Quantum Information Theory (NMMB537)

Peter Zeman

October 24, 2025

Abstract

These are lecture notes for the course NMMB537 Quantum Information Theory, taught at Faculty of Mathematics and Physics, Charles University. The lecture notes are mostly inspired by the course NMAK14020U Quantum Information Theory, offered at University of Copenhagen.

Contents

Chapter 1

Quantum States and Measurements

We start by describing a precise mathematical framework for quantum states.

1.1 Finite-dimensional Hilbert Spaces

Axiom 1 (Hilbert space)**.** To every quantum system we asocciate a Hilbert space $\mathcal{H}$.

We only consider finite-dimensional Hilbert spaces in this course. If $\dim(\mathcal{H}) = d$, then $\mathcal{H}$ is isomorphic to

$$\mathbb{C}^d = \left\{ \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_d \end{pmatrix} : \psi_i \in \mathbb{C} \right\}$$

with the standard inner product

$$\langle \psi | \phi \rangle = \sum_{i=1}^{d} \overline{\psi_i} \phi_i.$$

Bra-ket notation. We write $|\psi\rangle \in \mathcal{H}$ for any vector. By $\langle \psi | \in \mathcal{H}^\dagger$, we mean the linear map $\mathcal{H} \to \mathbb{C}$ defined by $\langle \psi | (|\phi\rangle) = \langle \psi | \phi \rangle$. In other words, $\langle \psi |$ is the row vector corresponding to the conjugate transpose of $|\psi\rangle$. We denote the standard basis of $\mathbb{C}^d$ by $|0\rangle, |1\rangle, \ldots, |d\rangle$. That is, $|i\rangle$ has 1 on the i^{th} coordinate and 0 elsewhere.

Linear algebra notation. Every basis that we will consider will be orthogonal. If there is no confusion, we will use the terms linear map and matrix interchangeably.

By $\mathrm{Lin}(\mathcal{H}, \mathcal{K})$, we denote the set of linear maps $\mathcal{H} \to \mathcal{K}$. When we fix the standard bases, we can express $M \in \mathrm{Lin}(\mathcal{H}, \mathcal{K})$ as

$$M = \sum_{i,j} M_{ij} |i\rangle\langle j|.$$

Note that $\langle i | M | j \rangle = M_{ij} \in \mathbb{C}$.

Exercise 1.1. Write $\sum_{i,j \in \{0,1\}} ij |i\rangle\langle j|$ as a matrix.

If $M \in \mathrm{Lin}(\mathcal{H}, \mathcal{H}) = \mathrm{Lin}(\mathcal{H})$, then the *trace* with respect to the standard basis is

$$\mathrm{tr}[M] = \sum_i M_{ii} = \sum_i \langle i|M|i\rangle.$$

An important fact about trace is that it is in fact independent of the basis. Recall also that $\mathrm{tr}[MN] = \mathrm{tr}[NM]$.

Here are some important types of matrices:

- $M \in \mathrm{Lin}(\mathcal{H})$ is *Hermitian* if $M^\dagger = M$.
- $P \in \mathrm{Lin}(\mathcal{H})$ is *positive* if $\langle\psi|P|\psi\rangle \geq 0$, for all $|\psi\rangle \in \mathcal{H}$. $\mathrm{PSD}(\mathcal{H})$ is the class of all positive matrices. We also use the notation $P \geq 0$ to indicate that P is positive. If $M, N \in \mathrm{Lin}(\mathcal{H})$, then we define $M \geq N$ if $M - N \geq 0$, that is, if $M - N$ is positive.
- $U \in \mathrm{Lin}(\mathcal{H})$ is *unitary* if $U^\dagger U = UU^\dagger = I$. $\mathcal{U}(\mathcal{H})$ is the class of all unitary matrices.
- $V \in \mathrm{Lin}(\mathcal{H}, \mathcal{K})$ is an *isometry* if $V^\dagger V = I$.
- $P \in \mathrm{Lin}(\mathcal{H})$ is a *projection* if $P^\dagger = P$ and $P^2 = P$.

Exercise 1.2. Are the following matrices $\{\text{Hermitian}, \text{PSD}, \text{unitary}, \text{projections}\}$?

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \begin{pmatrix} e^{i\frac{\pi}{4}} & 0 \\ 0 & e^{-i\frac{\pi}{4}} \end{pmatrix}.$$

Theorem 1.3 (Spectral theorem for Hermitian matrices)**.** *If $M \in \mathrm{Lin}(\mathcal{H})$ is Hermitian $d \times d$ matrix, then M has eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ and there is a basis of eigenvectors $\{|\psi_i\rangle\}_{i=1}^d$ such that*

$$M = \sum_{i=1}^d \lambda_i |\psi_i\rangle\langle\psi_i|.$$

Recall that the eigenvalues can have multiple occurences and the basis of eigenvectors may not be unique. Indeed, for instance,

$$I = \sum_{i=1}^d |\psi_i\rangle\langle\psi_i|$$

is true for any basis $\{|\psi_i\rangle\}_{i=1}^d$.

This allows us to define functions of Hermitian matrices. If $f\colon \mathbb{R} \to \mathbb{R}$ is a function, then we can define

$$f(M) = \sum_{i=1}^d f(\lambda_i)|\psi_i\rangle\langle\psi_i|.$$

For instance,

$$\sqrt{M} = \sum_{i=1}^d \sqrt{\lambda_i}|\psi_i\rangle\langle\psi_i|.$$

Note that if P is a projection, then the condition $P^2 = P$ implies that the eigenvalues are in $\{0, 1\}$. Thus, we can write $P = \sum_{i=1}^r |\psi_i\rangle\langle\psi_i|$. Here, $r = \mathrm{rank}(P)$ of the projection P.

Theorem 1.4 (Characterization of positive matrices)**.** *If $P \in \mathrm{Lin}(\mathcal{H})$, then the following are equivalent:*

(a) P is positive, i.e., $\langle\psi|P|\psi\rangle \geq 0$, for all $|\psi\rangle \in \mathcal{H}$.

(b) $P^\dagger = P$ and all eigenvalues are nonnegative.

(c) There is $M \in \mathrm{Lin}(\mathcal{H}, \mathcal{K})$ such that $P = M^\dagger M$.

(d) $\mathrm{tr}[PQ] \geq 0$, for every $Q \in \mathrm{PSD}(\mathcal{H})$.

Exercise 1.5. Prove that if $P \in \mathrm{PSD}(\mathcal{H})$ and $M \in \mathrm{Lin}(\mathcal{H}, \mathcal{K})$, then $MPM^\dagger$ is positive.

1.2 Quantum States

Definition 1.6. A *density matrix* is a positive operator $\rho \in \mathrm{PSD}(\mathcal{H})$ with $\mathrm{tr}[\rho] = 1$. We put $S(\mathcal{H}) = \{\rho \in \mathrm{PSD}(\mathcal{H}) : \mathrm{tr}[\rho] = 1\}$.

Axiom 2. The state of a quantum system with Hilbert space $\mathcal{H}$ is described by a density matrix ρ. We will refer to ρ as a quantum state.

Example 1.7 (classical states)**.** In probability theory we have a finite set of outcomes Σ with a probability distribution p such that $p(x) \geq 0$, for $x \in \Sigma$, and $\sum_{x\in\Sigma} p(x) = 1$. If $\mathcal{H}$ has a basis $\{|x\rangle\}_{x\in\Sigma}$, then

$$\rho = \sum_{x\in\Sigma} |x\rangle\langle x|$$

is a density matrix. Conversely, any diagonal density matrix corresponds to a probability distribution. □

Example 1.8 (pure states)**.** Let $|\psi\rangle \in \mathcal{H}$ be a unit vector. Then $\rho = |\psi\rangle\langle\psi|$ is a density matrix since $\mathrm{tr}[|\psi\rangle\langle\psi|] = \mathrm{tr}[\langle\psi|\psi\rangle] = 1$. If $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, then

$$|\psi\rangle\langle\psi| = \begin{pmatrix} |\alpha|^2 & \alpha\overline{\beta} \\ \overline{\alpha}\beta & |\beta|^2 \end{pmatrix}.$$

Note that replacing $|\psi\rangle$ by $e^{i\theta}|\psi\rangle$ does not change the density matrix. □

Exercise 1.9. Is $\rho = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ a pure state?

Solution. Yes, since $\rho = |+\rangle\langle+|$, where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. □

Example 1.10 (general state)**.** Let ρ be a general quantum state. By Theorem 1.4, we can write

$$\rho = \sum_{i=1}^{d} p_i |\psi_i\rangle\langle\psi_i|,$$

where $p_i \geq 0$. Since ρ is diagonal in the basis $\{|\psi_i\rangle\}_{i=1}^d$ and trace is independent of the chosen basis, $\mathrm{tr}[\rho] = 1$ implies that $\sum_{i=1}^d p_i = 1$. So, the eigenvalues p_i's define a probability distribution. We may interpret ρ as if we have the pure state $|\psi_i\rangle\langle\psi_i|$ with probability p_i. We call ρ a *mixed state* if it is not pure.

However, note that the decomposition of ρ given above is not unique. It is possible that ρ may be written as a sum over some different probabilities and different states, even different number of states. □

Exercise 1.11. What is the state corresponding to 50% chance of $|0\rangle$ and 50% chance of $|+\rangle$?

Solution.

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +| = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}.$$

Note that the eigenvalues of ρ are $1/2 + \sqrt{5}/4$ and $1/2 - \sqrt{5}/4$, so the spectral decomposition of ρ gives a completely different decomposition of ρ compared to the one we started with. □

Definition 1.12. The state $\tau = \frac{1}{d}I$ is called *the maximally mixed state.*

1.3 Bloch Sphere

How does the set $S(\mathcal{H})$ look like? For two states $\rho_1, \rho_2 \in S(\mathcal{H})$ and $t \in [0, 1]$, the convex combination $t\rho_1 + (1-t)\rho_2$ is also a state. In fact, we have the following lemma.

Lemma 1.13. *The set $S(\mathcal{H}) \subseteq \mathrm{Lin}(\mathcal{H})$ is convex. Moreoever, the extreme points of $S(\mathcal{H})$ are exactly the pure states.*

For the particular case of one qubit, there is a visualization known as the *Bloch sphere*. Hermitian 2×2 matrices have the following (real) basis:

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

The last three matrices are known as *Pauli matrices*. Here are some of their properties:

$$X^2 = Y^2 = Z^2 = I, \quad XY = iZ = -YX, \quad \mathrm{tr}[X] = \mathrm{tr}[Y] = \mathrm{tr}[Z] = 0.$$

In particular, Pauli matrices are unitary and have eigenvalues ± 1.

Every 2×2 Hermitian matrix can be written in a unique way as a linear combination of the matrices I, X, Y, Z with all coefficients being real numbers. An arbitrary 2×2 Hermitian matrix ρ with $\mathrm{tr}[\rho] = 1$ can be expressed as

$$\rho = \frac{1}{2}(I + xX + yY + zZ) = \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}. \tag{1.3.1}$$

The coefficient $1/2$ in front of the I is fixed by the condition $\mathrm{tr}[\rho] = 1$. Further, $\mathrm{tr}[\rho] = 1$ implies that the eigenvalues of ρ are λ and $1 - \lambda$, for some $\lambda \in \mathbb{R}$. If we want ρ to be a quantum state, it also has to be positive, which is equivalent to requiring that $\lambda \geq 0$ and $1 - \lambda \geq 0$. Now, λ and $1 - \lambda$ are nonnegative if and only if $\det(\rho) = \lambda(1 - \lambda) \geq 0$. We compute

$$\det(\rho) = \frac{1}{4}((1+z)(1-z) - (x+iy)(x-iy)) = \frac{1}{4}(1 - x^2 - y^2 - z^2).$$

If we let $r = (x, y, z) \in \mathbb{R}^3$, then $\rho \geq 0$ if and only if $\|r\| \leq 1$. If ρ is a pure state, then $\mathrm{rank}(\rho) = 1$ and $\det(\rho) = 0$, which is equivalent to $\|r\| = 1$. Thus, pure states correspond exactly to the unit sphere in $\mathbb{R}^3$ and mixed states correspond to the open unit ball.

Exercise 1.14. What is the pure state $|\psi\rangle$ and its density matrix $|\psi\rangle\langle\psi|$ corresnponding to $r_1 = (0, 1, 0)$ and to $r_2 = (0, -1, 0)$.

Solution. By plugging r_1 into Eq. (1.3.1), we get

$$\rho_1 = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}.$$

We can easily calculate, that the eigenvector corresnponding to the eigenvalue 1 of ρ_1 is $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$. By Theorem 1.3, we get $\rho_1 = |\psi_1\rangle\langle\psi_1|$. Similarly, we get that $\rho_2 = |\psi_2\rangle\langle\psi_2|$, where $|\psi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)$. □

We have the following important bases of pure states that correspond to the intersection points of x, y, z-axes with the Bloch sphere:

- Z-basis: $\{|0\rangle, |1\rangle\}$, $r = (0, 0, 1)$, $-r = (0, 0, -1)$.
- X-basis: $\{|+\rangle, |-\rangle\}$, $r = (1, 0, 0)$, $-r = (-1, 0, 0)$
- Y-basis: $\{|+i\rangle, |-i\rangle\}$, $r = (0, 1, 0)$, $-r = (0, -1, 0)$.

1.4 Measurements

Definition 1.15. A *measurement* or a *positive operator valued measure (POVM)* on a Hilbert space $\mathcal{H}$ with a finite set of outcomes Ω is a function $\mu\colon \Omega \to \mathrm{PSD}(\mathcal{H})$ such that $\sum_{x\in\Omega} \mu(x) = I$. If $\mu(x)$ are projections, then the measurement is called *projective*.

Axiom 3. If we measure a state ρ using μ, then the probability of getting an outcome $x \in \Omega$ is $p(x) = \mathrm{tr}[\mu(x)\rho]$.

Remark 1.16. In general in quantum mechanics, it is not possible to perform measurements on a state without changing it. However, for now, we will not consider what happens with a state after the measurement (we can think that it is destroyed). We only get classical information about which outcome occured. We will discuss post-measurement states later in the course.

Exercise 1.17. Prove that if ρ is a quantum state and μ is a measurement, then $p(x) = \mathrm{tr}[\mu(x)\rho]$ defines a probability distribution on Ω.

Solution. By Theorem 1.4(d), we have $\mathrm{tr}[\mu(x)\rho] \geq 0$ since both $\mu(x)$ and ρ are positive. By the linearity of trace,

$$\sum_{x\in\Omega} \mathrm{tr}[\mu(x)\rho] = \mathrm{tr}[\sum_{x\in\Omega} \mu(x)\rho] = \mathrm{tr}[I\rho] = 1.$$

□

Example 1.18 (basis measurement)**.** Let $\{|\psi_i\rangle\}_{i=1}^d$ be an orthonormal basis and let $\mu(i) = |\psi_i\rangle\langle\psi_i|$ be the projection onto $|\psi_i\rangle$. We get

$$p(i) = \mathrm{tr}[|\psi_i\rangle\langle\psi_i|\rho] = \mathrm{tr}[\langle\psi_i|\rho|\psi_i\rangle] = \langle\psi_i|\rho|\psi_i\rangle.$$

In particular, if $\rho = |\phi\rangle\langle\phi|$ is a pure state, then we get

$$p(i) = \langle\psi_i||\phi\rangle\langle\phi||\psi_i\rangle = |\langle\psi_i|\phi\rangle|^2.$$

□

Exercise 1.19. What are the outcome probabilities when we measure $|0\rangle$ the X-basis?

Solution. We get $|+\rangle$ with probability $|\langle+|0\rangle|^2 = 1/2$ and $|-\rangle$ with probability $|\langle-|0\rangle|^2 = 1/2$. □

1.5 Bloch Sphere Revisited

Given $r = (x, y, z)$ with $\|r\| = 1$, the positive matrices

$$\mu(0) := \rho(r) = \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix} \quad \text{and} \quad \mu(1) := \rho(-r) = \frac{1}{2}\begin{pmatrix} 1-z & -x+iy \\ -x-iy & 1+z \end{pmatrix}$$

define a general qubit basis measurement. Performing this measurement on a quantum state with Bloch vector $s = (x', y', z')$, then the probability of obtaining the outcome 0 is

$$p(0) = \frac{1}{2} + \frac{1}{2}(r \cdot s) = \frac{1}{2} + \frac{1}{2}(xx' + yy' + zz').$$

Geometrically $r \cdot s$ is the projection of the Bloch vector s of a quantum state onto the axis defining the measurment.

1.6 Observables

Sometimes it is convenient to reformulate measurements in terms of *observables*. A *projective measurement* is a measurement where all measurement operators are projections, that is, we have a set of outcomes Ω and projections $\{P_x : x \in \Omega\}$ such that $\sum_{x\in\Omega} P_x = I$. Suppose that $\Omega \subseteq \mathbb{R}$. The *observable* associated to the projective measurment is an operator $O \in \mathrm{Lin}(\mathcal{H})$ given by

$$O = \sum_{x\in\Omega} xP_x.$$

Since x are real, O is Hermitain, and, conversely, each Hermitian operator O has a spectral decomposition of this form, thereby defining a projective measurement.

Remark 1.20. In physics, observables are often the preferred way to reason about measurements. Measuring the observable O is the same as performing the projective measurement defined by the spectral decomposition of O.

Example 1.21. Since $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$, the observable Z corresponds to measuring in the Z-basis with outcomes $\{\pm 1\}$. Similarly, since $X = |+\rangle\langle +| - |-\rangle\langle -|$, measuring X corresponds to measuring in the X-basis with outcomes $\{\pm 1\}$. More generally, we may measure a qubit in the basis given by the antipodal points $r = (x, y, z)$ and $-r$ on the Bloch sphere with outcomes $\{\pm 1\}$. This gives an observable

$$\begin{aligned} O(r) = \rho(r) - \rho(-r) &= \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1-z & -x+iy \\ -x-iy & 1+z \end{pmatrix} \\ &= \begin{pmatrix} z & x-iy \\ x+iy & -z \end{pmatrix}. \end{aligned}$$

□

An observable provides a compact formula for the expectation value of the measurement outcome. If $O = \sum_{x\in\Omega} xP_x$ is an observable and $\rho \in S(\mathcal{H})$, then

$$\mathbb{E}(\text{outcome}) = \sum_{x\in\Omega} x\mathbb{P}(\text{outcome } x) = \mathrm{tr}[O\rho].$$

In case $\rho = |\psi\rangle\langle\psi|$, then the right-hand hand side is equal to $\langle\psi|O|\psi\rangle$.

1.7 Uncertainty Relation

Let $|\psi\rangle$ be a pure state. Note that since $X = |+\rangle\langle+| - |-\rangle\langle-|$, we get the following:

$$|\langle\psi|X|\psi\rangle| = |p_X(1) - p_X(-1)| = |2p_X(1) - 1| = 2\max\{p_X(1), p_X(-1)\} - 1,$$

where $p_X(x)$ denotes the probability of getting outcome x after measuring in the X-basis. Clearly,

$$0 \leq |\langle\psi|X|\psi\rangle| \leq 1.$$

The upper bound is attained precisely when either $p_X(1) = 1$ or $p_X(-1) = 1$, that is, when the measurement outcome is certain. The lower bound is attained precisely when $p_X(1) = p_X(-1) = 1/2$, which means that the measurement is completely uncertain. Thus, $|\langle\psi|X|\psi\rangle|$ provides a meaningful way to quantify our uncertainty about the measurement outcome.

Using $Z = |0\rangle\langle0| - |1\rangle\langle1|$, we similarly obtain $|\langle\psi|Z|\psi\rangle| = 2\max\{p_Z(1), p_Z(-1)\} - 1$. Combining the two, we get

$$|\langle\psi|X|\psi\rangle| + |\langle\psi|Z|\psi\rangle| \leq 2.$$

Note that by the previous discussion, the equality can never be attained. However, there is a significant strenghtening:

Theorem 1.22 (Uncertainty relation for Pauli matrices)**.** *For every state $|\psi\rangle$, we have*

$$|\langle\psi|X|\psi\rangle| + |\langle\psi|Z|\psi\rangle| \leq \sqrt{2}.$$

Proof. Let $s_X, s_Z \in \{\pm1\}$ and let $A = s_X X + s_Z Z$. We need to show that

$$s_X\langle\psi|X|\psi\rangle + s_Z\langle\psi|Z|\psi\rangle = \langle\psi|A|\psi\rangle \leq \sqrt{2}.$$

Using Cauchy-Schwarz inequality, we get

$$\langle\psi|A|\psi\rangle \overset{\text{CS}}{\leq} \|A|\psi\rangle\| \leq \|A\|,$$

where $\|A\| = \sup_{\||\psi\rangle\|=1} \|A|\psi\rangle\|$ is the operator norm of A. Further,

$$A^\dagger A = A^2 = (s_X X + s_Z Z)(s_X X + s_Z Z) = I + s_X s_Z (XZ + ZX) + I = 2I.$$

This calculation shows that $A/\sqrt{2}$ is unitary. Since the operator norm of a unitary is one, we get

$$\|A\| = \sqrt{2}\left\|A/\sqrt{2}\right\| = \sqrt{2}.$$

□

We can interpret the quantity $\max\{p_X(1), p_X(-1)\}$ as the *guessing probability* $p_{\text{guess},X}$, that is, the maximal probability of guessing the outcome of X-basis measurement on the state $|\psi\rangle$ – the best option is to just guess the outcome with larger probability. Using this notation, we can rewrite the uncertainty relation from Theorem 1.22 as follows:

$$\begin{aligned} |\langle\psi|X|\psi\rangle| + |\langle\psi|Z|\psi\rangle| &\leq \sqrt{2} \\ 2\max\{p_X(1), p_X(-1)\} - 1 + 2\max\{p_Z(1), p_Z(-1)\} - 1 &\leq \sqrt{2} \\ p_{\text{guess},X} + p_{\text{guess},Z} &\leq 1 + \frac{\sqrt{2}}{2}. \end{aligned}$$

So the uncertainty relation from Theorem 1.22 gives a bound on the sum of probabilities of guessing the two measurement outcomes correctly.

1.8 TODO

- simultaneously diagonalizable operators
- more about uncertainty principle

Chapter 2

Multiple Quantum Systems

2.1 Multiple Systems

If we have classical random variables with outcome sets $\Sigma_1, \ldots, \Sigma_n$, then their joint distribution is a probability distribution on the product set

$$\Sigma_1 \times \cdots \times \Sigma_n = \{(x_1, \ldots, x_n) : x_j \in \Sigma_j\}.$$

So n bits are represented by an n-tuples.

Tensor product is precisely the quantum version of this. If we have bases $\Sigma_1, \ldots, \Sigma_n$ for the Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_n$, then the Hilbert space

$$\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$$

is the *tensor product* with basis

$$\{|x_1\rangle \otimes |x_2\rangle \otimes \cdots \otimes |x_n\rangle : |x_j\rangle \in \Sigma_j\}.$$

We will sometimes use the shorthand $|x_1\rangle|x_2\rangle \cdots |x_n\rangle$ or $|x_1 \cdots x_n\rangle$. Clearly, we have $\dim(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n) = \dim(\mathcal{H}_1) \cdots \dim(\mathcal{H}_n)$ since the basis elements are labelled by the elements of $\Sigma_1 \times \cdots \times \Sigma_n$.

Axiom 4. If we have multiple quantum systems with Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_n$, then the joint system has associated Hilbert space $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$.

Example 2.1. If we have two qubits, the joint Hilbert space is $\mathbb{C}^2 \otimes \mathbb{C}^2$, which is 4-dimensional and has the standard basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$.

The Hilbert space of n qubits is $(\mathbb{C}^2)^{\otimes n}$, which is 2^n-dimensional and has the standard basis $\{|x_1 \cdots x_n\rangle\}_{x_1,\ldots,x_n \in \{0,1\}}$. □

Notation. We will often label different quantum systems by $A, B, C, \ldots$ and denote the associated Hilbert spaces as $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C, \ldots$. The bases of these Hilbert spaces are usually denoted by $\Sigma_A, \Sigma_B, \Sigma_C, \ldots$. We will also write $\mathrm{Lin}(A) = \mathrm{Lin}(\mathcal{H}_A)$, $S(A) = S(\mathcal{H}_A)$, $\mathrm{PSD}(A) = \mathrm{PSD}(\mathcal{H}_A)$. We also write for example AB instad of $\mathcal{H}_A \otimes \mathcal{H}_B$. So, we may write $\rho_{AB} \in S(AB)$ for a quantum state shared by Alice and Bob and μ_A for a measurement on Alice's quantum system.

Tensor product of operators. If $M_A \in \mathrm{Lin}(A)$ and $N_B \in \mathrm{Lin}(AB)$, then $M_A \otimes N_B \in \mathrm{Lin}(AB)$ is defined by extending the following formula by linearity:

$$(M_A \otimes N_B)|\psi_A\rangle \otimes |\phi_B\rangle = (M_A|\psi_A\rangle) \otimes (N_B|\phi_B\rangle),$$

where $|\psi_A\rangle = \sum_a \psi_a|a\rangle$, $|\phi_B\rangle = \sum_b \phi_b|b\rangle$, and $|\psi_A\rangle|\psi_B\rangle = \sum_{a,b} \psi_a\phi_b|ab\rangle$.

Example 2.2. If $M_A = |a\rangle\langle a'|$ and $N_B = |b\rangle b'$, then $M_A \otimes N_B = |a\rangle\langle a'| \otimes |b\rangle\langle b'| = |ab\rangle\langle a'b'|$.

Definition 2.3. Let A and B be quantum systems. States of the form $\rho = \rho_A \otimes \rho_B \in S(AB)$, for $\rho_A \in S(A)$ and $\rho_B \in S(B)$ are called *product states*. A state, which is not a product state is called *correlated*.

Note that the previous definition makes sense since the tensor product of positive operators is positive and $\mathrm{tr}[M_A \otimes N_B] = \mathrm{tr}[M_A]\,\mathrm{tr}[N_B]$.

Example 2.4 (classical states)**.** A *joint probability distribution* $p_{XY} \in \mathbb{P}(XY)$ associates a probability $p_{XY}(x, y)$ to each pair $(x, y) \in \Sigma_X \times \Sigma_Y$. The classical state corresponding to XY has the density matrix

$$\rho_{XY} = \sum_{x,y} p_{XY}(x, y)|x, y\rangle\langle x, y| = \sum_{x,y} p_{XY}(x, y)|x\rangle\langle x| \otimes |y\rangle\langle y|.$$

Moreover, any classical joint state is of this form. The state ρ_{XY} is a product state if and only if X and Y are indpendent under the probability distribution p_{XY}. Thus, one can think of product states as the quantum generalization of independence in probability theory. Most of the quantum states are neither classical nor product states. □

Example 2.5 (maximally correlated state)**.** A classical state that is not a product state is for example the *maximally correlated state*:

$$\sigma_{AB} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|) = \frac{1}{2}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|).$$

Writing this in the basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ gives

$$\sigma_{AB} = \frac{1}{2}\begin{pmatrix} 1 & & & \\ & & & \\ & & & \\ & & & 1 \end{pmatrix}.$$

□

Exercise 2.6. Let

$$\rho_A = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \rho_B = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Express $\rho_A \otimes \rho_B$ in the standard basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$.

Solution.

$$\rho_A \otimes \rho_B = \frac{1}{4}\begin{pmatrix} 1\rho_B & 1\rho_B \\ 1\rho_B & 1\rho_B \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & & 1 & \\ & 1 & & 1 \\ 1 & & 1 & \\ & 1 & & 1 \end{pmatrix}.$$

□

Example 2.7 (pure product states)**.** If $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$, then pure product states are for example the basis states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$. Another example is the state $|+\rangle \otimes |+\rangle$. □

We will use abbreviation $|\psi_A\rangle|\phi_B\rangle = |\psi_A\rangle \otimes |\phi_B\rangle$ for pure states.

Example 2.8 (maximally entangled state)**.** A state that is neither classical nor a product state is the *maximally entangled state*:

$$|\Phi^+_{AB}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

The density matrix is

$$\rho_{AB} = |\Phi^+_{AB}\rangle\langle\Phi^+_{AB}| = \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|).$$

Writing this in the basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ gives

$$\rho_{AB} = \frac{1}{2}\begin{pmatrix} 1 & & & 1 \\ & & & \\ & & & \\ 1 & & & 1 \end{pmatrix}.$$

Recall that $|xy\rangle = |x\rangle \otimes |y\rangle$, so we can also write

$$\rho_{AB} = \frac{1}{2}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 1| \otimes |0\rangle\langle 1| + |1\rangle\langle 0| \otimes |1\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|).$$

□

2.2 The Partial Trace

If Alice does a measurement $\mu_A \colon \Omega \to \mathrm{PSD}(A)$, how to describe this on AB? Define

$$\mu_A \otimes I_B := \{\mu_A \otimes I_B : x \in \Omega\}.$$

This is a measurement on the whole system since tensor products of positive operators are positive and

$$\sum_{x\in\Omega} \mu_A(x) \otimes I_B = \left(\sum_{x\in\Omega} \mu_A(x)\right) \otimes I_B) = I_A \otimes I_B = I_{AB}.$$

More generally, suppose that Alice and Bob have measurements $\mu_A \colon \Omega_1 \to \mathrm{PSD}(A)$ and $\mu_B \colon \Omega_2 \to \mathrm{PSD}(B)$. This corresponds to a measurement on AB given by

$$\mu_A \otimes \mu_B := \{\mu_A(x_1) \otimes \mu_B(x_2) : (x_1, x_2) \in \Omega_1 \times \Omega_2\}.$$

The outcome probabilities for Alice do not depend on the choice of measurement for Bob.

Given a state $\rho_{AB} \in S(AB)$, what is the state ρ_A of Alice? We would like

$$\mathrm{tr}[\mu_A(x)\rho_A] = \mathrm{tr}[(\mu_A(x) \otimes I_B)\rho_{AB}],$$

for any measurement operator $\mu_A(x)$. We can proceed by computing the trace on the RHS. We chose bases $\{|a\rangle\}_{a\in\Sigma_A}$ and $\{|b\rangle\}_{b\in\Sigma_B}$ of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, which gives the product basis $\{|a\rangle\otimes|b\rangle\}$ for $\mathcal{H}_A\otimes\mathcal{H}_B$. We can expand the RHS above

$$\begin{aligned}\operatorname{tr}[(\mu_A(x)\otimes I_B)\rho_{AB}] &= \sum_{\substack{a\in\Sigma_A\\ b\in\Sigma_B}} \langle ab|(\mu_A(x)\otimes I_B)\rho_{AB}|ab\rangle \quad \text{use } |a\rangle\otimes|b\rangle = (I_A\otimes|b\rangle)|a\rangle \\ &= \sum_{a,b}\langle a|(I_A\otimes\langle b|)(\mu_A(x)\otimes I_B)\rho_{AB}(I_A\otimes|b\rangle)|a\rangle \\ &= \sum_{a,b}\langle a|\mu_A(x)(I_A\otimes\langle b|)\rho_{AB}(I_A\otimes|b\rangle)|a\rangle \\ &= \sum_{a}\langle a|\mu_A(x)\left(\sum_b (I_A\otimes\langle b|)\rho_{AB}(I_A\otimes|b\rangle)\right)|a\rangle.\end{aligned}$$

Finally, we define

$$\rho_A = \sum_b (I_A\otimes\langle b|)\rho_{AB}(I_A\otimes|b\rangle).$$

Definition 2.9 (partial trace)**.** Let A and B be systems with Hilbert spaces $\mathcal{H}_A$ nad $\mathcal{H}_B$ and choose basis $\{|b\rangle\}_{b\in\Sigma_B}$ for $\mathcal{H}_B$. Let $M_{AB}\in\operatorname{Lin}(AB)$. Then the *partial trace over B* of M_{AB} is

$$\operatorname{tr}_B[M_{AB}] = \sum_b (I_A\otimes\langle b|)M_{AB}(I_A\otimes|b\rangle).$$

For $\rho_{AB}\in S(AB)$, we call $\rho_A = \operatorname{tr}_B[\rho_{AB}]\in S(A)$ the *reduced state* of ρ_{AB} on A.

If we choose a basis $|a\rangle$ for $\mathcal{H}_A$, then the entries of the partial trace are given by

$$\langle a|\operatorname{tr}_B[M_{AB}]|a'\rangle = \langle a|\sum_b (I_A\otimes\langle b|)M_{AB}(I_A\otimes|b\rangle)|a'\rangle = \sum_b \langle ab|M_{AB}|a'b\rangle.$$

Further, we can write

$$M_{AB} = \sum_{a,a'\in\Sigma_A}\sum_{b,b'\in\Sigma_B} M_{ab,a'b'}|ab\rangle\langle a'b'| = \sum_{a,a'\in\Sigma_A}\sum_{b,b'\in\Sigma_B} M_{ab,a'b'}|a\rangle\langle a'|\otimes|b\rangle\langle b'|.$$

The partial trace is then

$$\operatorname{tr}_B[M_{AB}] = \sum_{a,a',b} M_{ab,a'b}|a\rangle\langle a'| = \sum_{a,a'}\left(\sum_b M_{ab,a'b}\right)|a\rangle\langle a'|.$$

For tensor product of operators $M_{AB} = N_A\otimes O_B$, the partial trace is given by

$$\operatorname{tr}_B[N_A\otimes O_B] = N_A\operatorname{tr}[O_B] = \operatorname{tr}[O_B]N_A.$$

This follows from the above since $M_{ab,a'b'} = N_{a,a'}O_{b,b'}$ and $\sum_b M_{ab,a'b} = M_{a,a'}\operatorname{tr}[O_B]$. Every operator can be written as a linear combination of tensor product operators, so this formula is sufficient to compute partial traces of arbitrary operators. Moreover, it shows that the notation of the reduced state is compatible with the notation for product states: $\rho_{AB} = \rho_A\otimes\rho_B$, then ρ_A and ρ_B are reduced states of A and B, respectively.

Lemma 2.10 (properties of partial trace)**.** *(a) The map* $\mathrm{tr}_B\colon \mathrm{Lin}(AB) \to \mathrm{Lin}(A)$ *is linear.*

(b) For $N_A \in \mathrm{Lin}(A)$ *and* $M_{AB} \in \mathrm{Lin}(AB)$*, we have* $\mathrm{tr}[(N_A \otimes I_B)M_{AB}] = \mathrm{tr}[N_A \,\mathrm{tr}_B[M_{AB}]]$.

(c) The partial trace does not depend on the choice of basis Σ_B.

(d) For $M_{AB} \in \mathrm{Lin}(AB)$*, we have* $\mathrm{tr}[\mathrm{tr}_B[M_{AB}]] = \mathrm{tr}[M_{AB}]$.

(e) If $P_{AB} \in \mathrm{PSD}(AB)$*, then* $\mathrm{tr}_B[P_{AB}] \in \mathrm{PSD}(A)$.

Proof. (a) The formula in the definition is linear in M_{AB}.

(b) We can follow the same calculation that we used in the derivation of partial trace with M_{AB} instead of ρ_{AB} and N_A instead of $\mu_A(x)$.

(c) Consider the formula in (b). Since LHS does not depend on the choice of basis, neither does the right-hand side. From linear algebra, we know that $\mathrm{tr}[AX] = \mathrm{tr}[BX]$ for all $X \in \mathrm{Lin}(\mathcal{H})$ if and only if $A = B$. Since the equation in (b) holds for all N_A, using the fact, $\mathrm{tr}_B[M_{AB}]$ is determined.

(d) Immedietelly from (b) by using $N_A = I_A$.

(e) By Theorem 1.4(d) it suffices to check that

$$\mathrm{tr}[Q_A \,\mathrm{tr}_B[P_{AB}]] \geq 0, \quad \text{for all } Q_A \in \mathrm{PSD}(A).$$

If $Q_A \geq 0$, then $Q_A \otimes I_B \geq 0$. So we can apply first (b) and then Theorem 1.4(d):

$$\mathrm{tr}[Q_A \,\mathrm{tr}_B[P_{AB}]] = \mathrm{tr}[(Q_A \otimes I_B)P_{AB}] \geq 0.$$

□

Example 2.11. Let $|\Phi^+_{AB}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and $\rho_{AB} = |\Phi^+_{AB}\rangle\langle\Phi^+_{AB}|$. Then

$$\rho_A = \mathrm{tr}_B[\rho_{AB}] = \mathrm{tr}_B[\frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}] = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This is the maximally mixed state. This also discribes the situation where Alice has a classical bit equal to zero or one with equal probability. If Alice cannot communicate with Bob, she cannot see the difference. □

We have three sources of mixed states in the quantum formalism: probabilistic mixtures (we receive ρ_x with probability p(x)), restricting to subsystem (even if $|\psi_{AB}\rangle$ is pure, the reduced state ρ_A can be mixed), measurement (if we perform measurement μ on ρ, we have some probability distribution on Ω, which can be described by a classical state).

Example 2.12 (marginal distributions)**.** For classical states, the partial trace reduces to marginal distirbutions. Let

$$\rho_{XY} = \sum_{x,y} p_{XY}(x,y)|x,y\rangle\langle x,y| = \sum_{x,y} p_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y|,$$

where p_{XY} is a joint probability distribution on $\Sigma_X \times \Sigma_Y$. The reduced state

$$\rho_X = \mathrm{tr}_Y[\rho_{XY}] = \sum_x \left(\sum_y p_{XY}(x,y)\right)|x\rangle\langle x| = \sum_x p_X(x)|x\rangle\langle x|,$$

where p_X is the *marginal distribution* of random variable X given by

$$p_X(x) = \sum_y p_{XY}(x, y).$$

The two random variables X and Y are idependent, that is, $p_{XY}(x, y) = p_X(x)p_Y(y)$, if and only if ρ_{XY} is a product state, that is, $\rho_{XY} = \rho_X \otimes \rho_Y$.

Another important concept in classical probability is the notion of conditional probability. In quantum information, there is no direct generalization of conditional probabilities for general quantum states. We come to this when dealing with quantum entorpy later.

2.3 Purification

The situation when with probability $p(x)$ we output a state ρ_x is described by the density matrix $\rho_A = \sum_{x\in\Omega} p(x)\rho_x$. However, given the density matrix, the interpretation can be non-unique.

Suppose now that we want to model a different situation, namely with probability $p(x)$ we output a state ρ_x, but now we also receive the value of x as well. This can be described by introducing a reference system X with Hilbert space $\mathbb{C}^\Omega$ and taking the joint state

$$\rho_{AX} = \sum_{x\in\Omega} p(x)\rho_x \otimes |x\rangle\langle x|.$$

It is easy to check that $\mathrm{tr}_A[\rho_{AX}] = \rho_A$. Once we actually recive outcome x, the state must be ρ_x (like in probability if we see outcome of die to be 6, then the die is in the state 6 with probability 1). We may consider X to be side information. It turns out that it is usefull to have quantum side information.

Definition 2.13. Given $\rho_A \in S(\mathcal{H}_A)$, a *purification* of ρ_A is a pure state $|\phi_{AR}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_R$ such that

$$\mathrm{tr}_R[|\phi_{AR}\rangle\langle\phi_{AR}|] = \rho_A.$$

The system R is called a *reference* or *purifying system*. We will refer to both $|\phi_{AR}\rangle$ and $\rho_{AR} = |\phi_{AR}\rangle\langle\phi_{AR}|$ as purification of ρ_A.

Lemma 2.14. *Every $\rho_A \in S(A)$ has a purification. The dimension $|R|$ of the purifying system can be taken to be* $\mathrm{rank}(\rho_A)$.

Proof. Let $r = \mathrm{rank}(\rho_A)$ and $\rho_A = \sum_{j=0}^{r-1} p_j|e_j\rangle\langle e_j|$ be a spectral decomposition. Put $\mathcal{H}_R = \mathbb{C}^r$ and

$$|\phi_{AR}\rangle = \sum_{j=0}^{r-1} \sqrt{p_j}|e_j\rangle \otimes |j\rangle.$$

Then

$$\mathrm{tr}_R[|\phi_{AR}\rangle\langle\phi_{AR}|] = \mathrm{tr}_R\left[\sum_{j,k=0}^{r-1} \sqrt{p_j p_k}|e_j\rangle\langle e_k| \otimes |j\rangle\langle k|\right] = \sum_j^{r-1} p_j|e_j\rangle\langle e_j| = \rho_A.$$

□

The proof does notu seo orthogonality given by the spectral decomposition. So we may take any decomposition $\rho_A = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ and take $\sum_j \sqrt{p_j}|\psi_j\rangle \otimes |j\rangle$ as purification. This does not give the optimal dimension of the reference system.

Exercise 2.15. Compute the purification of $\rho_A = \frac{1}{3}(|0\rangle\langle 0| + |+\rangle\langle +| + |1\rangle\langle 1|)$.

Solution.

$$|\phi_{AR}\rangle = \frac{1}{\sqrt{3}}(|0\rangle|0\rangle + |+\rangle|1\rangle + |1\rangle|2\rangle).$$

□

Lemma 2.16. *If $|\phi_{AR}\rangle$ and $|\phi_{AS}\rangle$ are purifications of ρ_A and $|R| \leq |S|$, then there is an isometry $V_{R\to S} \in \mathrm{Isom}(R, S)$ such that*

$$(I_A \otimes V_{R\to S})|\phi_{AR}\rangle = |\phi_{AS}\rangle.$$

In particular, when $S = R$, then the purificaton is unique up to a unitary.

Proof. Uses Schmidt decomposition. □

2.4 Schmidt decomposition

A standard result in linear algebra is that every matrix has a *singular value decomposition (SVD)*. Let $M \in \mathrm{Lin}(\mathcal{H}, \mathcal{K})$, there are bases $\{|e_j\rangle\}$ and $\{|f_j\rangle\}$ of $\mathcal{K}$ and $\mathcal{H}$, respectively, and $s_1 \geq \cdots \geq s_r > 0$, the *singular values* of M, such that

$$M = \sum_{j=1}^{r} s_j |e_j\rangle\langle f_j|.$$

The number of singular values equals $r = \mathrm{rank}(M)$.

To find SVD, we can take $M^\dagger M$, which is positive. It has eigenvalues s_j^2. The eigenvectors of $M^\dagger M$ and $MM^\dagger$ are the bases $\{|e_j\rangle\}$ and $\{|f_j\rangle\}$.

We may interpret a pure state $\psi_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ as a linear map $M \in \mathrm{Lin}(\mathcal{H}_A^*, \mathcal{H}_B)$ and apply the singular value decomposition to M. This leads to the Schmidt decomposition.

Theorem 2.17 (Schmidt decomposition)**.** *Let $|\psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a pure quantum state. Then there are bases $\{|e_j\rangle\}$ and $\{|f_j\rangle\}$ of $\mathcal{H}_A$ and $\mathcal{H}_B$, and positive number $s_1 \geq \cdots \geq s_r > 0$, where $r \leq \min(|A|, |B|)$, such that $\sum_j s_j^2 = 1$ and*

$$|\psi_{AB}\rangle = \sum_{j=1}^{r} s_j |e_j\rangle \otimes |f_j\rangle.$$

The numbers $s_1, \ldots, s_r$ are called Schmidt coefficients and r is calleed the Schmidt rank.

Proof. Let $\{|a\rangle\}$ and $\{|b\rangle\}$ be arbitrary basis for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Let M be the $|A| \times |B|$-matrix defined by $M_{ab} = \langle ab|\psi_{AB}\rangle$. We apply SVD to M and get

$$M = \sum_{j=1}^{r} s_j |e_j\rangle\langle f_j|.$$

Thus,

$$\begin{aligned} M_{ab} = \langle a|M|b\rangle &= \sum_{j=1}^{r} s_j \langle a|e_j\rangle\langle g_j|b\rangle \\ &= \sum_{j=1}^{r} s_j \langle a|e_j\rangle\overline{\langle b|g_j\rangle} \\ &= \sum_{j=1}^{r} s_j \langle a|e_j\rangle\langle b|f_j\rangle, \end{aligned}$$

here $|f_j\rangle$ denotes the vectors whose entries with respect to the basis $\{|b\rangle\}$ are the complex conjugate of the entries of $|g_j\rangle$, i.e., $\langle b|f_j\rangle = \overline{\langle b|g_j\rangle} = \langle g_j|b\rangle$, for all b. Note that $\{|f_j\rangle\}$ is also an orthonormal basis. We have

$$\begin{aligned} |\psi_{AB}\rangle &= \sum_{a,b} M_{ab}|a\rangle \otimes |b\rangle \\ &= \sum_{a,b}\sum_{j=1}^{r} s_j \langle a|e_j\rangle\langle b|f_j\rangle|a\rangle \otimes |b\rangle \\ &= \sum_{j=1}^{r}\sum_{a,b} s_j |a\rangle\langle a|e_j\rangle \otimes |b\rangle\langle b|f_j\rangle \\ &= \sum_{j=1}^{r} s_j |e_j\rangle \otimes |f_j\rangle. \end{aligned}$$

The state is normalized, so $\sum_j s_j^2 = 1$. □

Exercise 2.18. Compute the Schmidt decomposition of

$$|\psi_{AB}\rangle = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 1 \\ 2 \\ -2 \end{pmatrix}.$$

Solution. To do it systematically, reorganize the vector into the matrix M, e.g., we take $|01\rangle$ to $|0\rangle\langle 1|$. We get

$$M = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

Then we compute the SVD.

To get the left singular vectors, we take

$$MM^\dagger = \begin{pmatrix} 1/5 & 0 \\ 0 & 4/5 \end{pmatrix}.$$

The eigenvalues of $MM^\dagger$ are $1/5$ and $4/5$, so the Schmidt coefficients are

$$s_1 = \frac{1}{\sqrt{5}} \quad \text{and} \quad s_2 = \frac{2}{\sqrt{5}}$$

and the eigenvectors are just the computational basis $|0\rangle$ and $|1\rangle$.

To get the right singular vectors, we take

$$M^\dagger M = \begin{pmatrix} 1/2 & -3/10 \\ -3/10 & 1/2 \end{pmatrix}.$$

We have eigenvalue $1/5$ with eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $4/5$ with eigenvector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. After normalization, we get the right singular vectors $|+\rangle$ and $|-\rangle$. Finally, we get

$$|\psi_{AB}\rangle = \frac{1}{\sqrt{5}}|0\rangle \otimes |+\rangle + \frac{2}{\sqrt{5}}|1\rangle \otimes |-\rangle.$$

□

Lemma 2.19. *If $|\phi_{AB}\rangle$ is a pure state with Schmidt decomposition*

$$|\phi_{AB}\rangle = \sum_{i=1}^{r} s_i |e_i\rangle \otimes |f_i\rangle,$$

then the reduced density matrices are given by

$$\rho_A = \sum_{i=1}^{r} s_i^2 |e_i\rangle\langle e_i|, \quad \textit{and} \quad \rho_B = \sum_{i=1}^{r} s_i^2 |f_i\rangle\langle f_i|.$$

In particular, Schmidt rank and Schmidt coefficients are uniquely determined by the rank and the non-zero eigenvalues of the reduced states, respectively.

Proof. We compute

$$\rho_A = \mathrm{tr}_B[|\psi_{AB}\rangle\langle\psi_{AB}|] = \mathrm{tr}_B \left[\sum_{j,k=1}^{r} s_j s_k |e_j\rangle\langle e_k| \otimes |f_i\rangle\langle f_k| \right] = \sum_{j=1}^{r} s_j^2 |e_j\rangle\langle e_j|.$$

Similarly for ρ_B. □

Corollary 2.20. *A pure state $|\phi_{AB}\rangle$ with density matrix ρ_{AB} is a product state if and only if ρ_A is pure if and only if ρ_B is pure.*

2.5 Entanglement

Two random variables X and Y are independent if the corresponding classical state is a product state $\rho_{XY} = \rho_X \otimes \rho_Y$. If X and Y are not independent, the are correlated. For instance the maximally correlated state on two qubits

$$\rho_{XY} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|).$$

Non-classical correlations, so-called *entanglement*, creates fundamental difference between classical and quantum information theory.

Definition 2.21. A pure state $|\psi_{AB}\rangle \in \mathcal{H}_{AB}$ is called *entangled* if it is not a product state, i.e., there are no $|\psi_A\rangle$ and $|\psi_B\rangle$ such that $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$.

The following lemma follows directly from Theorem 2.20.

Lemma 2.22. *A pure state $\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$ is entangled if and only if the reduced density matrix ρ_A (or ρ_B) is not pure.*

Definition 2.23. A state $\rho_{AB} \in S(AB)$ is *maximally entangled state* if ρ_{AB} is pure and both reduced states are maximally mixed, i.e., $\rho_A = \frac{1}{|A|} I_A$ and $\rho_B = \frac{1}{|B|} I_B$.

The pervious lemma implies that maximally entangled states are indeed entangled since their reduced states are maximally mixed. We are not going to make this precise at this moment, but the "more mixed" the reduced states are, the "more entangled" the state is. We will come back to this later.

By Schmidt decomposition a pure state is maximally entangled if and only if its Schmidt rank is d and its Schmidt coefficients are all equal to $1/\sqrt{d}$, where $|A| = |B| = d$. In particualr maximally entangled states exists if and only if $|A| = |B|$, and they are of the form

$$|\Phi^+_{AB}\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^{d} |e_j\rangle \otimes |f_j\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B.$$

For example using the standard basis for both, we get

$$|\Phi^+_{AB}\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d.$$

Lemma 2.24 (properties of maximally entangled states)**.** *Let $|\Phi^+_{AB}\rangle$ be the maximally entangled state and let ρ_{AB} be its density matrix.*

(a) The reduced density matrices are maximally mixed: $\rho_A = \frac{1}{d} I_A$.

(b) For any $d \times d$ matrix M, we have $(M_A \otimes I_B)|\Phi^+_{AB}\rangle = (I_A \otimes M_B^T)|\Phi^+_{AB}\rangle$.

(c) For two $d \times d$ matrices M and N, we have $\langle\Phi^+_{AB}|M_A \otimes N_B|\Phi^+_{AB}\rangle = \frac{1}{d}\operatorname{tr}[M^T N] = \frac{1}{d}\operatorname{tr}[MN^T]$.

When $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^d$ and we use the standard basis, then the notation is clear. In general, the operators M_A, M_B^T, etc., are defined with respect to the same basis as those obtained in the Schmidt decomposition.

Exercise 2.25. Show that $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|{+}{+}\rangle + |{-}{-}\rangle)$.

Solution. Use Theorem 2.24(b). We have $(U \otimes U)|\Phi^+_{AB}\rangle = |\Phi^+_{AB}\rangle$. Now just apply the unitary

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

□

If we have classical state on systems X and Y, it is of the form

$$\rho_{XY} = \sum_{x,y} p_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y|.$$

It is a convex combination of the state $|x\rangle\langle x| \otimes |y\rangle\langle y|$. The next definition capters wider class of states where correlations between A and B are of classical nature and defines entanglement as its complement.

Definition 2.26. A state $\rho_{AB} \in S(AB)$ is *separable* if there is a collection of states $\rho_{A,x} \in S(A)$, $\rho_{B,x} \in S(B)$, for $x \in \Omega$, and a probability distribution p on Ω, for soem set Ω, such that

$$\rho_{AB} = \sum_{x\in\Omega} p(x)\rho_{A,x} \otimes \rho_{B,x}.$$

A state is called *entangled* if it is not separable.

Clearly classical states are separable. If a state is entangled, there is no choice of basis for A and B such that the state is classical in that basis. We can interpret separable states as follows: (1) Alice and Bob generate some shared classical random variable with outcome $x \in \Omega$. (2) Based on the outcome, Alice prepares $\rho_{A,x}$ and Bob prepares $\rho_{B,x}$. Thus, spearable states form a class of states where the correlations between Alice and Bob are of classical nature.

The definition of entanglement for pure states is consistent with the latter definition. Indeed, a pure product state is separable, and conversely we know that if a pure state is a convex combination of product states, then it must be product state itself. Also, if $|\psi_{AB}\rangle\langle\psi_{AB}| = \rho_A \otimes \rho_B$, then ρ_A and ρ_B must be pure since $1 = \operatorname{rank}(\rho_{AB}) = \operatorname{rank}(\rho_A)\operatorname{rank}(\rho_B)$.

2.6 TODO

- revise the proof Schmidt decomposition
- prove Theorem 2.16
- the no-communication theorem
- add problem about tensor product of observables

Chapter 3

Non-local games

Non-local games provide a mathematical framework to study the power of entanglement.

3.1 Basic definitions

Definition 3.1. A *non-local game* is a 6-tuple $\mathcal{G} = (I_A, I_B, O_A, O_B, \pi, V)$, where I_A and I_B are finite sets of *questions*, O_A and O_B are finite sets of *answers*, π is a probability distribution on $I_A \times I_B$, and $V \colon O_A \times O_B \times I_A \times I_B \to \{0, 1\}$ is a *winning predicate*.

The game is played as follows: the referee selects a pair of questions $(x, y) \in I_A \times I_B$ according to π and sends x to Alice and y to Bob. Alice responds with $a \in O_A$ and Bob with $b \in O_B$. The players win if $V(a, b, x, y) = 1$. The players can agree on a strategy before the game starts, but cannot communicate afterwards.

Definition 3.2. th Given a non-local game $\mathcal{G} = (I_A, I_B, O_A, O_A, \pi, V)$, a *strategy* is an element $p = (p(a, b|x, y))_{a,b,x,y} \in [0, 1]^{O_A \times O_B \times I_A \times I_B}$ such that for each $(x, y) \in I_A \times I_B$, we have

$$\sum_{(a,b) \in O_A \times O_B} p(a, b|x, y) = 1.$$

Given a strategy S for the game $\mathcal{G}$, the *succes probability* is defined as

$$\omega(\mathcal{G}, S) = \sum_{x,y} \pi(x, y) \sum_{a,b} V(a, b, x, y) p(a, b|x, y).$$

We say that a strategy S is *perfect* if $\omega(\mathcal{G}, S) = 1$. For a fixed collection of possible strategies $\mathcal{S}$, then we define the associated *value* of $\mathcal{G}$ by

$$\omega(\mathcal{G}, \mathcal{S}) = \sup_{S \in \mathcal{S}} \omega(\mathcal{G}, S).$$

Definition 3.3. A *deterministic strategy* for a non-local game $\mathcal{G} = (I_A, I_B, O_A, O_A, \pi, V)$ is given by a pair of functions $f \colon I_A \to O_A$ and $g \colon I_B \to O_B$.

Definition 3.4. A strategy p for a non-local game $\mathcal{G}$ is called *classical/randomized* if there exists a probability space (Ω, μ) and for all a, x and b, y measurable functions $p_A(a|x, \cdot), p_B(b|y, \cdot) \colon \Omega \to [0, 1]$ such that for all x, y, ω,

$$\sum_a p_A(a|x, \omega) = \sum_b p_B(b|y, \omega) = 1,$$

and

$$p(a,b|x,y) = \int_\omega p_A(a|x,\omega)p_B(b|y,\omega)d\mu(w).$$

The *classical value* of a non-local game $\mathcal{G}$ is denoted by $\omega_c(\mathcal{G})$.

Proposition 3.5. *If $\mathcal{G}$ is a non-local game, then $\omega_c(\mathcal{G})$ is attained by a deterministic strategy.*

Definition 3.6. A strategy p for a game $\mathcal{G}$ is *non-signalling* if every for every a, b, x, x', y, y' it holds

$$\sum_a p(a,b|x,y) = \sum_a p(a,b|x',y), \quad \text{and} \quad \sum_b p(a,b|x,y) = \sum_b p(a,b|x,y').$$

Definition 3.7. In a *(finite-dimensional) quantum strategy* Alice and Bob share a quantum state $\rho_{AB} \in S(\mathcal{H}_A \otimes \mathcal{H}_B)$, and their answers are the result of a measurement on their system. For each $x \in I_A$, there is a measurement $\mu_A^x \colon O_A \to \mathrm{PSD}(\mathcal{H}_A)$ on Alice's system and Alice answers the outcome of the measurement. Similalry, for each $y \in I_B$, Bob has measurement $\mu_B^y \colon O_B \to \mathrm{PSD}(\mathcal{H}_B)$ and answers the outcome of his measurement. This means that Alice and Bob answer a and b, when asked x and y, with probability

$$p(a,b|x,y) = \mathrm{tr}\left[(\mu_A^x(a) \otimes \mu_B^y(b))\rho_{AB}\right].$$

The *quantum value* of a game $\mathcal{G}$ is denoted by $\omega_q(\mathcal{G})$.

3.2 CHSH game

The *Clauser-Horne-Shimony-Holt (CHSH) game* has $I_A = I_B = O_A = O_B = \{0,1\}$. The winning predicate is given by

$$x \cdot y = (a+b) \mod 2 \iff x \wedge y = a \oplus b.$$

The probability distribution π is just the uniform distribution on $\{0,1\}^2$.

Classical strategy. We first consider classical strategies. Suppose that there exist a deterministic strategy, given by functions $f, g \colon \{0,1\} \to \{0,1\}$ such that $a = f(x)$ and $b = f(y)$, that wins for all pairs of questions. The winning condition implies that

$$\sum_{x,y\in\{0,1\}} f(x) + g(y) \mod 2 = \sum_{x,y\in\{0,1\}} x \cdot y = 1.$$

On the other hand,

$$\sum_{x,y\in\{0,1\}} f(x) + g(y) = 2\sum_{x\in\{0,1\}} f(x) + 2\sum_{y\in\{0,1\}} g(y),$$

which is even. Thus, $\omega_c(\mathcal{G}) \leq 3/4$. There is a deterministic strategy that achives this value, e.g., always answer $a = b = 0$. Therefore $\omega_c(\mathcal{G}) = 3/4$.

Quantum strategy. We will use observables with outcomes ± 1 corresponding to answers a and b as $(-1)^a$ and $(-1)^b$. We encode Alice's projective measurement with operators $\mu_A^x(a) = P_a^x$ corresponding to the question x by the observable $A^x = P_0^x - P_1^x$. Similarly, we encode Bob's projective measurement with operators $\mu_B^x(b) = Q_b^y$ corresponding to the question y by the observable $B^y = Q_0^y - Q_1^y$.

A quantum strategy is determined by a state $\rho_{AB} = |\psi\rangle\langle\psi|$, Alice's obserables $\{A^0, A^1\}$, and Bob's observables $\{B^0, B^1\}$. The probabilty that Alice and Bob give the same answer to x and y minus the probability that hey give different answers is given by

$$\mathbb{P}(a \neq b) - \mathbb{P}(a = b) = \langle\psi|A^x \otimes B^y|\psi\rangle.$$

Further,

$$\beta := 2\mathbb{P}(\text{win}) - 1 = \mathbb{P}(\text{win}) - \mathbb{P}(\text{lose}) = \frac{1}{4}\langle\psi|A^0 \otimes B^0 + A^0 \otimes B^1 + A^1 \otimes B^0 - A^1 \otimes B^1|\psi\rangle.$$

The quantity β is called the *bias* of the strategy.

Consider the states

$$|\psi_0(\theta)\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle, \qquad |\psi_0(\theta)\rangle = -\sin\theta|0\rangle + \cos\theta|1\rangle.$$

Clearly these form an orthonormal basis. The corresponding vectors on the Bloch sphere are $\vec{r}_\theta = (\sin 2\theta, 0, \cos 2\theta)$ and $-\vec{r}_\theta$, respectively. The corresponding observable is

$$O_\theta(\vec{r}) = \rho(\vec{r}) - \rho(-\vec{r}) = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}.$$

With the notation in place, we are ready to choose a strategy. As a quantum state, we pick the maximally entangled state $|\psi\rangle := |\Phi_{AB}^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. The Alice's measurement is given by the observables

$$A^0 = O_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = Z, \qquad A^1 = O_{\frac{\pi}{4}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = X.$$

Bob's measurement is the basis measurement corresponding to $\vec{r} = \frac{1}{\sqrt{2}}(1, 0, 1)$, for $y = 0$, and $\vec{s} = \frac{1}{\sqrt{2}}(-1, 0, 1)$, for $y = 1$. This corresponds to the measurement in the bases

$$\begin{aligned} \Big\{\cos(\frac{\pi}{8})|0\rangle + \sin(\frac{\pi}{8})|1\rangle, -\sin(\frac{\pi}{8})|0\rangle + \cos(\frac{\pi}{8})|1\rangle\Big\}, &\quad \text{for } y = 0, \\ \Big\{\cos(-\frac{\pi}{8})|0\rangle + \sin(-\frac{\pi}{8})|1\rangle, -\sin(\frac{\pi}{8})|0\rangle + \cos(-\frac{\pi}{8})|1\rangle\Big\}, &\quad \text{for } y = 1. \end{aligned}$$

In terms of observables, we have

$$B^0 = O_{\frac{\pi}{8}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{Z + X}{\sqrt{2}}, \qquad B^1 = O_{-\frac{\pi}{8}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix} = \frac{Z - X}{\sqrt{2}}.$$

We have $\vec{r} = (x, y, z)$ and $\vec{s} = (x', y', z')$ in the Bloch sphere. Recall that

$$O(\vec{r}) = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}.$$

If $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, by Theorem 2.24(c), we get

$$\langle\psi|O(\vec{r}) \otimes O(\vec{s})|\psi\rangle = \frac{1}{2}\operatorname{tr}[O(\vec{r})^T O(\vec{s})].$$

By directly computing this, trace, we get $\langle\psi|O(\vec{r}) \otimes O(\vec{s})|\psi\rangle = xx' - yy' + zz$. We use this to calculate β. Note that

$$\begin{aligned}\beta &= \frac{1}{4}\langle\psi|O_0 \otimes O_{\pi/8} + O_0 \otimes O_{-\pi/8} + O_{\pi/4} \otimes O_{\pi/8} - O_{\pi/4} \otimes O_{-\pi/8}|\psi\rangle \\ &= \frac{1}{4}\left((0 - 0 + \frac{1}{\sqrt{2}}) + (0 - 0 + \frac{1}{\sqrt{2}}) + (\frac{1}{\sqrt{2}} - 0 + 0) - (-\frac{1}{\sqrt{2}} + 0 - 0)\right) \\ &= \frac{1}{\sqrt{2}}.\end{aligned}$$

From this, we Immedietelly get $\mathbb{P}(\text{win}) = \frac{1}{2} + \frac{1}{2\sqrt{2}} = \cos^2(\pi/8) \approx 0.85$. We have proven the following theorem.

Theorem 3.8. *If $\mathcal{G}$ is the CHSH game, then*

$$\omega_q(\mathcal{G}) \geq \frac{1}{2} + \frac{1}{2\sqrt{2}} > \omega_c(\mathcal{G}).$$

The following *Tsirelson bound* shows that in fact we cannot do better.

Theorem 3.9 (Tsirelson bound)**.** *Let $\mathcal{G}$ be the CHSH game. Then*

$$\omega_q(\mathcal{G}) = \frac{1}{2} + \frac{1}{2\sqrt{2}}.$$

Proof. It remains to prove the upper bound. The key is that the bias β, defined above, is derived for an arbitrary quantum strategy. Let $|\psi_{AB}\rangle$ be the shared quantum state, which we assume to be pure and the measurements to be projective (we will later see that this is possible).

When we construct the observables A^x and B^y from a projective two-outcome measurement, we see that A^x and B^y are Hermitian with eigenvalues ± 1. Consequently $(A^x)^2 = I_A$ and $(B^y)^2 = I_B$. Let

$$M_{AB} = A^0 \otimes B^0 + A^0 \otimes B^1 + A^1 \otimes B^0 - A^1 \otimes B^1$$

We have

$$4\beta = \langle\psi_{AB}|M_{AB}|\psi_{AB}\rangle.$$

Using CS-inequality, we get

$$\begin{aligned}|\langle\psi_{AB}|M_{AB}|\psi_{AB}\rangle| &\leq \sqrt{\langle\psi_{AB}|M_{AB}M_{AB}^\dagger|\psi_{AB}\rangle}\sqrt{\langle\psi_{AB}|\psi_{AB}\rangle} \\ &= \sqrt{\langle\psi_{AB}|M_{AB}^2|\psi_{AB}\rangle}\end{aligned}$$

since M_{AB} is Hermitian and $|\psi_{AB}\rangle$ is normalized. We will now bound the right-hand side. To that end, we rewrite

$$\begin{aligned}
M_{AB}^2 &= \left(A^0 \otimes (B^0 + B^1) + A^1 \otimes (B^0 - B^1)\right)^2 \\
&= \left(A^0 \otimes (B^0 + B^1) + A^1 \otimes (B^0 - B^1)\right)\left(A^0 \otimes (B^0 + B^1) + A^1 \otimes (B^0 - B^1)\right) \\
&= (A^0)^2 \otimes (B^0 + B^1)^2 + (A^1)^2 \otimes (B^0 - B^1)^2 \\
&\quad + A^0A^1 \otimes (B^0 + B^1)(B^0 - B^1) + A^1A^0 \otimes (B^0 - B^1)(B^0 + B^1) \\
&= I_A^2 \otimes 4I_B - [A^0, A^1] \otimes [B^0, B^1].
\end{aligned}$$

Note that, in general, for hermitian matrices M and N with $M^2 = N^2 = I$, we have by triangle inequality and submultiplicativity

$$\|[M, N]\| = \|MN - NM\| \leq \|MN\| + \|NM\| \leq 2\|M\|\|N\| = 2.$$

Now, using the CS-inequality, we get

$$\begin{aligned}
|\langle\psi_{AB}|[A^0, A^1] \otimes [B^0, B^1]|\psi_{AB}\rangle| &\leq \big\|[A^0, A^1] \otimes [B^0, B^1]|\psi_{AB}\rangle\big\|\||\psi_{AB}\rangle\| \\
&\leq \big\|[A^0, A^1] \otimes [B^0, B^1]\big\|\||\psi_{AB}\rangle\|\||\psi_{AB}\rangle\| \\
&= \big\|[A^0, A^1]\big\|\big\|[B^0, B^1]\big\| \leq 4.
\end{aligned}$$

So, $\langle\psi_{AB}|M_{AB}^2|\psi_{AB}\rangle \leq 8$ and we get

$$4\beta = \langle\psi_{AB}|M_{AB}|\psi_{AB}\rangle \leq \sqrt{8} = 2\sqrt{2} \quad \Longrightarrow \quad \omega_q(\mathcal{G}) \leq \frac{1}{2} + \frac{1}{2\sqrt{2}}.$$

□

3.3 Mermin-Peris magic square game

Consider the magic square given in the Fig. 3.1 on the left. Each variable can be assigned values in $\{\pm 1\}$. We have six equations: each row and column corresponds to an equation $x_ix_jx_k = b$. So we have six Boolean linear equations.

The corresponding non-local game is defined as follows. The sets of questions are $X = Y = \{1, 2, 3\}$. For a pair of quesitons $(x, y) \in X \times Y$, Alice and Bob must respond with $a = (a_1, a_2, a_3), b = (b_1, b_2, b_3) \in \{-1, 1\}^3$, respectively. Alice and Bob win the game if the following three conditions are satisfied:

$$a_1 \cdot a_2 \cdot a_3 = r_x, \quad b_1 \cdot b_2 \cdot b_3 = c_y, \quad a_x = b_y.$$

In other words, their answers could form a part of the solution to the system of linear equations.

Classical strategy. First note that if there is a solution, then Alice and Bob can just agree beforehand, defining a deterministic strategy. On the other hand if $\omega(\mathcal{G}) = 1$, then there must be a determinstc strategy attaining this value. After Alice and Bob receive questions (x, y), then they must agree on the value of the (x, y) element, by the winning condition $a_x = b_y$. Doing this for every (x, y), we see that the strategy fixes a specific filling of the magic square

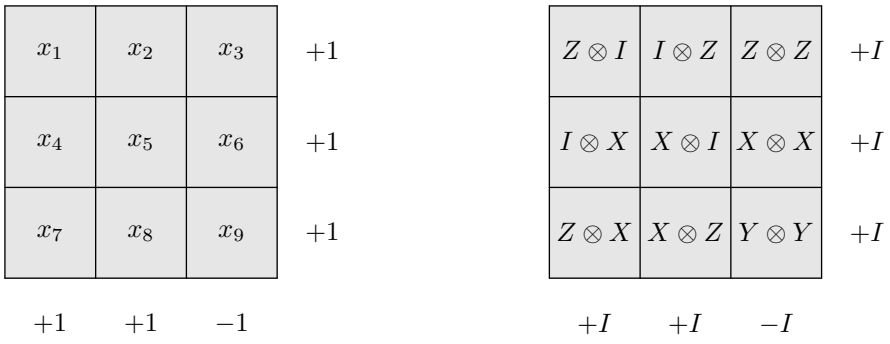

Figure 3.1: The Mermin-Peris magic square on the left. A 4-dimensional operator solution on the right.

which Alice and Bob are both using. Moreover, the other winning conditions require this filling to be a solution.

However, the system clearly has no solution. Indeed, suppose we have a solution such that the row products are all 1, the first two column products are 1 and the third column prodcut is -1. We get contradiction by taking all row products and obtaining 1, and taking all column products and obtaining -1.

Quantum strategy. To obtain a quantum strategy, we will make use of the *operator solution* given in Fig. 3.1 on the right. Each element of the operator solution is a tensor product of Pauli matrices on $\mathbb{C}^2 \otimes \mathbb{C}^2$. It is easy to see that each operator squares to the identity, operators in each row and each column pairwise commute, the product of each row and of the first two columns is the identity, and the product of the third column is the minus identity. In particular, the operators in each column and each row can be simultaneously diagonalized.

Alice and Bob share the following entangled state in $(\mathcal{H}_{A_1} \otimes \mathcal{H}_{B_1}) \otimes (\mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2})$:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_{A_1} \otimes |0\rangle_{B_1} + |1\rangle_{A_1} \otimes |1\rangle_{B_1}) \otimes \frac{1}{\sqrt{2}}(|0\rangle_{A_2} \otimes |0\rangle_{B_2} + |1\rangle_{A_2} \otimes |1\rangle_{B_2}). \tag{3.3.1}$$

We can use the Schmidt decomposition to check that $|\psi\rangle$ is in fact an entangled state. By multiplying and reordering the factors of the tensor product, we get

$$\frac{1}{2}(|00\rangle_A|00\rangle_B + |01\rangle_A|01\rangle_B + |10\rangle_A|10\rangle_B + |11\rangle_A|11\rangle_B), \tag{3.3.2}$$

which is already a Schmidt decomposition with all $s_i = 1/2$ and the standard bases. By Theorem 2.19, we have

$$\rho_A = \frac{1}{4}(|00\rangle_A + |01\rangle_A + |10\rangle_A + |11\rangle_A),$$

which is not pure. So by Theorem 2.22 it is entangled.

Upon receiving questions from the referee, Alice and Bob each measure their two quibits with the operator from the corresponding row or column in the magic square to determine their outputs. For example, if Alice receives $x = 3$, she measures $Z \otimes X$, $X \otimes Z$, $Y \otimes Y$, and answers (a_1, a_2, a_3) according to the outcomes. From the properties of the operators we discussed above it directly follows that their answers will satisfy the corresponding equations.

It remains to argue that their answers will agree on the common entry. We can check this individually, for example, consider the entry $Z \otimes X$. Using the equations

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

and the Eq. (3.3.2), we can conclude that

$$(Z \otimes X)_A \otimes (Z \otimes X)_B |\psi\rangle = |\psi\rangle.$$

Another way to see this is to rearrange $(Z \otimes X) \otimes (Z \otimes X)$ to $(Z_{A_1} \otimes Z_{B_1}) \otimes (X_{A_1} \otimes X_{A_2})$ ans use the fact that for maximally entangled state $|\Phi^+_{AB}\rangle$, we have

$$|\Phi^+_{AB}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|{+}{+}\rangle + |{-}{-}\rangle) = \frac{1}{\sqrt{2}}(|i, i\rangle + |{-}i, {-}i\rangle).$$

Then we can deduce this from Eq. (3.3.1).

Appendix A

Tensor Product