reduced dynamics, if it is initially in the state $|\psi\rangle = \cos\theta\,|0\rangle + \sin\theta\,|1\rangle$?

Solution. Our initial state is

$$|\psi+\rangle = \frac{1}{\sqrt{2}}(\cos\theta\,|00\rangle + \cos\theta\,|01\rangle + \sin\theta\,|10\rangle + \sin\theta\,|11\rangle).$$

Applying CZ gives us

$$\mathsf{CZ}\,|\psi+\rangle = \frac{1}{\sqrt{2}}(\cos\theta\,|00\rangle + \cos\theta\,|01\rangle + \sin\theta\,|10\rangle - \sin\theta\,|11\rangle)^{QR}.$$

Let $|\tau\rangle^{QR} = M_\tau \otimes \mathbb{1}\,|\varphi^+\rangle$ be this state. Then,

$$M_\tau = \frac{1}{\sqrt{2}}\begin{pmatrix}\cos\theta & \cos\theta\\ \sin\theta & -\sin\theta\end{pmatrix},$$

so

$$\operatorname{tr}_R |\tau\rangle\langle\tau|^{QR} = M_\tau M_\tau^\dagger = \begin{pmatrix}\cos^2\theta & 0\\ 0 & \sin^2\theta\end{pmatrix}.$$

Let us look at what happened:

$$|\psi\rangle\langle\psi| = \begin{pmatrix}\cos^2\theta & \cos\theta\sin\theta\\ \cos\theta\sin\theta & \sin^2\theta\end{pmatrix} \mapsto \begin{pmatrix}\cos^2\theta & 0\\ 0 & \sin^2\theta\end{pmatrix}.$$

The off-diagonals vanished! Since we often call these off-diagonals "coherence terms," this process is dubbed *decoherence*.[18] □

[16] That is, they are prepared separately (uncoupled) initially.

[17] This CZ maps

$$\begin{aligned}|00\rangle &\mapsto |00\rangle\\ |01\rangle &\mapsto |01\rangle\\ |10\rangle &\mapsto |10\rangle\\ |11\rangle &\mapsto -|11\rangle.\end{aligned}$$

[18] In some ways, this is even worse. Our final state is like a classical biased coin!

Exercise 3.2.1 Prove that $\mathbb{F}_2 \left|\alpha\right\rangle^A \left|\beta\right\rangle^B = \left|\beta\right\rangle^A \left|\alpha\right\rangle^B$.

Mathematically, we having been doing transformations which are linear maps on the space of linear operators :

$$\mathscr{E}^{Q \to Q'} : \mathbb{B}(\mathscr{H}^Q) \to \mathbb{B}(\mathscr{H}^{Q'}).$$

3.3 Superoperators and Channels

Definition 3.3.1 (Superoperator) *Linear maps on the space of linear operators are known as superoperators.*

Remember, linear maps $\mathscr{H}^Q \to \mathscr{H}^Q$ are called operators and constitute $\mathbb{B}(\mathscr{H}^Q)$. We know consider the set $\mathbb{B}(\mathbb{B}(\mathscr{H}^Q) : \mathbb{B}(\mathscr{H}^{Q'}))$. The identity id_Q is the identity superoperator. Reduced dynamics represent a nice subclass of superoperators with the partial trace form we looked at.

Definition 3.3.2 (Trace-Preserving) *A superoperator* $\mathscr{E} : \mathbb{B}(\mathscr{H}^Q) \to \mathbb{B}(\mathscr{H}^{Q'})$ *is called trace-preserving (TP) if* $\operatorname{tr} \mathscr{E}(X) = \operatorname{tr} X$.

Definition 3.3.3 (Positive) *A superoperator* $\mathscr{E} : \mathbb{B}(\mathscr{H}^Q) \to \mathbb{B}(\mathscr{H}^{Q'})$ *is called positive if* $\mathscr{E}(X) \geq 0$ *for all* $X \geq 0$.

Let Q_k denote a k-dimensional system; i.e., $\mathscr{H}^{Q_k} \simeq \mathbb{C}^k$.

Definition 3.3.4 (k-Positive) *A superoperator* $\mathscr{E}^{Q \to Q'}$ *is* k*-positive if the superoperator*

$$\mathrm{id}_{Q_k} \otimes \mathscr{E}^{Q \to Q'} : \mathbb{B}(\mathscr{H}^{Q_k Q}) \to \mathbb{B}(\mathscr{H}^{Q_k Q'})$$

is positive.[19]

Definition 3.3.5 *A superoperator* $\mathscr{E}$ *is completely positive if it is* k*-positive for all* $k \in \mathbb{Z}_+$.

The superoperator $\mathscr{E}^{Q \to Q'}$ maps operators in $\mathscr{D}(\mathscr{H}^Q)$ to density operators in $\mathscr{D}(\mathscr{H}^{Q'})$ *even when acting on only half of an entangled state*. Reduced dynamics superoperators are CPTP.

Definition 3.3.6 (Quantum Channel) *CPTP superoperators are known as quantum channels.*

Example 3.3.1 (States as Channels) You can think of quantum states as quantum channels! Let Q be a one-dimensional Hilbert space.[20] We can

[19] That is, $\mathrm{id}_{Q_k} \otimes \mathscr{E}(T^{Q_k Q}) \geq \mathbf{0}$ for any positive $T^{Q_k Q}$.

[20] That is, $Q \simeq \mathbb{C}$

think of

$$\mathsf{CPTP}(\mathbb{C} \to Q') = \mathfrak{D}(Q'),$$

where $\mathsf{CPTP}(\mathbb{C} \to Q') \subseteq \mathbb{B}(\mathbb{B}(\mathbb{C}) : \mathbb{B}(Q'))$ is the set of CPTP superoperators between the two spaces' operators.

Example 3.3.2 (Transpose and Partial Transpose) We can think of the transpose map as a superoperator

$$X \mapsto t(X) := X^t,$$

where X^t is a transpose wrt the computation basis. It is clearly trace-preserving, since we do not change the diagonals. Now, let $|\varphi\rangle := \sum_i c_i |i\rangle$. Let $X \geq 0$. Then, for arbitrary $|\varphi\rangle$,

$$\langle\varphi|X^t|\varphi\rangle = \operatorname{tr}\big(X^t\, |\varphi\rangle\langle\varphi|\big) = \operatorname{tr}\big(X\, |\varphi\rangle\langle\varphi|^t\big) = \langle\varphi^*|X|\varphi^*\rangle \geq 0.$$

Thus, t is trace-preserving and positive. Is it completely positive? It turns out, *the answer is no*. Consider $t : \mathbb{B}(\mathbb{C}^2) \to \mathbb{B}(\mathbb{C}^2)$ and $\mathrm{id} : \mathbb{B}(\mathbb{C}^2) \to \mathbb{B}(\mathbb{C}^2)$. Define the bipartite superoperator[21]

$$\Gamma_A := t \otimes \mathrm{id}\,.$$

If we apply Γ_A on $|\Phi^+\rangle\langle\Phi^+|$, then we get $\mathbb{F}_2$, which is *not* positive.[22]

Note that while $t \otimes \mathrm{id}$ is not positive on all $\mathbb{B}(\mathbb{C}^d \otimes \mathbb{C}^d)$, it is positive on the set of *separable* states; that is, those states which have the form

$$\rho^{AB} = \sum_i p_i\, |\alpha_i\rangle\langle\alpha_i|^A \otimes |\beta_i\rangle\langle\beta_i|^B\,.$$

We have

$$\begin{aligned}\Gamma_A(\rho^{AB}) &= \sum_i p_i\, |\alpha_i\rangle\langle\alpha_i|^t \otimes |\beta_i\rangle\langle\beta_i| \\ &= \sum_i p_i\, |\alpha_i^*\rangle\langle\alpha_i^*| \otimes |\beta_i\rangle\langle\beta_i|\,,\end{aligned}$$

which is also separable. Thus, it is a positive operator.

Definition 3.3.7 (PPT) *A state ρ^{AB} is called PPT if $\Gamma_A \rho^{AB} \geq 0$*

Theorem 3.3.1 (PPT Crtierion) *If ρ^{AB} is separable, then its partial transpose Γ_A is positive.*[23]

Corollary 3.3.2 *If ρ^{AB} is not PPT, then ρ^{AB} is entangled!*

Remark 3.3.1 There exist entangled states which are PPT.

[21] This map is called the *partial transpose*.

[22] Hence, t is not CP, and thus, not a quantum channel.

[23] That is, separable implies PPT.

Example 3.3.3 Define

$$\begin{tikzcd} \mathbb{B}(\mathbb{C}^d) \arrow[r, "\varphi"] & \mathbb{B}(\mathbb{C}^d) \\ X \arrow[r, mapsto] & \operatorname{tr}(X)\mathbb{1} - X. \end{tikzcd}$$

Is φ positive? Is φ CP?

Proof. Assume $X \geq 0$. Let $|\psi\rangle$ be arbitrary. Then,

$$\langle\psi|\varphi(X)|\psi\rangle = \operatorname{tr} X\, |\psi\rangle\langle\psi| - \langle\psi|X|\psi\rangle = \operatorname{tr} X - \langle\psi|X|\psi\rangle\,.$$

We claim that $\langle\psi|X|\psi\rangle \leq \operatorname{tr} X$. Complete $|\psi\rangle$ with $\left|\psi_i^{\perp}\right\rangle$ until it is an orthonormal basis. Then,

$$\operatorname{tr} X = \langle\psi|X|\psi\rangle + \sum_{i=2}^{d} \left\langle\psi_i^{\perp}\middle|X\middle|\psi_i^{\perp}\right\rangle \geq \langle\psi|X|\psi\rangle\,.$$

Thus, φ is positive. Now, for CP:[24]

$$\varphi \otimes \operatorname{id}(\left|\Phi^{+}\right\rangle) = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

and the outer block is simply $-\sigma_x$, which is not a positive operator. Thus, φ is *not* a CP superoperator; i.e., not a channel. □

Theorem 3.3.3 (Reduction Criterion) *If we have* $\rho^{AB} \mapsto \varphi \otimes \operatorname{id} \rho^{AB} \not\geq 0$.

It turns out, this criterion is *no stronger* than the PPT one.

Definition 3.3.8 (Choi Matrix) *Let* $\mathscr{E} : \mathbb{B}(\mathscr{H}^B) \to \mathbb{B}(\mathscr{H}^A)$. *Its Choi matrix is the operator* $J_{\mathscr{E}} \in \mathbb{B}(\mathscr{H}^{AB})$ *defined by*

$$J_{\mathscr{E}}^{AB} = \mathscr{E}^{\widetilde{B}\to A} \otimes \operatorname{id}^B(\varphi^{+\widetilde{B}B}),$$

where

$$\varphi^{+\widetilde{B}B} = \sum_{i,j=1}^{\dim B} |ii\rangle\langle jj|^{\widetilde{B}B}\,.$$

Theorem 3.3.4 (Second Canonical Isomorphism) *We have an isomorphism*[25]

$$\begin{tikzcd} \mathbb{B}(\mathbb{B}(\mathscr{H}^B) : \mathbb{B}(\mathscr{H}^A)) \arrow[r, "\simeq"] & \mathbb{B}(\mathscr{H}^A \otimes \mathscr{H}^B) \\ \mathscr{E} \arrow[r, mapsto] & J_{\mathscr{E}}. \end{tikzcd}$$

Sketch of Proof. The action of $\mathscr{E} \in \mathbb{B}(\mathbb{B}(\mathscr{H}^B) : \mathbb{B}(\mathscr{H}^A))$ can be described fully by its Choi matrix. That is, for $X \in \mathbb{B}(\mathscr{H}^B)$,[26]

[24] We try the same thing we did for Γ_A.

[25] This isomorphism goes by a few names, including the Choi-Jamiołkowski isomorphism.

[26] Note that the transpose on X is the same as the basis we use on $\left|\varphi^{+}\right\rangle$.

$$\mathcal{E}(X) = \operatorname{tr}_B \left((\mathbb{1}^A \otimes (X^t)^B) J^A B_{\mathcal{E}} \right).$$

Conversely, if $J^{AB} \in \mathbb{B}(\mathcal{H}^{AB})$, then we can define the superoperator $\mathcal{E}_J : \mathbb{B}(\mathcal{H}^B) \to \mathbb{B}(\mathcal{H}^A)$ by

$$\mathcal{E}_J(X) = \operatorname{tr}_B \left((\mathbb{1}^A \otimes (X^t)^B) J^{AB} \right),$$

which precisely aligns with our definition of the Choi matrix. □

Now, suppose $\mathcal{E} \in \mathsf{CP}(A, B)$. Then,

$$J_{\mathcal{E}}^{AB} = \mathcal{E}^{\widetilde{B} \to A} \otimes \mathrm{id}^B (\varphi^{+\widetilde{B}B}) \geq 0.$$

That is, $\mathcal{E}$ being CP implies that $J_{\mathcal{E}}$ is positive. What about the converse? Let $J \geq 0$. Take the spectral decomposition

$$J^{AB} = \sum_{k=1}^{r} \lambda_k \left|\psi_k\right\rangle\!\left\langle\psi_k\right| = \sum_{k=1}^{r} \left|\widetilde{\psi_k}\right\rangle\!\left\langle\widetilde{\psi_k}\right|.$$

Write $\left|\widetilde{\psi_k}\right\rangle = M_{\widetilde{\psi_k}} \otimes \mathbb{1} \left|\varphi_{dB}^{+}\right\rangle$. Thus, we can write

$$\mathcal{E}_J(X) = \operatorname{tr}_B \left((\mathbb{1}^A \otimes (X^t)^B) \sum_{k=1}^{r} (M_{\widetilde{\psi_k}} \otimes \mathbb{1}^B) \varphi^{+\widetilde{B}B} (M_{\widetilde{\psi_k}}^{\dagger} \otimes \mathbb{1}^B) \right).$$

Using ricochet, we get

$$\mathcal{E}_J(X) = \sum_{k=1}^{r} M_{\widetilde{\psi_k}} X M_{\widetilde{\psi_k}}^{\dagger} = \sum_{k=1}^{r} M_k X M_k^{\dagger}.$$

Then, using this new form, we can see that $\mathcal{E}_J(X)$ is CP

Definition 3.3.9 (Kraus Operators) *We call the $M_k \in \mathbb{B}(\mathcal{H}^B : \mathcal{H}^A)$ the Kraus operators for $\mathcal{E}_J$.*

Theorem 3.3.5 *For a superoperator $\mathcal{E} : \mathbb{B}(\mathcal{H}^A) \to \mathbb{B}(\mathcal{H}^B)$, the following are equivalent:*

(i) $\mathcal{E}$ is completely positive.
(ii) Its Choi operator $J_{\mathcal{E}}^{AB} \geq 0$.
(iii) There exist Kraus operators $\{M_k\}_k$ such that

$$\mathcal{E}(X) = \sum_k M_k X M_k^{\dagger}.$$

Are the Kraus operators of a CP map unique? As it turns out, the answer is *yes, but only up to unitaries*. Since the convex combination of a pure state ensemble is not unique, but unique up to unitary equivalence, the same can be said for the Kraus operators.

Proposition 3.3.6 *We have*[27]

$$\mathscr{E}(\cdot) = \sum_{j=1}^{s} N_j(\cdot)N_j^\dagger = \sum_{k=1}^{r} M_k(\cdot)M_k^\dagger$$

if and only if there exists a unitary matrix $U = (u_{jk})$ *so that*

$$N_j = \sum_k u_{jk} M_k.$$

Thus far, we have a good understanding of what happens when our map is CP, but what about TP? Suppose that $\mathscr{E}$ is trace-preserving. Take a Kraus representation

$$\operatorname{tr}\mathscr{E}(X) = \sum_{k=1}^{r} \operatorname{tr}\left(M_k^\dagger M_k X\right) = \operatorname{tr}\left(\left(\sum_{k=1}^{r} M_k^\dagger M_k\right)X\right),$$

which only works if we meet the completion condition

$$\sum_{k=1}^{r} M_k^\dagger M_k = \mathbb{1}^B.$$

If our map is TP, then $\operatorname{tr}_A J_{\mathscr{E}}^{AB} = \mathbb{1}^B$.

Example 3.3.4 Consider the qubit superoperator that kills all off-diagonal terms:

$$\mathscr{T} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}.$$

What is the Choi matrix of this map. Is the map CP? If so, find an operator-sum decomposition.

Solution. We have $\mathscr{T} : |0\rangle\langle 0| \mapsto |0\rangle\langle 0|$, $\mathscr{T} : |1\rangle\langle 1| \mapsto |1\rangle\langle 1|$, $\mathscr{T} : |0\rangle\langle 1| \mapsto 0$ and $\mathscr{T} : |1\rangle\langle 0| \mapsto 0$. Then,

$$J_{\mathscr{T}} = \mathscr{T} \otimes \mathrm{id}[\varphi^+] = \mathscr{T} \otimes \mathrm{id} \sum |ii\rangle\langle jj| = \sum_{i,j=0}^{1} \mathscr{T}(|i\rangle\langle j|) \otimes |i\rangle\langle j|.$$

Thus, we kill our off-diagonals and get

$$J_{\mathscr{T}} = |0\rangle\langle 0|^{\otimes 2} + |1\rangle\langle 1|^{\otimes 2}.$$

Now, is the map CP?[28] Well, remember our equivalence theorems. We just need to check $J_{\mathscr{E}} \geq 0$. It certainly is positive. Write

$$|00\rangle = M_0 \otimes \mathbb{1}\left|\varphi^+\right\rangle = |0\rangle\langle 0| \otimes \mathbb{1}\left|\varphi^+\right\rangle,$$

which we can just write as

$$M_0 \otimes \mathbb{1}(|00\rangle + |11\rangle).$$

[27] The M_k and N_j are Kraus operators.

[28] Remember, this is the decoherence map.

We can do the same for the other eigenvector:

$$|11\rangle = |1\rangle\langle 1| \otimes \mathbb{1} \left|\varphi^+\right\rangle.$$

Thus,[29]

$$\mathscr{T}(\cdot) = |0\rangle\langle 0| \, (\cdot) \, |0\rangle\langle 0| + |1\rangle\langle 1| \, (\cdot) \, |1\rangle\langle 1| \, .$$

□

Remark 3.3.2 What is the Choi matrix for a unitary transformation?

Well, let $U^{B \to B}$ be a unitary transformation. Take the density matrix interpretation of its action:

$$|\psi\rangle\langle\psi| \mapsto U \, |\psi\rangle\langle\psi| \, U^\dagger.$$

Then,

$$\mathscr{U}(\rho) = U\rho U^\dagger,$$

the channel form of U. We have

$$J_{\mathscr{U}} = \mathscr{U} \otimes \mathrm{id}[\varphi^+] = \left|\hat{\varphi}^+\middle\rangle\middle\langle\hat{\varphi}^+\right| ,$$

where $\left|\widehat{\varphi^+}\right\rangle = U \otimes \mathbb{1} \left|\varphi^+\right\rangle$. What are the Kraus operators? The Kraus operators form a singleton $\{U\}$![30]

Proposition 3.3.7 *The set of Kraus operators of a unitary channel is simply the singleton containing the unitary itself.*

Remark 3.3.3 We have a rk 1 Choi matrix if and only if we have 1 Kraus operator. If our channel is TP, then $M^\dagger M = \mathbb{1}$, so M is unitary.

Does every CPTP map have a physical interpretation?

Theorem 3.3.8 (Stinespring Dilation) *For every CPTP map $\mathscr{E} : Q \to Q'$, there exists a system R, state ω^R, and unitary $U \in \mathbb{B}(\mathscr{H}^{QR})$ so that*

$$\mathscr{E}(X) = \mathrm{tr}_{R'}\left(U(X \otimes \omega^R)U^\dagger\right).$$

The unitary $U : QR \to Q'R'$ is called a unitary dilation of the channel $\mathscr{E}$. In general, for every CP map $\mathscr{E} : Q \to Q'$, there exists a system R, state ω^R, and operator $M \in \mathbb{B}(\mathscr{H}^{QR})$ such that

$$\mathscr{E}(X) = \mathrm{tr}_{R'}\left(M(X \otimes \omega^R)M^\dagger\right).$$

Proof. For $\mathscr{E} \in \mathsf{CP}(Q \to Q')$, consider the Choi matrix $J_{\mathscr{E}}^{Q'Q}$. Purify it:

$$J_{\mathscr{E}}^{Q'Q} = d_Q \, \mathrm{tr}_{\widetilde{Q}\,\widetilde{Q}'}\left(|J_{\mathscr{E}}\rangle\langle J_{\mathscr{E}}|^{Q'Q\widetilde{Q}'\widetilde{Q}}\right).$$

Then, doing a lot of algebra, we find that[31]

[29] This is our Kraus representation.

[30] In other words, our eigenvector is just $\left|\widehat{\varphi^+}\right\rangle$.

[31] We use that $Q \simeq \widetilde{Q} \simeq R'$ and $Q' \simeq \widetilde{Q}' \simeq R$.

$$\mathscr{E}(X) = \mathrm{tr}_{R'}\left(\hat{M}_{J_{\mathscr{E}}}^{QR\to R'Q'}\left(X \otimes \frac{\mathbb{1}^R}{dQ'}\hat{M}_{J_{\mathscr{E}}}^{\dagger QR\to R'Q'}\right)\right).$$

□

3.4 Qubit Channels

We now restrict our attention to qubit channels $\mathsf{CPTP}(\mathbb{B}(\mathbb{C}^2) \to \mathbb{B}(\mathbb{C}^2))$. If $\mathscr{N}$ is a channel transforming $\rho \mapsto \rho'$, where $\rho = 1/2(\mathbb{1} + \mathbf{r}\cdot\vec{\sigma})$. What happens to $\mathbf{r} \mapsto \mathbf{r}'$. For a unitary channel, the pair is related by an SO(3) operation. Now, for a qubit channel $\mathscr{N}$, its Choi matrix is the two-qubit operator

$$J_{\mathscr{N}} = \mathscr{N} \otimes \mathrm{id}(\varphi_2^+),$$

where

$$\varphi_2^+ = \left|\varphi_2^+\right\rangle\!\left\langle\varphi_2^+\right| \quad \text{and} \quad \left|\varphi_2^+\right\rangle = |00\rangle + |11\rangle\,.$$

We can write $J_{\mathscr{N}}$ in the two-qubit Pauli basis. We want $J_{\mathscr{N}} \geq 0$ and $\mathrm{tr}_A(J_{\mathscr{N}}) = \mathbb{1}$. To check the TP, we compute

$$\mathrm{tr}_A(J_{\mathscr{N}}) = a\,\mathrm{tr}(\mathbb{1})\mathbb{1} + \mathrm{tr}\big(\mathbf{m}\cdot\vec{\sigma}\big)\mathbb{1} + \mathrm{tr}(\mathbb{1})\mathbf{n}\cdot\vec{\sigma} + \sum_{i,j=1}^{3} t_{ij}\,\mathrm{tr}(\sigma_i)\sigma_j = 2a\mathbb{1} + 2\mathbf{n}\cdot\vec{\sigma},$$

so $a = 1/2$ and $\mathbf{n} = 0$:

$$J_{\mathscr{N}} = \frac{1}{2}\mathbf{1}\otimes\mathbf{1} + \mathbf{m}\cdot\vec{\sigma} + \sum_{i,j=1}^{3} t_{ij}\sigma_i \otimes \sigma_j \geq 0.$$

Let ρ be as before. Remember, the action of channel $\mathscr{N}$ can be written as

$$\mathscr{N}(\rho) = \mathrm{tr}_B\left((\mathbb{1}\otimes\rho^t)J_{\mathscr{N}}\right),$$

which we compute to be

$$\frac{1}{2}\left(\mathbb{1}\,\mathrm{tr}\big(\rho^t\big) + \mathbf{m}\cdot\vec{\sigma}\,\mathrm{tr}\big(\rho^t\big) + \sum_{i,j=1}^{3} t_{ij}\sigma_i\,\mathrm{tr}\big(\sigma_i\rho^t\big)\right).$$

Now, recall that

$$\mathrm{tr}\big(\sigma_x^t\rho\big) = \mathrm{tr}(\sigma_x\rho) = r_x,$$

and likewise for y and z, so we get

$$\frac{1}{2}\left(\mathbb{1} + \mathbf{m}\cdot\vec{\sigma} + \sum_{i,j=1}^{3} t_{ij}r_j(-1)^{\delta_{j2}}\sigma_i\right),$$

using the Kronecker delta δ_{j2}.[32] Well, notice that the sum is just $\mathbf{s}\cdot\vec{\sigma}$, where

$$s_i = \sum_{j=1}^{3} t_{ij}r_j(-1)^{\delta_{j2}}.$$

[32] We need this to adjust for the negative sign on $-r_y$ from σ_y.

That is,

$$\mathbf{s} = T \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{r} =: T'\mathbf{r}.$$

Thus, the transformed state is

$$\mathcal{N}(\rho) = \frac{1}{2}(\mathbb{1} + (\mathbf{m} + \mathbf{s}) \cdot \vec{\sigma}),$$

so

$$\mathcal{N} : \mathbf{r} \mapsto \mathbf{s} + \mathbf{m} = T'\mathbf{r} + \mathbf{m}.$$

Theorem 3.4.1 *The action of every qubit channel $\mathcal{N}$ on the Bloch sphere is described by an affine transformation*

$$\mathbf{r} \mapsto T'\mathbf{r} + \mathbf{m},$$

where $\mathbf{m}$ is the Bloch vector of the Choi matrix $J_{\mathcal{N}}$ and $T' = T$ times the diagonal matrix above.[33]

Remark 3.4.1 For a unitary transformation, $\mathbf{m} = 0$, so $T' \in \mathrm{SO}(3)$.

Definition 3.4.1 (Partially Depolarizing Channel) *We define $\mathcal{D}_\lambda$ to be defined by*

$$\mathcal{D}_\lambda(X) := \lambda X + (1 - \lambda)\,\mathrm{tr}(X)\frac{\mathbb{1}}{2}.$$

The Choi matrix is

$$J_{\mathcal{D}_\lambda} = \mathcal{D}_\lambda \otimes \mathrm{id}(\varphi^+) = \sum_{i,j=0}^{1} \mathcal{D}_\lambda(|i\rangle\langle j|) \otimes |i\rangle\langle j| = \lambda\varphi^+ + (1 - \lambda)\frac{\mathbb{1}}{2} \otimes \mathbb{1}.$$

Let us take a spectral decomposition of $J_{\mathcal{D}_\lambda}$ to get the Kraus operators. We have that $|\varphi^\pm\rangle = \sqrt{2}\,|\Phi^\pm\rangle$ and $|\psi^\pm\rangle = \sqrt{2}\,|\Psi^\pm\rangle$, so

$$\mathbb{1}^{\otimes 2} = \frac{1}{2}(\varphi^+ + \varphi^- + \psi^+ + \psi^-).$$

Thus,[34]

$$J_{\mathcal{D}_\lambda} = \left(\frac{1 + 3\lambda}{4}\right)\varphi^+ + \frac{1 - \lambda}{4}(\varphi^- + \psi^+ + \psi^-).$$

Hence, our Kraus representation is

$$\mathcal{D}_\lambda(X) = aX + b\sigma_x X \sigma_x + b\sigma_y X \sigma_y + b\sigma_z X \sigma_z,$$

where

$$a = \frac{1 + 3\lambda}{4} \quad \text{and} \quad b = \frac{1 - \lambda}{4}.$$

Note that when $\lambda = 0$,[35]

$$\mathcal{D}_\lambda(X) = \frac{1}{4}(X + \sigma_x X \sigma_x + \sigma_y X \sigma_y + \sigma_z X \sigma_z) = \frac{\mathbb{1}}{2}.$$

[33] Remember, T is the correlation matrix of $J_{\mathcal{N}}$.

[34] Recall that

$$|\Phi_{ij}\rangle = \sigma_z^i \sigma_x^j \otimes \mathbb{1}\,|\Phi_{00}\rangle\,.$$

[35] This is the maximally mixed state, which is precisely what we *do not want*.

We have

$$t_{ij} = \mathrm{tr}\big(\sigma_i \otimes \sigma_j \varphi^+\big) = \mathrm{tr}\big(\sigma_i \sigma_j^t\big) = \begin{cases} 0, & i \neq j \\ 2, & i = j \in \{x, z\} \\ -2, & i = j = y. \end{cases}$$

via ricochet on σ_j. Thus,

$$J_{\mathcal{D}_\lambda} = \frac{1}{2}(\mathbb{1}^{\otimes 2} + \lambda(\sigma_x^{\otimes 2} - \sigma_y^{\otimes 2} + \sigma_z^{\otimes 2})).$$

Since $\mathbf{m} = 0$, we get that the effect of the channel on the Bloch sphere is $\mathbf{r} \mapsto T'\mathbf{r} = \lambda \mathbf{r}$.

Remark 3.4.2 Geometrically, taking our $\mathbf{r} \mapsto T'\mathbf{r} = \lambda\mathbf{r}$ action by $\mathcal{D}_\lambda$, we see that the partially depolarizing channel shrinks the Bloch sphere by λ.[36]

Definition 3.4.2 (Partially Dephasing Channel) *We define*

$$\Delta_\lambda(X) = \lambda X + (1-\lambda)\Delta(X),$$

where $\Delta(X) = |0\rangle\langle 0|\, X\, |0\rangle\langle 0| + |1\rangle\langle 1|\, X\, |1\rangle\langle 1|$ *is the complete dephasing (or decohering map).*

The Choi matrix of Δ_λ is given by

$$J_{\Delta_\lambda} = \lambda \varphi^+ + (1-\lambda)\left(\sum_{i=0}^{1} |i\rangle\langle i| \otimes |i\rangle\langle i|\,.\right)$$

To get the Kraus representation we will again use the spectral decomposition:

$$J_{\Delta_\lambda} = \left(\lambda + \frac{1-\lambda}{2}\right)\varphi^+ + \frac{1-\lambda}{2}\varphi^-.$$

Once again, we take the matrix representations and get the Kraus representation

$$\Delta_\lambda(X) = aX + b\sigma_z X \sigma_z,$$

where $a = (1+\lambda)/2$ and $b = (1-\lambda)/2$. Lastly, for the affine transformation, we can write

$$J_{\mathcal{D}_\lambda} = \frac{1}{2}(\mathbb{1}^{\otimes 2} + \lambda(\sigma_x^{\otimes 2} - \sigma_y^{\otimes 2}) + \sigma_z^{\otimes 2}).$$

Thus, our correlation matrix is[37]

$$T = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so

$$\mathbf{r} \mapsto T'\mathbf{r} = (\lambda r_x, \lambda r_y, r_z).$$

[36] Since we are shrinking by λ, when $\lambda = 0$ we get the center of the sphere, as desired.

[37] Once again, $\mathbf{m} = \mathbf{0}$. When $\mathbf{m} \neq \mathbf{0}$, the sphere also *shifts* in space.

Remark 3.4.3 This map, geometrically, takes the Bloch sphere to an ellipsoid around z-axis, shrinking the Bloch "ellipsoid" inward by λ, keeping the north and south poles fixed.

Definition 3.4.3 (General Pauli Channel) *A general Pauli channel $\mathcal{P}$ is defined by*[38]

$$\mathcal{P}(X) = p_{00}X + p_{01}\sigma_x X\sigma_x + p_{10}\sigma_z X\sigma_z + p_{11}\sigma_y X\sigma_y,$$

where $p_{00} + p_{01} + p_{10} + p_{11} = 1.$

The Kraus operators are just $\{\sqrt{p_{00}}\mathbb{1}, \sqrt{p_{01}}\sigma_x, \sqrt{p_{10}}\sigma_z, \sqrt{p_{11}}\sigma_y\}$. The affine transformation of Bloch vectors is obtained by just writing the Choi matrix in the Pauli basis. Well, we can write

$$\varphi^+ = \frac{1}{2}\left(\mathbb{1}^{\otimes 2} + \sigma_x^{\otimes 2} - \sigma_y^{\otimes 2} + \sigma_z^{\otimes 2}\right).$$

We can compute $\psi^+, \varphi^-, \psi^-$ in a similar fashion, just conjugating φ^+ by $(\sigma_i \otimes \mathbb{1})$ for $i \in \{x, z, y\}$, respectively. Then,

$$J_{\mathcal{P}} = \frac{1}{2}\left(\mathbb{1}^{\otimes} + t_{11}\sigma_x^{\otimes 2} + t_{22}\sigma_y^{\otimes 2} + t_{33}\sigma_z^{\otimes 2}\right),$$

where

$$\begin{aligned} t_{11} &= p_{00} + p_{01} - p_{10} - p_{11} \\ t_{22} &= -p_{00} + p_{01} + p_{10} - p_{11} \\ t_{33} &= p_{00} - p_{01} + p_{10} - p_{11}. \end{aligned}$$

Thus, the Bloch vector transforms as

$$\mathbf{r} \mapsto \begin{pmatrix} t_{11}r_x \\ -t_{22}r_y \\ t_{33}r_z \end{pmatrix}.$$

Proposition 3.4.2 *Every Pauli channel is a unital channel; i.e., $\mathcal{P}(\mathbb{1}) = \mathbb{1}$.*[39]

The *partially depolarizing* and *partially dephasing* channels are special cases of general Pauli channels.

Example 3.4.1 (Amplitude Damping Channel) The amplitude damping channel is a non-unital qubit channel with Kraus representation

$$\mathcal{A}_\lambda(X) = M_0 X M_0^\dagger + M_1 X M_1^\dagger,$$

where

$$M_0 = |0\rangle\langle 0| + \sqrt{1-\lambda}\,|1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{pmatrix}$$

and

$$M_1 = \sqrt{\lambda}\,|0\rangle\langle 1| = \begin{pmatrix} 0 & \sqrt{\lambda} \\ 0 & 0 \end{pmatrix}.$$

[38] We interpret this as saying "with probability $p_{b_0 b_1}$, the channel input X incurs a Pauli error $\sigma_z^{b_0}\sigma_x^{b_1}$."

[39] That is, $\mathbf{0} \mapsto \mathbf{0}$, leaving the origin of the Bloch sphere invariant.

The Choi matrix is given by[40]

$$J_{\mathscr{A}_\lambda} = \left(|00\rangle + \sqrt{1-\lambda}\,|11\rangle\right)\left(\langle 00| + \sqrt{1-\lambda}\,\langle 11|\right) + \lambda\,|01\rangle\!\langle 01|\,.$$

We express the Choi matrix in the Pauli basis by computing the components

$$m_i = \frac{1}{2}\operatorname{tr}\left(J_{\mathscr{A}_\lambda}(\sigma_i \otimes \mathbb{1})\right)$$

and

$$t_{ij} = \frac{1}{2}\operatorname{tr}\left(J_{\mathscr{A}_\lambda}(\sigma_i \otimes \sigma_j)\right).$$

We can read off the trace as a trace out on Bob's side and then a trace out on Alice's side. Thus, the channel $\mathscr{A}_\lambda$ acts by

$$\mathbf{r} \mapsto \begin{pmatrix} \sqrt{1-\lambda}r_x \\ \sqrt{1-\lambda}r_y \\ r_z(1-\lambda)+\lambda \end{pmatrix}.$$

When $\lambda = 0$, it does nothing. When $\lambda = 1$, we map $\mathbf{r} \mapsto \hat{k}$. In general, between 0 and 1, we have the Bloch sphere squishing and shrinking upwards as a point at the north pole.

Remark 3.4.4 (Replacement Channel) One neat channel (which may be on a future exam) is the replacement channel[41]

$$\mathscr{R}_\lambda(X) = \lambda X + (1-\lambda)\operatorname{tr}[X]\omega.$$

Consider an arbitrary qubit channel $\mathscr{N}$. Suppose we perform a unitary $U \in \mathrm{SU}(2)$ to the channel and another unitary $V \in \mathrm{SU}(2)$ after the output: $\mathscr{N}' := V\mathscr{N}U$. How are the Choi matrices of $\mathscr{N}$ and $\mathscr{N}'$ related? Well, using richochet on the wires, $J_{\mathscr{N}'} = (V \otimes U^t)J_{\mathscr{N}}(V \otimes U^t)^\dagger$. Well, we can identify $V \in \mathrm{SU}(2) \mapsto R \in \mathrm{SO}(3)$ and $U^t \in \mathrm{SU}(2) \mapsto S \in \mathrm{SO}(3)$, so

$$(V \otimes U^t)\left(\sum_{i,j=1}^{3} t_{ij}\sigma_i \otimes \sigma_j\right)(V \otimes U^t)^\dagger$$

can be written as

$$\sum_{i,j=1}^{3} t_{ij} \sum_{k,l=1}^{3} r_{ki}s_{lj}(e_k \cdot \vec{\sigma}) \otimes (e_l \cdot \vec{\sigma}).$$

Rewriting again, we get[42]

$$\sum_{k,l=1}^{3} t'_{kl}\sigma_k \otimes \sigma_l = [[RTS^t]]_{kl}.$$

Remark 3.4.5 Thus, pre- and post- SU(2) rotations on $\mathscr{N}$ correspond to left and right SO(3) rotations on the correlation matrix of $J_{\mathscr{N}}$.

By the SVD of T, there exist orthogonal R, S^t so that RTS^t is *diagonal* with

[40] We just switch Bob's bra to a ket.

[41] We can just have ω to be arbitrary.

[42] We substitute
$$t'_{kl} = \sum_{i,j}^{3} r_{ki}t_{ij}s_{lj}.$$

non-negative diagonal elements.

Theorem 3.4.3 (Canonical Form of Qubit Channels) *Every qubit channel $\mathcal{N}$ can be transformed into canonical form $\mathcal{N}'$ by applying pre- and post-* SU(2) *rotations such that $\mathcal{N}'$ has a Choi matrix of the form*

$$J_{\mathcal{N}'} = \frac{1}{2}\left(\mathbb{1}^{\otimes 2} + \mathbf{m}'\cdot\vec{\sigma}\otimes\mathbb{1} + \sum_{i=1}^{3}\sqrt{t_i}\sigma_i\otimes\sigma_i\right).$$

The canonical channel $\mathcal{N}'$ induces an affine transformation on the Bloch sphere given by

$$\mathbf{r} \mapsto \begin{pmatrix} \sqrt{t_1} & 0 & 0 \\ 0 & -\sqrt{t_2} & 0 \\ 0 & 0 & \sqrt{t_3} \end{pmatrix}\mathbf{r} + \mathbf{m}.$$

Corollary 3.4.4 *Every unital qubit channel can be transformed into a Pauli channel by a pre- and post- unitary transformation.*

Proof. If $\mathbf{m} = 0$, then $\mathbf{m}' = 0$, and via our $J_{\mathscr{P}}$ of general Pauli channels, we have our result.[43] □

3.5 Pauli Twirling

Consider an arbitrary qubit channel $\mathcal{N}$. Suppose we choose a random Pauli and apply it both before and after the channel: $\mathcal{N}' := \sigma_z^{b_0}\sigma_x^{b_1}\mathcal{N}\sigma_z^{b_0}\sigma_x^{b_1}$. This operation is called *Pauli twirling*. What does the resulting channel $\mathcal{N}'$ look like?[44]

Remark 3.5.1

(i) For all σ_k,

$$\sum_{b_0,b_1=0}^{1}(\sigma_z^{b_0}\sigma_x^{b_1})\sigma_k(\sigma_z^{b_0}\sigma_x^{b_1}) = 0,$$

since σ_k commutes with exactly 2 elements in $\mathscr{P}$ generators and anti-commutes with the other 2 elements.

(ii) For all $\sigma_k \neq \sigma_l$,

$$\sum_{b_0,b_1=0}^{1}(\sigma_z^{b_0}\sigma_x^{b_1}\otimes\sigma_z^{b_0}\sigma_x^{b_1})\sigma_k\otimes\sigma_k(\sigma_z^{b_0}\sigma_x^{b_1}\otimes\sigma_z^{b_0}\sigma_x^{b_1}) = 0,$$

since $\sigma_k\otimes\sigma_k$ commutes with exactly 2 elements in

$$\{\mathbb{1}\otimes\mathbb{1}, \sigma_x\otimes\sigma_x, \sigma_y\otimes\sigma_y, \sigma_z\otimes\sigma_z\}.$$

(iii) For all σ_k,

$$\sum_{b_0,b_1=0}^{1}(\sigma_z^{b_0}\sigma_x^{b_1}\otimes\sigma_z^{b_0}\sigma_x^{b_1})\sigma_k\otimes\sigma_k(\sigma_z^{b_0}\sigma_x^{b_1}\otimes\sigma_z^{b_0}\sigma_x^{b_1}) = 4\sigma_k\otimes\sigma_k.$$

[43] It is for these reasons that our understanding of Pauli channels gives us information about all unital channels.

[44] Each bit b_0, b_1 is chosen with uniform probability $1/2$.

Theorem 3.5.1 *Pauli twirling an arbitrary qubit channel $\mathcal{N}$ transforms it into a Pauli channel $\mathcal{P}$, whose Choi matrix is*

$$J_{\mathcal{P}} = \frac{1}{2}\left(\mathbb{1}^{\otimes 2} + \sum_{i=1}^{3} t_{ii}\sigma_i \otimes \sigma_i\right),$$

with $t_{ii} = 2\,\mathrm{tr}(\sigma_i \otimes \sigma_i\, J_{\mathcal{N}})$.

Remark 3.5.2 For an arbitrary channel $\mathcal{N}$ which is unknown, we can first transform it into a Pauli channel by twirling, and then try and correct the noise of the simpler problem.

Measurement Axiom 4

This is the fourth, and final axiom we will consider. It is also perhaps the most *controversial*, forcing a sort of stochastic nature on our understanding of measurable quantities. Now, at some point of every experiment, we eventually connect our quantum system to a measurement apparatus to extract *classical data*.[1] Abstractly, we will consider a quantum measurement as a process which *extracts information from a quantum system*.

4.1 Registers

Definition 4.1.1 (Classical Register) *Every quantum measurement generates classical data, and we refer to the system* X *recording this data as a classical register.*[2]

Remark 4.1.1 What happens if someone does not have access to a classical register X or know its state? The best we can do is assign a probability to the different outcomes based on the individual's *state of knowledge*. In other words, the individual describes the state of the classical register X by a random variable X with distribution p_X over alphabet $\mathcal{X}$. Alright, but then how to we decide the probability values $p_X(x)$ for $x \in \mathcal{X}$.

The measurement axiom tells us how to associate a probability distribution to our "read-out" process. States of classical registers are always represented by density matrices *diagonal* in the computational basis:

$$\rho^X = \sum_{x \in \mathcal{X}} p_X(x) \, |x\rangle\langle x|^{\mathsf{X}} .$$

Classical random variables are *always* diagonal. This is a convex combination of the outcomes $|x\rangle\langle x|$.

Example 4.1.1 (Classical States) Suppose at time t_1 Alice and Bob witness the rolling of a die land at four. We can model their classical registers via $\rho^{\mathsf{X}} = |4\rangle\langle 4|$ for both. At time t_2, Alice's memory becomes fuzzy, and she only remembers within ±1 of the outcome. Thus, her matrix becomes

$$\rho^{\mathsf{X}} = \frac{1}{3}(|3\rangle\langle 3| + |4\rangle\langle 4| + |5\rangle\langle 5|),$$

whereas Bob remains the same. Then, at time t_3, Bob's memory also becomes fuzzy, so his state of knowledge becomes

$$\rho^{\mathsf{X}} = \frac{1}{3}(|2\rangle\langle 2| + |4\rangle\langle 4| \, |6\rangle\langle 6|),$$

only remembering the number was even.

[1] That is, anything representable as finite strings of data.

[2] We will denote the set of possible outcomes as $\mathcal{X}$ with classical registers X, Y.

In general, a change of information is described by a classical channel with transitional probabilities $p(x' \mid x)$.

Example 4.1.2 (CQ States) Now, suppose a quantum state ρ_x of system Q is prepared whenever x is rolled on the die. We have two pieces of information we need to store: the classical value $x \in [6]$ and the quantum state $\rho_x \in \mathscr{D}(Q)$. We represent both systems by a classical-quantum (CQ) state

$$\rho^{\mathsf{X}Q} = \sum_{x\in\mathcal{X}} p_X(x)\, |x\rangle\langle x|^{\mathsf{X}} \otimes \rho_x^Q .$$

The rule of thumb is to average over the different outcomes. Then, Alice (as in the previous example) has the descriptions

$$|4\rangle\langle 4| \otimes \rho_4 \mapsto \frac{1}{3}(|3\rangle\langle 3| \otimes \rho_3 + |4\rangle\langle 4| \otimes \rho_4 + |5\rangle\langle 5| \otimes \rho_5),$$

whereas Bob has

$$|4\rangle\langle 4| \otimes \rho_4 \mapsto \frac{1}{3}(|2\rangle\langle 2| \otimes \rho_2 + |4\rangle\langle 4| \otimes \rho_4 + |6\rangle\langle 6| \otimes \rho_6).$$

The most general type of quantum measurement can be seen as a Q-to-QC *stochastic* mapping called a *quantum instrument*:

$$\rho^Q \mapsto \rho_x^{Q'} \otimes |x\rangle\langle x|^{\mathsf{X}} \qquad \text{with probability } p(x).$$

If we model the state of the classical register as being unknown, then the instrument is a *deterministic* mapping[3]

$$\rho^Q \mapsto \sum_{x\in\mathcal{X}} p_X(x) \rho_x^{Q'} \otimes |x\rangle\langle x|^{\mathsf{X}} .$$

Physically, it is built in four steps:

(i) Introduce ancilla system R.
(ii) Unitary evolution.
(iii) Projective measurement.
(iv) Output QC.

This is helpful because we can understand a quantum instrument as a generalization of a quantum channel in which we receive a classical output.[4]

Definition 4.1.2 (Measurement Axiom) *Every measurement of a (finite-dimensional) quantum system Q is described by a set of orthogonal projectors $\{P_x\}_{x=1}^r$ (that is, $P_x P_y = \delta_{xy} P_X$) such that $\sum_{x=1}^r P_x = \mathbb{1}^Q$. That is to say, we have a decomposition*

$$\mathscr{H}^Q = \bigoplus_{i=1}^{r} \mathscr{H}_i , \qquad \textit{with projector } P_i .$$

If ρ is the state of Q prior to measurement, then with probability $p(x) = \mathrm{tr}[P_x \rho]$,

[3] In fact, this is a quantum channel!

[4] That is, tr_{X} gives us our channel.

the post-measurement state will be

$$\rho_x = \frac{P_x \rho P_x}{p(x)}.$$

If the projectors P_x are rank-one projectors, $P_x = |\varphi_x\rangle\langle\varphi_x|$, then $\mathscr{B} := \{|\varphi_x\rangle\}_{x=1}^{dQ}$ forms an orthogonal basis for $\mathscr{H}^{Q}$. We then say we "measure in the $\mathscr{B}$-basis."

Example 4.1.3 For example, measuring in the computational basis is a projective measurement with $P_x = |x\rangle\langle x|$.

The measurement axiom describes an inherently stochastic transformation

$$\rho \mapsto \rho_x = \frac{P_x \rho P_x}{p(x)}, \quad \text{with } p(x) = \mathrm{tr}[P_x \rho].$$

Including a classical register that store the measurement outcome, we have

$$\rho \mapsto \rho_x \otimes |x\rangle\langle x|^{\mathsf{X}}.$$

From the perspective of someone not learning x, we have that[5]

$$\rho \mapsto \sum_{x \in \mathcal{X}} P_x \rho P_x \otimes |x\rangle\langle x|^{\mathsf{X}}.$$

Remark 4.1.2 The equation above for ρ's transformation is the description of both the Q and C systems. The state of just Q is

$$\sum_{x \in cal X} P_x \rho P_x.$$

Example 4.1.4 (Qubit I) A projective measurement in the $\hat{n}$ direction:

$$|\pm\hat{n}\rangle\langle\pm\hat{n}| = P_{\pm\hat{n}} = \frac{1}{2}(\mathbb{1} \pm \hat{n} \cdot \vec{\sigma}).$$

Suppose initially that the qubit system is $|\hat{m}\rangle$. Well, recall that $|\langle\hat{m}|\pm\hat{n}\rangle|^2$ is[6]

$$\mathrm{tr}\,(|\hat{m}\rangle\langle\hat{m}| \pm |\hat{n}\rangle\langle\hat{n}|),$$

and plugging in our Pauli form for $|\hat{n}\rangle\langle\hat{n}|$ gives us

$$\frac{1}{2}(1 \pm \hat{m} \cdot \hat{n}).$$

Now, if we have an arbitrary state

$$\rho = \frac{1+\lambda}{2}\,|{+\hat{m}}\rangle\langle{+\hat{m}}| + \frac{1-\lambda}{2}\,|{-\hat{m}}\rangle\langle{-\hat{m}}|,$$

so

$$p(\pm) = \mathrm{tr}[P_{\pm\hat{n}}\rho] = \frac{1 \pm \lambda\hat{m}\cdot\hat{n}}{2}.$$

[5] We average over the different outcomes and plug in ρ_x.

[6] Via the measurment axiom, this is precisely $p(\pm\hat{n})$.

Example 4.1.5 (Qubit II) Suppose Alice measures in spin direction $\hat{n}$ on her qubit state $\rho = (1/2)(|0\rangle\langle 0| + |+\rangle\langle +|)$. Denote her $\pm$ measurement outcome by $a = +1$ and $a = -1$, respectively. What is the probability distribution of outcome a?

Solution. Using the trace formula, we get

$$p(a = +1) = \operatorname{tr}(|\hat{n}\rangle\langle\hat{n}|\,\rho) = \langle\hat{n}|\rho|\hat{n}\rangle\,,$$

and rewriting gives us

$$\frac{1}{2}|\langle\hat{n}|0\rangle|^2 + \frac{1}{2}\,\langle\langle\hat{n}|+\rangle|^2 = \frac{1}{4}(1+n_z) + \frac{1}{4}(1+n_x) = \frac{1}{2}\Big(1 + \frac{n_z + n_x}{2}\Big).$$

What about $p(a = -1) = \operatorname{tr}(|-\hat{n}\rangle\langle -\hat{n}|\,\rho)$? Well, looking back at our computation, it is just

$$\frac{1}{2}\Big(1 - \frac{n_z + n_x}{2}\Big),$$

since we need the two to add up to one.[7] □

4.2 Wave Function Collapse and Observables

Since the projectors P_x in the measurement axiom are orthogonal and satisfy the completion relation, they are subspace projectors for the direct sum decomposition

$$\mathcal{H}^Q = V_1 \oplus V_2 \oplus \cdots \oplus V_r,$$

where P_x is the projector onto V_x. Any $|\psi\rangle^Q$ can be uniquely decomposed as[8]

$$|\psi\rangle = \sum_{x=1}^{r} \sqrt{p(x)}\,|\varphi_x\rangle\,.$$

Projecting down onto V_y, we get

$$P_y\,|\psi\rangle = \sum_{x=1}^{r} \sqrt{p(x)}P_y\,|\varphi_x\rangle = \sqrt{p(y)}\,|\varphi_y\rangle,$$

so $\langle\psi|P_y|\psi\rangle = p(y)$.

Remark 4.2.1 Thus, with probability $p(y)$, the measurement will *collapse* $|\psi\rangle$ into its V_y component $|\varphi_y\rangle$.

Historically, measurement has been described in terms of observables.

Definition 4.2.1 (Quantum Observable) *A quantum observable for a system Q is a hermitian operator K acting on $\mathcal{H}^Q$.*

[7] Remembering that these are probabilities can make computations way quicker.

[8] Here, $|\varphi_x\rangle \in V_x$.

We get a spectral decomposition

$$K = \sum_{k=1}^{r} \lambda_k P_k,$$

and to measure observable K means to perform the projective measurement with projectors $\{P_k\}_k$ and store the outcome λ_k in the classical register. Measuring collapses the state ρ into an eigenspace of the observable. Thus, changing the eigenvalues gives you a different observable, but one that is physically equivalent after relabeling the values on the classical register. That is, we take

$$K' = \sum_{k=1}^{r} \lambda'_k P_k,$$

we get a physically identical measurement.[9]

Example 4.2.1 For instance, every unit vector $\hat{n}$ in $\mathbb{R}^3$ identifies a spin observable

$$\hat{n} \cdot \vec{\sigma} = |{+\hat{n}}\rangle\langle{+\hat{n}}| - |{-\hat{n}}\rangle\langle{-\hat{n}}|\,.$$

Outcome $\pm\hat{n}$ is associated with the classical value ± 1. Yet, we could consider the observable

$$S = 45\,|{+\hat{n}}\rangle\langle{+\hat{n}}| + 46\,|{-\hat{n}}\rangle\langle{-\hat{n}}|\,.$$

These outcomes are physically identical.

Definition 4.2.2 (Expectation Value) *Consider observable* $K = \sum_{k=1}^{r} \lambda_k P_k$*. Its expectation value is given by*

$$\mathsf{E}[K]_\rho := \sum_{k=1}^{r} \lambda_k\, p(\lambda_k) = \operatorname{tr}(K\rho).$$

Definition 4.2.3 (Variance) *The variance is given by*

$$\operatorname{Var}[K]_\rho := \sum_{k=1}^{r} p(\lambda_k)(\lambda_k - \mathsf{E}[K])^2.$$

Remark 4.2.2 (Zero Uncertainty) Variance measures the average deviation from the expectation value. Measuring observable K on state ρ has *zero uncertainty* if and only if $\operatorname{Var}[K]_\rho = 0$.

Per usual, some algebra shows us

$$\operatorname{Var}[K]_\rho = \sum_{k=1}^{r} \operatorname{tr}[P_k\rho](\lambda_k - \mathsf{E}[K])^2 = \mathsf{E}[K^2] - \mathsf{E}[K]^2$$

Example 4.2.2 Suppose Alice and Bob share the singlet state $|\Psi^-\rangle =$

[9] This is exactly just relabeling things.

$\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. Alice measures in spin direction $\hat{n}$ and Bob measures in spin direction $\hat{m}$. Denote their measurement outcomes by $a, b \in \{0, 1\}$. What is the expected product of their outcomes ab.

Solution. Define the product observable $K = P_{+1} - P_{-1}$. We have

$$\begin{aligned} P_+ :=& \frac{1}{2}(\mathbb{1} + \hat{n} \cdot \vec{\sigma}) \otimes \frac{1}{2}(\mathbb{1} - \hat{m} \cdot \vec{\sigma}) \\ &+ \frac{1}{2}(\mathbb{1} - \hat{n} \cdot \vec{\sigma}) \otimes \frac{1}{2}(\mathbb{1} - \hat{m} \cdot \vec{\sigma}), \end{aligned}$$

and P_- is the cross terms of parity. Thus,

$$K = \hat{n} \cdot \vec{\sigma} \otimes \hat{m} \cdot \vec{\sigma}.$$

Using ricochet, plus some algebra, we find

$$\mathsf{E}[K] = \frac{1}{2} \operatorname{tr}\left(\sigma_y (\hat{n} \cdot \vec{\sigma}) \sigma_y (\hat{m} \cdot \vec{\sigma})^t\right),$$

and using the anti-commutation relation of the Paulis, this is precisely

$$-\frac{1}{2} \operatorname{tr}((\hat{n} \cdot \vec{\sigma})(\hat{m} \cdot \vec{\sigma}) = -\hat{m} \cdot \hat{n}.)$$

□

4.3 Quantum Instruments and POVMs

Using the four steps of the quantum instruments we mentioned before, we get a CP map

$$\mathcal{E}_x(\rho) = \operatorname{tr}_{R'}(P_x U(\rho \otimes \omega^R) U^\dagger P_x).$$

Definition 4.3.1 (Quantum Instrument) *A quantum instrument is a collection of CP maps $\{\mathcal{E}_x\}_x$ such that $\sum_x \mathcal{E}_x$ is trace-preserving.*[10]

Remark 4.3.1 (Instrument Performance) Performing instrument $\{\mathcal{E}_x\}_x$ on state ρ

(i) outputs classical value x with probability $p(x) = \operatorname{tr}(\mathcal{E}_x(\rho))$.
(ii) outputs the post-measurement quantum state $p_x = \frac{1}{p(x)}\mathcal{E}_x(\rho)$.

$$\rho \mapsto \frac{1}{p(x)}\mathcal{E}_x(\rho) \otimes |x\rangle\langle x|^{\mathsf{X}}.$$

Well, what is the point? Using quantum instruments we can extract classical information and evolve our system in ways that are *not possible* using projective measurements. Remember, if the classical register is unknown,

$$\rho \mapsto \sum_x \mathcal{E}_x(\rho) \otimes |x\rangle\langle x|^{\mathsf{X}}.$$

[10] Instruments were introduced in the 70s, but they have only caught on recently. It is a bit of a departure from the classical view of system evolution. Yet, it aligns very well with the way we have approached QIP.

Remark 4.3.2 There are two extremes:

(i) Completely ignore the classical output and any correlation it has with the quantum output.

$$\rho \mapsto \sum_x \mathscr{E}_x(\rho).$$

(ii) Completely ignore the quantum output and any correlation it has with the classical output.[11]

$$\rho \mapsto \sum_x \operatorname{tr}(\mathscr{E}_x(\rho)) \, |x\rangle\langle x|^{\mathsf{X}} = \sum_x p(x) \, |x\rangle\langle x|^{\mathsf{X}} .$$

Definition 4.3.2 (Positive Operator-Valued Measure (POVM)) *A POVM is a mathematical object used to describe the classical outcomes of a quantum measurement.*

$$\rho \mapsto \sum_{x \in \mathcal{X}} \operatorname{tr}(\mathscr{E}_x(\rho)) \, |x\rangle\langle x|^{\mathsf{X}} .$$

Essentially, the quantum measurement is defining a probability measure $p(x)$ over the output set $\mathcal{X}$. A POVM is a *representation* of this measure in terms of positive operators. Take an operator-sum representation

$$\mathscr{E}_x(\rho) = \sum_k M_{x,k} \rho M_{x,k}^\dagger .$$

Then,

$$p(x) = \operatorname{tr}(\mathscr{E}_x(\rho)) = \operatorname{tr}(\Pi_x \rho),$$

where[12]

$$\Pi_x := \sum_k M_{x,k}^\dagger M_{x,k} .$$

Moreover, since $\sum_x \mathscr{E}_x$ is TP we have $\sum_x \Pi_x = \mathbb{1}$.

Definition 4.3.3 (POVM Generalized Born's Rule) *A POVM on system Q is a collection of positive operators $\{\Pi_x\}_x$ such that $\sum_x \Pi_x = \mathbb{1}^Q$.*[13] *Measuring POVM $\{\Pi_x\}_x$ on a state ρ returns outcome x with probability*

$$p(x) = \operatorname{tr}(\Pi_x \rho).$$

Example 4.3.1 Consider the four qubit states

$$\begin{aligned}
|\varphi_1\rangle &= |0\rangle \\
|\varphi_2\rangle &= \frac{1}{\sqrt{3}} |0\rangle + \frac{\sqrt{2}}{\sqrt{3}} |1\rangle \\
|\varphi_3\rangle &= \frac{1}{\sqrt{3}} |0\rangle + \frac{\sqrt{2}}{\sqrt{3}} e^{i\frac{2\pi}{3}} |1\rangle \\
|\varphi_4\rangle &= \frac{1}{\sqrt{3}} |0\rangle + \frac{\sqrt{2}}{\sqrt{3}} e^{i\frac{4\pi}{3}} |1\rangle .
\end{aligned}$$

[11] Maybe all we care about is the measurement, classically.

[12] This tells us that the outcome $p(x)$ is represented by the family $\{\Pi_x\}_x$.

[13] This is the actual starting place for when you have a POVM.

The $\{\frac{1}{2}\,|\varphi_x\rangle\langle\varphi_x|\}_{x=1}^{4}$ form a qubit POVM.

We now consider an application in *quantum state tomography*. Suppose you have a quantum source that is generating some unknown state ρ. A POVM $\{\Pi_x\}$ on $\mathcal{H}^Q$ is called *informationally complete* if the Π_x form a linearly independent set that spans $\mathbb{B}(\mathcal{H}^Q)$.

Remark 4.3.3 Projective measurement are not informationally complete! For instance,

$$\{|1\rangle\langle 1|\,,|2\rangle\langle 2|\,,\ldots,|d\rangle\langle d|\}$$

does not span $\mathbb{B}(\mathbb{C}^d)$!.

For informationally complete POVMs, we can uniquely write

$$\rho = \sum_x c_x \Pi_x.$$

Now, take the sequence[14]

$$\rho = \sum_x c_x \Pi_x \xrightarrow[\text{POVM } \{\Pi_y\}_{y=1}^{d^2}]{} y.$$

We can solve for the coefficients c_x by measuring n copies of ρ and using the sampled average of measurement outcomes. Thus,

$$\frac{\#\text{ outcomes } y}{n} \approx p(y) = \operatorname{tr}\big(\Pi_y \rho\big) = \sum_x c_x \operatorname{tr}\big(\Pi_y \Pi_x\big).$$

Define the matrix M with elements $[[M]]_{yx} = \operatorname{tr}\big(\Pi_y \Pi_x\big)$, so that

$$p(y) = \sum_x [[M]]_{yx} c_x.$$

We get the linear equation

$$\begin{pmatrix} p(1) \\ p(2) \\ \vdots \\ p(d^2) \end{pmatrix} = M \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{d^2}, \end{pmatrix}$$

which we will solve by taking M^{-1}.

[14] We go to d^2 since $\dim \mathbb{B}(\mathbb{C}^d) = d^2$.

On Quantum Communication and Nonlocality

Communication 5

5.1 Multi-System Measurements

Recalling the measurement axiom, what if the system Q consists of multiple subsystems

$$Q = Q_1 Q_2 \cdots Q_N?$$

Then, the projectors P_x might be entangled across the Q_i subsystems. These are called *entangled measurements*:

$$\bigotimes_{i=1}^{N} \mathcal{H}^{Q_i} \simeq \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_N .$$

Note that the $\mathcal{V}_i$ are *spaces*, not subsystems. That is, we are not just taking each qubit. The most important example of an entangled measurement is a two-qubit projective measurement into the Bell basis.[1] We take $\mathcal{H}^Q \simeq \mathbb{C}^2 \otimes \mathbb{C}^2$, where the projectors are

$$\begin{aligned} P_{00} &= |\Phi^+\rangle\langle\Phi^+| \\ P_{01} &= |\Psi^+\rangle\langle\Psi^+| \\ P_{10} &= |\Phi^-\rangle\langle\Phi^-| \\ P_{11} &= |\Psi^-\rangle\langle\Psi^-| . \end{aligned}$$

Example 5.1.1 Compute the post-measurement states and probabilities when performing a Bell measurement on a generic two-qubit state

$$|\psi\rangle = a\,|00\rangle + b\,|01\rangle + c\,|10\rangle + d\,|11\rangle .$$

Solution. We have the projector $P_{ij} = |\Phi^+_{ij}\rangle\langle\Phi^+_{ij}|$, so

$$p(ij) = \operatorname{tr}\left(|\psi\rangle\langle\psi|\; |\Phi^+_{ij}\rangle\langle\Phi^+_{ij}|\right) = \left|\langle\psi|\Phi^+_{ij}\rangle\right|^2 .$$

□

There are various ways to implement a Bell measurement. One way is by inverting the circuit that builds the Bell states. Build the Bell states from the computational basis.[2] Now, invert the circuit, taking an arbitrary state ρ^{AB} through $U^\dagger$, then measuring in the computational basis. What is the probability of measuring (i, j) in the computational basis? Well,

$$p(i,j) = \operatorname{tr}\left(U^\dagger \rho^{AB} (\,|i\rangle\langle i| \otimes |j\rangle\langle j|)\right) = \operatorname{tr}\left(\rho^{AB}\, |\Phi^+_{ij}\rangle\langle\Phi^+_{ij}|\right),$$

which is equivalent to measuring ρ^{AB} in the Bell basis.

[1] We call this a *Bell measurement*.

[2] The unitary U we need is just the Hadamard H followed by CNOT. Recall that

$$H\,|i\rangle = \frac{|0\rangle + (-1)^i\,|1\rangle}{\sqrt{2}} = \sigma_z^i\,|+\rangle .$$

Remark 5.1.1 This inversion of circuit trick can be used to realize a projective measurement in any orthonormal basis $\{ |\psi_{\mathbf{b}}\rangle\}_{\mathbf{b}\in C_2^n}$.

Local measurements are independent POVMs performed on two or more subsystems. In general, we have

$$p(i,j) = \mathrm{tr}\left(\rho^{AB}\left(\Pi_i^A \otimes \Sigma_j^B\right)\right).$$

5.2 LOCC and Entanglement

Local operations (LO) are independent CPTP maps performed on two or more subsystems. Then,

$$\sigma^{AB} = \mathscr{E}^A \otimes \mathscr{F}^B(\rho^A B).$$

We refer to quantum operations in which the parties perform an interactive sequence of local instruments while exchanging classical data *local operations and classical communication* (LOCC). We send

$$\rho^{AB} \mapsto \sigma^{AB} := \sum_{i_{\mathrm{tot}}} \mathscr{A}_{i_{\mathrm{tot}}} \otimes \mathscr{B}_{i_{\mathrm{tot}}}(\rho^{AB}).$$

In general, LOCC operations can have a very complex structure.

Example 5.2.1 Suppose that Alice and Bob are given one of the following four bipartite states $|\psi_1\rangle = |00\rangle$, $|\psi_2\rangle = |01\rangle$, $|\psi_3\rangle = |1+\rangle$, $|\psi_4\rangle = |1-\rangle$. Show that they cannot be perfectly distinguished by LO, but they can be perfectly distinguished by LOCC.

Proof. Bob must distinguish between his parts

$$\{ |0\rangle, |1\rangle, |+\rangle, |-\rangle\}.$$

Yet, non-orthogonal states means there exists some error in Bob's measurement.[3] Now, for the LOCC, let Alice measure in the $\{ |0\rangle, |1\rangle\}$ basis. On Bob's part, if Alice gets 0, then Bob has orthogonal states, so let him measure in the computational basis. Similarly, if Alice gets 1, Bob can measure in $\{ |\pm\rangle\}$. Then, Bob announces what his outcome was. □

Definition 5.2.1 (Separable) *Remember, a bipartite state ρ^{AB} is called separable if it is a convex combination of product states*

$$\rho^{AB} = \sum_\lambda q_\lambda \omega_\lambda^A \otimes \tau_\lambda^B.$$

Suppose we perform some LOCC map on a separable state. Then, we get out

$$\sum_\lambda q_\lambda \sum_{i_{\mathrm{tot}}} \mathscr{A}_{i_{\mathrm{tot}}}(\omega_\lambda^A) \otimes \mathscr{B}_{i_{\mathrm{tot}}}(\tau_\lambda^B),$$

[3] See the previous discussion of measurement and POVMs.

which is still a convex combination of product states. On the other hand, if we consider an arbitrary separable state ρ^{AB}, we can build it using LOCC via

(i) Alice samples from random variable λ with probability q_λ.
(ii) She prepares the state ω_λ^A and communicates λ to Bob.
(iii) Bob prepares the state τ_λ^B.

Theorem 5.2.1 *A state* ρ^{AB} *is separable if and only if it can be generated by LOCC.*

Remark 5.2.1 When asked the question "What is entanglement?," we answer "It is a property in all quantum states that cannot be generated by LOCC."[4]

Proposition 5.2.2 *It is impossible to perform a Bell measurement by LOCC.*

Proof. We want to send ρ^{AB} through an LOCC with output (i, j) such that

$$p(i,j) = \mathrm{tr}\Big(\rho^{AB} \, |\Phi_{ij}^+\rangle\!\langle\Phi_{ij}^+|\Big).$$

Suppose Alice and Bob initially share the separable state $|\Phi^+\rangle^{AA'} \otimes |\Phi^+\rangle^{B'B}$. We write out

$$\begin{aligned} |\Phi^+\rangle^{AA'} \otimes |\Phi^+\rangle^{BB'} &= (\,|00\rangle + |11\rangle)^{AA'} \otimes (\,|00\rangle + |11\rangle)^{AB} \\ &= |00\rangle^{A'B'} \, |00\rangle^{AB} + |01\rangle^{A'B'} \, |01\rangle^{AB} \\ &\quad + |10\rangle^{A'B'} \, |10\rangle^{AB} + |11\rangle^{A'B'} \, |11\rangle^{AB} \,. \end{aligned}$$

Let us look at the Φ^+ projection onto the $A'B'$, we see

$$\begin{aligned} &\big(\mathbb{1}^{AB} \otimes |\Phi^+\rangle\!\langle\Phi^+|^{A'B'}\big)\, |\Phi^+\rangle^{AA'} \, |\Phi^+\rangle^{B'B} \\ &= |\Phi^+\rangle^{A'B'} \otimes |\Phi^+\rangle^{AB} \,. \end{aligned}$$

Thus, projecting $A'B'$ onto $|\Phi^+\rangle$ implies AB is in state $|\Phi^+\rangle^{AB}$, so we have generated entanglement on AB via LOCC, a contradiction.[5] □

5.3 Classical and Quantum Communication Resources

We will start with a unifying perspective on these types of resources.

Definition 5.3.1 (Classical Information) *We have defined classical information to be mixtures of bit strings. That is, states of classical registers.*

Definition 5.3.2 (Quantum Information) *Similarly, quantum information is mixtures of qubit strings. That is, states of quantum systems.*

[4] That is, we place LOCC as a more fundamental concept. Then, non-separability follows from the structure of LOCC. This is the "physical process" perspective on entanglement.

[5] Check the details of the last computation, it takes some careful observation.

In a high level sense, there are *classical resources* and *quantum resources*. Within these sections, we can divide into *static resources* and *dynamic resources*. Classical static resources [cc] have one bit of shared randomness (SR):

$$\gamma^{\mathsf{XY}} = \frac{1}{2}\,|00\rangle\!\langle 00|^{\mathsf{XY}} + \frac{1}{2}\,|11\rangle\!\langle 11|^{\mathsf{XY}}.$$

Shared randomness requires *some sort* of interaction.[6] Quantum static resource [qq] have on entangled bit (ebit)

$$|\Phi^{+}\rangle^{AB} = \frac{1}{\sqrt{2}}(\,|00\rangle + |11\rangle).$$

Classical dynamic resources [c → c] are one bit noiseless classical channels

$$\mathrm{id}^{\mathsf{X}\to\mathsf{Y}} : |i\rangle^{\mathsf{X}} \mapsto |i\rangle\!\langle i|^{\mathsf{Y}}.$$

These are precisely classical channels, in the sense we are used to. We are assuming the ideal case.[7] Finally, quantum dynamic resources [q → q] are one bit noiseless quantum channels

$$\mathrm{id}^{A\to B} : |\psi\rangle^{A} \mapsto |\psi\rangle^{B}.$$

Above, $i \in \{0, 1\}$ and $|\psi\rangle \in \mathbb{C}^2$. Our goal is to understand the relationship between these resources. At an abstract level, *all of information theory* can be thought of as transformations and trade-offs between resources like these four.

One instance, is one bit shared randomness (SR) from a one bit classical channel [c → c] ≥ [cc]. Alice locally prepares the state

$$\gamma^{\mathsf{X}\tilde{\mathsf{X}}} = \frac{1}{2}\,|00\rangle\!\langle 00|^{\mathsf{X}\tilde{\mathsf{X}}} + \frac{1}{2}\,|11\rangle\!\langle 11|^{\mathsf{X}\tilde{\mathsf{X}}}.$$

She uses the noiseless channel $\mathrm{id}^{\tilde{\mathsf{X}}\to\mathsf{Y}}$ to send system $\tilde{\mathsf{X}}$ to Bob. Then, we get

$$\gamma^{\mathsf{XY}} = \frac{1}{2}\,|00\rangle\!\langle 00|^{\mathsf{XY}} + \frac{1}{2}\,|11\rangle\!\langle 11|^{\mathsf{XY}}.$$

In fact, [c → c] > [cc], meaning anything we can do with one bit shared randomness can be done with a one bit classical channel, but the converse is untrue.

For the quantum analog, we could get one ebit from one qubit quantum channel. Alice starts with $|\Phi_2^{+}\rangle$, sends one qubit over the quantum channel, yielding [q → q] > [qq].

The fun stuff happens when we take one bit SR from one ebit [qq] ≥ [cc]. Alice and Bob initially share $|\Phi_2^{+}\rangle$, they measure in the computational basis $\{0, 1\}$. The joint state of the classical registers XY is γ^{XY}.[8]

Remark 5.3.1 (Classical information over quantum channels) Suppose that $\mathrm{id}^{A\to B}$ is a d-dimensional quantum channel. Clearly, it is possible to send $|\mathcal{X}| = d$ classical messages over this channel. Why? Well we can just use the computational basis states $|1\rangle, \ldots, |d\rangle$.[9] We have the c → q encoding: Alice copies the bit $x \in \{1, \ldots, d\}$ into the computational

[6] The notation [cc] means both X, Y means both are classical systems, and the brackets mean we are defining our resource.

[7] All channels are subject to error. Note that the [c → c] implies the dynamic nature, with an arrow time.

[8] The randomness is from the perspective of someone who does not have access to the measurement outcomes.

[9] That is, there are $\log d$ bits.

basis of the quantum system

$$|x\rangle\langle x|^{\mathsf{X}} \mapsto \rho_x^{\mathsf{A}} = |x\rangle\langle x|^{\mathsf{A}}.$$

Then, for q → c decoding: Bob measures his system in the computational basis $\{\Pi_y = |y\rangle\langle y|\}_1^d$. The probability of decoding is

$$p(y \mid x) = \mathrm{tr}\,(\Pi_y \rho_x).$$

We ended up with $\log_2 d$ [q → q] ≥ $\log_2 d$ [c → c]. The LHS here means we have a d-dimensional, noiseless quantum channel.

Theorem 5.3.1 *We cannot send more than d classical messages over a d-dimensional quantum channel.*

Proof. Suppose Alice tries to send $s > d$ messages by encoding $|x\rangle\langle x| \mapsto \rho_x$ for $x \in \{1, \dots, s\}$. Suppose Bob performs a decoder POVM $\{\Pi_y\}_{y=1}^s$. Then, the average probability of successfully decoding is

$$\frac{1}{s}\sum_{x=1}^{s} p(x \mid x) = \frac{1}{s}\sum_{x=1}^{s} \mathrm{tr}\,(\Pi_x \rho_x) \leq \frac{1}{s}\sum_{x=1}^{s} \mathrm{tr}(\Pi_x) = \frac{1}{s}\,\mathrm{tr}\big(\mathbb{1}^B\big) = \frac{d}{s} < 1.$$

Bummer. □

Remark 5.3.2 (Entanglement-assisted classical communication over a quantum channel) We start with the same setup, except we also have a static resource with the quantum channel. We end up with

$$\rho_x^{\mathsf{AB'}} = \mathscr{E}_x^{\mathsf{A}\to\mathsf{A}} \otimes \mathrm{id}(\varphi) \to \rho_x.$$

If B and B' are d-dimensional then $\{\rho_x^{BB'}\}_{x=1}^s$ is a collection of states on $\mathbb{C}^d \otimes \mathbb{C}^d$. Then, for any decoder POVM on Bob's side, the average probability is upper bounded by d^2/s![10]

5.4 Dense Coding

Let us begin by looking at two-qubit dense coding:

(i) Let $\mathcal{X} := \mathbb{Z}_2^2 = \{00, 01, 10, 11\}$.
(ii) Alice encodes her message $b_0 b_1 \in \mathcal{X}$ by applying $\sigma_z^{b_0}\sigma_x^{b_1}$ to A. This transforms $|\Phi_{00}^+\rangle \mapsto |\Phi_{b_0 b_1}^+\rangle$.
(iii) Alice sends qubit A over the channel to Bob. He performs a Bell measurement on systems $\mathsf{BB'}$.[11]

In our resource-theoretic terms, we have [qq] + [q → q] ≥ 2 [c → c]. The key point is that we are using *entanglement* to increase how much classical data we can send over a quantum channel.

[10] In fact, we can saturate this upper bound of sending $s = d^2$ perfect messages using *dense coding*.

[11] Why does this work? Alice can transform $|\Phi_{00}^+\rangle$ into 4 different orthogonal states by just acting on her half of $|\Phi_{00}^+\rangle$.

Let us now extend to higher dimensions for $\mathbb{C}^d \otimes \mathbb{C}^d$. We need to introduce a d-dimensional Bell basis for $(\mathbb{C}^d)^{\otimes 2}$, so define

$$X(l) = \sum_{j=0}^{d-1} |l + j\rangle\langle j| \quad (\text{mod } d) \quad \text{for} \quad l = 0, 1, \ldots, d-1$$

and

$$Z(s) = \sum_{j=0}^{d-1} e^{2\pi i \frac{sj}{d}} |j\rangle\langle j| \quad \text{for} \quad s = 0, 1, \ldots, d-1.$$

The d-dimensional *Bell basis* is the collection of d^2 states given by

$$\left\{ |\Phi^+_{ls}\rangle := Z(l)X(s) \otimes \mathbb{1} |\Phi^+_d\rangle : l, s = 0, \ldots, d-1 \right\}.$$

It is not clear that these are orthonormal, but it can be shown with simple computation.[12] Now, we can do $d \otimes d$-dimensional coding:

(i) Let $\mathcal{X} := \mathbb{Z}_d^2$.
(ii) Alice encodes her message $(l, s) \in \mathcal{X}$ by applying $Z(l)X(s)$ to A. This transforms $|\Phi^+_{00}\rangle \mapsto |\Phi^+_{ls}\rangle$.
(iii) Then, Bob performs a projective Bell measurement.

The resource inequality for general dense coding is $\log d$ [qq] + $\log d$ [q → q] $\geq 2 \log d$ [c → c].[13]

Remark 5.4.1 In case it was unclear, if we d bits, then we have 2^d possible outcomes (since they take values in $\mathbb{Z}_2$). Then, a d-dimensional channel would have associated $\log_2 d$ bits.

5.5 Teleportation

What if we wanted to send quantum information over *a classical channel*? Suppose Alice has a Hilbert space $\mathcal{H}^A$, and she wants to completely send it to Bob. Alice measures her quantum state ρ^A using some POVM $\{\Pi_x\}_x$: q → c encoding Then Bob can prepare some quantum state ρ_x given that Alice has outcome x from her measurement: c → q decoding.

Well, for a given input ρ, Bob's state averaging over outcomes x:[14]

$$\sigma_\rho := \sum_x p(x)\rho_x = \sum_x \text{tr}(\Pi_x)\rho_x.$$

Our goal: For *every* state of Alice's system ρ^A, we want $\sigma_\rho^B = \rho^A$. That would be perfect communication! Note that[15]

$$\min_{|\varphi\rangle \in \mathcal{H}} \langle\varphi|\sigma_\varphi|\varphi\rangle = \min_{|\varphi\rangle \in \mathcal{H}} \sum_x \langle\varphi|\Pi_x|\varphi\rangle \ \langle\varphi|\rho_x|\varphi\rangle \leq \frac{1}{d}.$$

our "worst-case" scenario. Unsurprisingly, this means the fidelity drops off via $1/d$, so classical communication is insufficient. Recall dense coding: What if, instead, we have some entanglement assistance for our classical channel–as we did with dense coding?

[12] Use the standard sum and product rules for roots of unity.

[13] In a sense, this means entanglement allows us to double our classical communication. Practically, we have to go beyond dense coding to polynomial or even *exponential* enhancement.

[14] That is, tracing over the classical register Y.

[15] We omit the proof for time.

$$\begin{tikzcd}
\mathbb{Z}_d^2 \arrow[r, dash, "|sl\rangle\langle sl|"] & Z(s)X(l) \arrow[ddl, dash, "A"'] \arrow[dr, Rightarrow, "A"] & & & \\
 & & \mathrm{id}^{A\to B} \arrow[dr, Rightarrow, "B"] & & \\
|\Phi_{00}\rangle^{AB'} \arrow[rrr, dash, "B'"] & & & \text{Bell Measure} \arrow[r, dash] & |s'l'\rangle\langle s'l'|
\end{tikzcd}$$

Figure 5.1: Communication diagram for $d \otimes d$-dimensional dense coding, where $Z(s), X(l)$ are the higher dimensional Sylvester operators

$$\begin{tikzcd}
\mathscr{D}(\mathbb{C}^d) \arrow[r, dash, "\rho^A"] & \{\Pi_m^{AA'}\} \arrow[ddl, dash] \arrow[dr, Rightarrow, "|x\rangle\langle x|^X"] & & & \\
 & & \mathrm{id}^{X\to Y} \arrow[dr, Rightarrow, "|x\rangle\langle x|^Y"] & & \\
\Psi^{A'B'} \arrow[rrr, dash] & & & \mathscr{D}^{B'\to B} \arrow[r, dash] & \rho^B
\end{tikzcd}$$

Figure 5.2: After the message encoding (the POVM), we have $|m\rangle\langle m|^X \otimes \mathrm{tr}_{AA'}((\Pi_m^{AA'} \otimes \mathbb{1}^{B'})(\rho^A \otimes \Psi^{A'B'}))$. Then, after the classical channel ($\mathrm{id}^{X\to Y}$), we have $|m\rangle\langle m|^Y \otimes \mathrm{tr}_{AA'}((\Pi_m^{AA'} \otimes \mathbb{1}^{B'})(\rho^A \otimes \Psi^{A'B'}))$.

We can employ a technique called *teleportation* to achieve this sort of mechanism:

$$\begin{tikzcd}
\mathscr{D}(\mathbb{C}^d) \arrow[r, dash, "\rho^A"] & \{\,|\Phi_{sl}\rangle\langle\Phi_{sl}|\}_{sl} \arrow[ddl, dash] \arrow[dr, Rightarrow, "|x\rangle\langle x|^X"] & & & \\
 & & \mathrm{id}^{X\to Y} \arrow[dr, Rightarrow, "|x\rangle\langle x|^Y"] & & \\
|\Phi_{00}\rangle^{A'B'} \arrow[rrr, dash] & & & Z(s)^t X(l)^t \arrow[r, dash] & \rho^B
\end{tikzcd}$$

Figure 5.3: Communication diagram for teleportation: Alice teleports her quantum state ρ^A to Bob.

Mathematically, the way entanglements helps us in teleportation is precisely *the ricochet property*. We omit the proof, but we are really using the special property of $|\Phi^+\rangle$ to do this.

Then, the inequality is $(2\log d)$ [c → c] $+(\log d)$ [qq] $\geq (\log d)$ [q → q].

Example 5.5.1 What is the probability of each outcome (s,l) on Alice's side, using standard teleportation? Well,[16]

$$\begin{aligned}
\mathbb{P}(s,l) &= \mathrm{tr}(\overbrace{(\mathbb{1}^B \otimes \Phi_{sl}^{A'A})}^{\text{measurement}}\ \overbrace{(\Phi_{00}^{A'B} \otimes \rho^A)}^{\text{initial state}}) \\
&= \frac{1}{d}\,\mathrm{tr}\left(\Phi_{sl}^{A'A}(\mathbb{1}^{A'} \otimes \rho^A)\right) \\
&= \frac{1}{d^2}\,\mathrm{tr}\left(\mathbb{1}^A \rho^A\right) = \frac{1}{d^2}.
\end{aligned}$$

[16] Every density matrix is unit-trace, so we cannot bias our outcome. The probability is uniform over (s,l).

Local Correlations and CHSH 6

Suppose we have two parties, Alice and Bob, who each have access to a pair of "black boxes." Alice chooses x among M_A inputs. Her device outputs a among N_A outputs. Bob does the same with y and M_B, N_B. The entire input/output behavior can be described by a family of conditional probabilities $p(a, b|x, y)$–*correlations*–held between Alice and Bob.

6.1 Bipartite Correlations

Suppose we are in the situation given above.

Definition 6.1.1 (Bipartite Device) *A bipartite input/output device is then represented by a point*

$$\mathbb{R}^{M_A M_B N_A N_B}$$

with coordinates

$$\{p(a, b|x, y)\}_{a,b,x,y}.$$

The collection of all possible input/output devices forms a subset

$$\mathcal{P}_{M_A,M_B,N_A,N_B} \subseteq \mathbb{R}^{M_A M_B N_A N_B}.$$

It is natural to ask what the structure of $\mathcal{P}_{M_A,M_B,N_A,N_B}$ is. Let $M_A = M_B = N_A = N_B = 2$. Then,

(i) every $v \in \mathcal{P}_{2222}$ is described by 16 coordinates:

$$\{p(a, b|x, y)\}_{a,b,x,y=0}^{1}.$$

(ii) there are special elements in $\mathcal{P}_{2222}$ that behave *deterministically*, meaning that $p(a, b|x, y) \in \mathbb{2}$ for all a, b, x, y.

How many deterministic boxes are there? In other words, how many functions $f : \mathbb{2}^2 \to \mathbb{2}^2$? The answer is $4^4 = 256$.[1] Label the deterministic boxes as $v_i(a, b|x, y)$.

Theorem 6.1.1 *Every element of $\mathcal{P}_{2222}$ can be written as a convex combination of deterministic boxes:*

$$p(a, b|x, y) = \sum_{\lambda=1}^{256} p(\lambda) v_\lambda(a, b|x, y).$$

Then, $\mathcal{P}_{2222}$ forms a polytope, where the deterministic boxes are the extremes.

[1] This is just counting.

6.2 Local Correlations

We now consider a special family of correlations motivated by *locality*. The intuitive meaning of locality is that Alice and Bob's devices are influenced only by what is happening in the nearby surrounding physical space. Consequently, if the boxes are separated by a *large distance*, then they should be have independently of each other. Thus,[2]

$$p(a,b|x,y) = p(a|x)p(b|y).$$

However, it may have happened that at some time in the past, the boxes, or parts of their boxes, were near each other and interacted. Specifically, the boxes may have shared some classical information with each other, which we will represent by the character λ.[3] We can think of λ as a classical register:

$$\rho_\lambda^{AB} = \sum_\lambda p(\lambda)\,|\lambda\rangle\langle\lambda|^X \otimes |\lambda\rangle\langle\lambda|^Y .$$

The behavior of the boxes can then depend on the particular value of λ:

$$p(a,b|x,y,\lambda) = p(a|x,\lambda)p(b|y,\lambda).$$

Suppose that λ has a distribution independent of x, y. Then,

$$p(a,b|x,y) - \sum_\lambda p(\lambda)p(a|x,\lambda)p(b|y,\lambda).$$

Definition 6.2.1 (Local Correlation) *A correlation, as above, is called a local correlation. They form a polytope* $\mathscr{L}_{M_A,M_B,N_A,N_B} \subseteq \mathscr{P}_{2222}$.[4]

In the discussion so far, we have not mentioned quantum systems. In fact, the only assumptions we have made is that the inputs and outputs are *classical*. We assumed that there could be some shared randomness, also classical. Our plan is to replace the instructions with a *quantum state* ρ^{AB}.

The outputs (a,b) are obtained by measuring ρ^{AB}. The inputs (x,y) determine which measurement Alice and Bob perform on their subsystem of ρ^{AB}. Specifically, for Alice, $\{\Pi^A_{a|x}\}_{a,x}$ is a family of POVMs for Alice.[5] Assume the same for Bob. Then, using Born's rule,

$$p(a,b|x,y) = \operatorname{tr}\left(\Pi^A_{a|x} \otimes \Pi^B_{b|y}\right)\rho^{AB}.$$

Definition 6.2.2 (Quantum Correlation) *A correlation, as above, is called a quantum correlation. The set of all quantum correlations is denoted* $\mathbb{Q}_{M_A,M_B,N_A,N_B}$.

Proposition 6.2.1 $\mathscr{L}_{M_A,M_B,N_A,N_B} \subseteq \mathbb{Q}_{M_A,M_B,N_A,N_B}$.

Proof.

$$\sum_\lambda p(\lambda)p(a|x,\lambda)p(b|y,\lambda) = \operatorname{tr}\left(\Pi^A_{a|x} \otimes \Pi^B_{b|y}\right)\rho^{AB}.$$

[2] We will sometimes call this the *Bell locality condition*.

[3] Think of this classical information like a shared "set of instructions" λ.

[4] Local correlations are said to satisfy a *local hidden variable* (LHV) model. This sort of idea dates even back to Einstein.

[5] Recall that POVMs satisfy the standard completion relation.

□

However, it turns out that $\mathscr{L}_{M_A,M_B,N_A,N_B} \subsetneq \mathbb{Q}_{M_A,M_B,N_A,N_B}$. How do we prove that there exist nonlocal quantum correlations?[6]

6.3 Polytope Membership Problem

In fact, the question above is a part of a more general type of problem, called the polytope membership problem. Start with a simpler problem. In $\mathbb{R}^2$, consider the polytope T whose vertices are $v_0 = (0, 0)$, $v_1 = (0, 1)$, and $v_2 = (1, 0)$. An element $x \in \mathbb{R}^2$ belongs to T if and only if it can be written as a convex combination of the v_i. We call this interpretation the $\mathscr{V}$-representation of T.

However, there is an alternative description of T that is often easier to work with. We can equivalently think of T as the intersection of three half planes:

$$\begin{aligned} \mathscr{H}_1 &:= \{(x, y) : x \geq 0\} \\ \mathscr{H}_2 &:= \{(x, y) : y \geq 0\} \\ \mathscr{H}_3 &:= \{(x, y) : x + y \leq 1\}, \end{aligned}$$

taking

$$T = \mathscr{H}_1 \cap \mathscr{H}_2 \cap \mathscr{H}_2.$$

This is known as the $\mathscr{H}$-representation of T. Thus, to decide whether or not $x \in \mathbb{R}^2$ is in T, we just need to check a system of three inequalities.

Definition 6.3.1 (Facet Inequalities) *The inequalities defining the half planes $\mathscr{H}_i$ the facet inequalities of T.*

Let us return to $\mathbb{R}^{16}$. We can form the associated half planes for $\mathscr{L}_{2222}$. The facet inequalities for the local polytope are called *Bell inequalities*.

Remark 6.3.1 In order to prove that a quantum correlation is nonlocal, we must show that it violates a Bell inequality. The number of Bell inequalities to characterize the polytope $\mathscr{L}$ depends on (M_A, M_B) and (N_A, N_B). Unfortunately, this number scales exponentially in M_i, N_i.[7]

Theorem 6.3.1 (Bell, CHSH, Fine) *Let $M_i, N_i = 2$. The polytope $\mathscr{L}_{2222}$ is characterized entirely by two types of inequalities:*

(i) positivity*: $p(a, b|x, y) \geq 0$ for all $a, b, x, y \in \mathbb{2}$.*

(ii) CHSH*:*

$$\alpha \leq p(a, b|x, y) - p(a, \overline{b}|x, \overline{y}) - p(\overline{a}, \overline{b}|\overline{x}, y) - p(a, b|\overline{x}, \overline{y}) \leq \beta$$

for all $a, b, x, y \in \mathbb{2}$, defining $\overline{j} := j \oplus 1$, where $\oplus$ is modular addition in $\mathbb{2}$; i.e., in $(\mathbb{Z}/2, +)$.

[6] It is worth noting that $\mathbb{Q}$ is *not a polytope*. The topological reasons why are a bit subtle.

[7] In fact, for general membership, the problem is NP-hard.

What are the values of α and β? Well, it is clear that $\alpha \geq -3$ and $\beta \leq 1$. We want to tighten this.[8] Let us first find the largest value of α such that this inequality is satisfied by all correlations in $\mathscr{L}_{2222}$. Assume we have a local hidden variable model. Suppose

$$p(a,b|x,y) = \sum_{\lambda} p(\lambda) p_A(a|x,\lambda) p_B(b|y,\lambda).$$

Substitute this into the LHS of the CHSH inequality. After a good amount of algebra, we get

$$2\alpha_0\beta_0 - \alpha_0\beta_0 - (\alpha_0 - \alpha_1)(\beta_0 - \beta_1),$$

bounding

$$0 \leq \alpha_i, \beta_i \leq 1,$$

where $\alpha_i := p_A(0|i)$, and likewise for β_i. Then,

$$2\alpha_0\beta_0 - \alpha_0\beta_0 - (\alpha_0 - \alpha_1)(\beta_0 - \beta_1) \geq -1.$$

Now, let us find the upper bound on β. Following the same procedure, we get that

$$2\alpha_0\beta_0 - \alpha_0\beta_0 - (\alpha_0 - \alpha_1)(\beta_0 - \beta_1) \leq 0.$$

Remark 6.3.2 It is a bit of an exercise to do these computations, taking maximums via some partials. It is just a bounded optimization problem.

Thus, our CHSH becomes

$$-1 \leq p(a,b|x,y) - p(a,\overline{b}|x,\overline{y}) - p(\overline{a},\overline{b}|\overline{x},y) - p(a,b|\overline{x},\overline{y}) \leq 0$$

in $\mathbb{R}^{16}$, where $x, y \in \mathbb{2}$.

6.4 Violation of CHSH

Suppose Alice and Bob each perform a two-outcome projective measurement on a qubit system. A two-outcome projective measurement on a qubit system means they project onto some orthonormal basis. ALice and Bob are each choosing between two directions on the Bloch sphere, and their outcome is either spin-up or spin-down.

Our projection for Alice takes the form

$$P_{a|\hat{m}_x} := \frac{1}{2}(\mathbb{1} + (-1)^a \hat{m}_x \cdot \vec{\sigma}),$$

and likewise for Bob.

[8] There is an algorithm to go between $\mathscr{V}$- and $\mathscr{H}$-representations, but it is complicated.

Suppose they measure on the Bell state $|\Phi^+\rangle$:

$$\begin{aligned} p(a,b|x,y) &= \left\langle \Phi^+ \middle| P_{a|\hat{m}_x} \otimes P_{b|\hat{n}_y} \middle| \Phi^+ \right\rangle \\ &= \frac{1}{4} \left\langle \Phi^+ \middle| (\mathbb{1} + (-1)^a \hat{m}_x \cdot \vec{\sigma})(\mathbb{1} + (-1)^b \hat{n}_y \cdot \vec{\sigma})^t \otimes \mathbb{1} \middle| \Phi^+ \right\rangle \\ &= \frac{1}{8} \operatorname{tr}\left((\mathbb{1} + (-1)^a \hat{m}_x \cdot \vec{\sigma})(\mathbb{1} + (-1)^b \hat{n}_y \cdot \vec{\sigma})^t \right) \\ &= \frac{1}{4}\Big(1 + (-1)^{a \oplus b} \hat{m}_x \cdot \hat{n}'_y \Big). \end{aligned}$$

Note that we took $\hat{n}'_y := (n_x, -n_y, n_z)$. Substitute this into the CHSH inequality. Our goal is to get a violation. How could we maximize

$$\frac{1}{4}(\hat{m}_0 \cdot (\hat{n}'_0 - \hat{n}'_1) - \hat{m}_1 \cdot (\hat{n}'_0 - \hat{n}'_1)) - \frac{1}{2}.$$

We need to find two vectors whose sum and difference have a large norm. Thus, we pick $\hat{n}'_0 := \hat{z}$, $\hat{m}_0 := \frac{1}{\sqrt{2}}(\hat{z} - \hat{x})$, $\hat{n}'_1 := \hat{x}$, and $\hat{m}_1 := -\frac{1}{\sqrt{2}}(\hat{z} + \hat{x})$.

Thus,

$$\frac{1}{4}(\sqrt{2} + \sqrt{2}) - \frac{1}{2} = \frac{1}{2}(\sqrt{2} - 1) > 0,$$

and we have caused nonlocality, finding something in $\mathbb{Q}_{2222}$ which is not in the $\mathscr{L}_{2222}$ polytope.

Remark 6.4.1 It is reasonable to ask if we could violate this more. Perhaps surprisingly, the answer is actually no. The proof sketched above shows the fundamental limits given by quantum mechanics, which cannot be surpassed. That is, the quantum bound $\beta_Q \leq 1$, which is greater than the $\beta = 0$ bound.[9]

Quantum mechanics, and notably entanglement, has given us correlations that we *fundamentally* cannot produce classically.

[9] That is, there are bounds on the quantum correlations, which stems from something natural. There are nonlinear inequalities which bound $\mathbb{Q}_{2222}$, but it gets really complicated. In our picture, our bounds also allow some classical, nonlocal correlations. This is irrelevant to our discussion.