Density matrices:

\[ \rho = \sum_k p_k \vert \psi_k \rangle \langle \psi_k \vert \quad 0 \leq p_k, \sum_k p_k = 1 \]

\[ \text{Tr}(\rho) = 1, \quad \rho = \rho^+, \quad \rho \geq 0. \]

We can also construct density matrices as mixtures of other density matrices:

\[ \rho = \sum_k p_k \rho_k \quad \text{also a valid density matrix} \]

So the set of density matrices is convex.

Unitary (noiseless) evolution:

\[ \rho \rightarrow U \rho U^+ = \sum_k p_k U \vert \psi_k \rangle \langle \psi_k \vert U^+. \]

Projective measurement:

\[ \exists P \in \mathbb{P} \mathbb{P}, \quad \rho_l = P_l \rho P_l^+, \quad \sum_l P_l = I, \quad P_l P_l = \delta_{ll} P_l. \]

\[ \rho \rightarrow \rho_l = \frac{P_l \rho P_l}{\text{Tr} \rho_l} \quad \text{w/ prob} \quad \rho_l = \frac{\text{Tr} \rho P_l \rho_l}{\text{Tr} \rho \rho_l}. \]
Composite systems:

I. It is possible for each subsystem to have its own density matrix:

$$\rho = \rho_A \otimes \rho_B$$

Measurements on the subsystems are **uncorrelated**.

II. More generally, there can be a joint density matrix $\rho_{AB}$. If we can write

$$\rho_{AB} = \sum_k p_k \rho_A^k \otimes \rho_B^k,$$

where $p_k \geq 0$ and $\rho_A^k$, $\rho_B^k$ are valid density matrices on subsystems, then $\rho_{AB}$ is **separable**. Measurements of the subsystems are **classically correlated**.

Determining if a given matrix $\rho_{AB}$ is separable is, in general, a hard problem.

III. If $\rho_{AB}$ is **not separable**, then it is **entangled**.
With composite systems, we can introduce the concept of reduced or effective states on subsystem using the Partial Trace:

\[ \rho_A = \text{Tr}_B \{ \rho_{AB} \} \quad \rho_B = \text{Tr}_A \{ \rho_{AB} \} \]

The matrix \( \rho_A \) makes exactly the same predictions as \( \rho_{AB} \) for measurements on subsystem A alone:

\[ P_k = \text{Tr} \{ P_k \rho_A \} = \text{Tr} \{ (P_k \otimes I) \rho_{AB} \} \]

So \( \rho_A \) is like a marginal distribution on A. Likewise, \( \rho_B \). But note that

\[ \rho_{AB} \neq \rho_A \otimes \rho_B \]

in general; they give the same probabilities on the subsystems, but \( \rho_A \otimes \rho_B \) loses any correlations of \( \rho_{AB} \).

Partial trace:

\[ \text{Tr}_A \{ \sum_{A} \rho_{A} \otimes \rho_{B} \} = \text{Tr}_A \rho_{A} \otimes \text{Tr}_B \rho_{B} \]

Extends to general operators by linearity:

\[ \text{Tr}_A \{ \sum_{A} O_A \otimes O_B \} = \sum_{A} \text{Tr}_A O_A \otimes \text{Tr}_B O_B \]

In terms of a basis, \( \rho_{AB} = \sum_{ijkl} \rho_{ijkl} |ij\rangle \langle kl| \)

\[ \rho_A = \sum_{ijkl} \rho_{ijkl} |ij\rangle \langle kl| = \sum_{i} \left( \sum_{kj} \rho_{ijkl} |j\rangle \langle k| \right) |i\rangle \langle l| \]

\[ \rho_B = \sum_{ijkl} \rho_{ijkl} |ij\rangle \langle kl| = \sum_{j} \left( \sum_{k} \rho_{ijkl} |j\rangle \langle k| \right) |i\rangle \langle l| \]
An important fact is that we can define a reduced density matrix even in an entangled pure state (where there is no local state vector):

$$\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}| \Rightarrow \begin{cases} 
\rho_A = \text{Tr}_B \left( |\psi_{AB}\rangle\langle\psi_{AB}| \right) \\
\rho_B = \text{Tr}_A \left( |\psi_{AB}\rangle\langle\psi_{AB}| \right)
\end{cases}$$

Even though the global state is pure, the local state is mixed.

There is an important connection to the Schmidt decomposition:

$$|\psi_{AB}\rangle = \sum_j \sqrt{\lambda_j} |i_j\rangle_A |j\rangle_B$$

$$\Rightarrow |\psi_{AB}\rangle\langle\psi_{AB}| = \sum_{ij} \sqrt{\lambda_i \lambda_j} |i\rangle_A \langle i|_A \otimes |j\rangle_B \langle j|_B$$

$$\Rightarrow \rho_A = \sum_{ij} \sqrt{\lambda_i \lambda_j} |i\rangle_A \langle i|_A \text{Tr}_B \left( |i\rangle_B \langle i|_B \right) = \delta_{ij} \sum_i \lambda_i |i\rangle_A \langle i|_A$$

$$\rho_B = \sum_i \lambda_i |i\rangle_B \langle i|_B$$

so the Schmidt coefficients $\lambda_{ij}$ are the eigenvalues of both $\rho_A$ and $\rho_B$, and the Schmidt bases $|i\rangle_A$ and $|i\rangle_B$ are their respective eigenbases.
Representing classical prob. theory with density matrices.

Here, we just require the density matrix, and any projectors and/or observables, to all be diagonal in the same standard basis \( \{ |x\rangle \}_{x \in \mathbb{X}} \)

\[
P_x(x) \rightarrow \rho = \sum_x P_x(x) |x\rangle \langle x| \]

\[
\mathbb{X} \rightarrow \sum_x |x\rangle \langle x| \equiv \sum_x P_x
\]

so the expectation becomes

\[
\mathbb{E}[\mathbb{X}] = \text{Tr} \sum_x X \rho_x = \sum_x P_x(x) x
\]

and the probability of measurement outcome \( x \) becomes

\[
P_x = \text{Tr} \sum_X P_x \rho X = P_x(x).
\]

We can more generally define indicator functions, which are just projectors:

\[
I_A = \sum_{x \in A} P_x = \sum_{x \in A} |x\rangle \langle x|.
\]

This lets us define intersections and unions of subsets in the usual way.
Generalized measurements

In a projective measurement, the outcomes are specified by a set of orthogonal projectors \( \{P_j\} \).

A generalized measurement uses a set of general measurement operators \( \{M_j\} \). The only restriction is
\[
\sum_j M_j^d M_j = I
\]

The outcome probabilities are \( p_j = \langle \psi | M_j | \psi \rangle \) or \( p_j = \text{Tr} \{ M_j \rho M_j^d \} = \text{Tr} \{ M_j^d M_j \rho \} \).

After measurement, the state becomes
\[
|\psi\rangle \rightarrow |\psi_j\rangle / \sqrt{p_j}
\]
or
\[
P \rightarrow P_j = \frac{M_j \rho M_j^d}{P_j}
\]

In many cases (e.g., when a measurement is destructive), we only care about the probabilities. In that case, it suffices to specify the POVM elements \( E_j = M_j^d M_j \). \( P_j = \text{Tr} \{ E_j \rho \} \).

It is always possible to do a generalized measurement indirectly, using a unitary interaction with an ancilla, followed by a projective measurement
\[
U |\psi\rangle |0\rangle = \sum_j M_j |\psi\rangle |j\rangle \rightarrow |\psi_j\rangle |j\rangle \text{ w/ prob } P_j
\]

In this case
\[
M_j = \text{diag}(I \otimes |j\rangle \langle j|) U(I \otimes |0\rangle)
\]
CPTP maps

Unitary maps are not the most general linear evolution that takes density matrices to density matrices. Those are completely positive, trace-preserving maps: \( \rho \rightarrow N(\rho) \).

1. A positive map takes positive operators to positive operators.
2. \( N \) is completely positive if it is positive and \( N \otimes I_d \) is also positive for all \( d \).

Any CP map can be written as a Kraus decomposition:

\[
\rho \rightarrow \sum_k A_k \rho A_k^+. \quad \text{The } \{A_k\} \text{ are Kraus operators}
\]

For any CP map there are infinitely many Kraus decomps:

\[
N(\rho) = \sum_k A_k \rho A_k^+ = \sum_k \bar{B}_k \rho \bar{B}_k^+. \quad \bar{B}_k = \sum_k U_{lk} A_k.
\]

To be trace-preserving, we need

\[
\sum_k A_k^+ A_k = I.
\]

This is the same condition as a generalized measurement! We can think of a CPTP map as a generalized meas, where we are ignorant of the outcome. This also means that any CPTP map can be done by a unitary w/ an ancillary system followed by a partial trace:

\[
\rho \rightarrow N(\rho) = \text{Tr}_{\text{anc}} \left\{ U(\rho \otimes I_\text{anc}) U^+ \right\}
\]

\[
U |\psi\rangle_\text{anc} = \sum_k A_k |\psi\rangle_\text{anc}.
\]
The ancilla is often called the environment of the reference system. The process of being coupled to the environment is called decoherence.

**Q Channels** — In QIT we will often refer to a CPTP map as a quantum channel. Here are some examples:
- **Unitary channel**: \( p \rightarrow UpU^\dagger \).
- **Bit-flip channel**: \( p \rightarrow (1-p)p + xpX \).
- **Phase-flip channel**: \( p \rightarrow (1-p)p + zpZ \).
- **Pauli channel**: \( p \rightarrow (1-px-py-pz)p + px XpX + py YpY + pz ZpZ \).

When \( px = py = pz = \frac{p}{3} \) this is the depolarizing channel. This channel is equivalent to removing the qubit and replacing it with a maximally mixed state with some probability \( q \):

\[
\frac{1}{4}(p + xpX + ypY + zpZ) = \Pi = \frac{1}{2} I \quad \text{for all} \quad p, \quad \text{so}
\]

\[
p \rightarrow (1-q)p + q \Pi = (1-p)p + \frac{p}{3}(xpX + ypY + zpZ)
\]

\[
q = \frac{4}{3} p.
\]

Note that Kraus operators need not be square. We can map between spaces of different dimensions.
e.g., Partial trace \( P_{AB} \rightarrow P_A = \sum A_k P_{AB} A_k^+ \)
\( A_k = I \otimes b_k \), where \( \{ b_k \}_B \) is a basis.

**Erasure channel** \( \rho \rightarrow (1-p) \rho + p |e \rangle \langle e | \)
where \( | e \rangle \) is an extra state orthogonal to \( |0 \rangle \) and \( |1 \rangle \). (Erasures can thus be detected.)

**Classical-quantum channel** First measure \( \rho \) in an orthonormal basis \( \{ |k \rangle \} \), then output a state \( \sigma_k \) conditioned on the outcome:
\( \rho \rightarrow \sum_k |k \rangle \langle k | \otimes \sigma_k \), output state
(This is an entanglement-breaking channel.)

**Quantum Instrument** This is partway between a generalized meas and a CPTP map—like a meas, with partial info about the outcome:
\( \rho \rightarrow \sum_{j} E_j (\rho) \otimes |j \rangle \langle j | \)
(transformed state)
\( E_j (\rho) = \sum_{k} A_{jk} \rho A_{jk}^+ \), \( \sum_{j} \sum_{k} A_{jk}^+ A_{jk} = I \)

**Conditional quantum channels** take both a classical and quantum input and produce a Q output:
\( 1 \otimes m_1 \otimes \rho \rightarrow E_m (\rho), \ E_m (\rho) = \sum_k A_{mk} \rho A_{mk}^+ \)
\( \sum_k A_{mk}^+ A_{mk} = I \) for all Kraus operators:
\( \sum_{k} |m_1 \rangle \langle m_1 | \otimes A_{mk}^2 \)
Graphically, we can include Q channels in circuit diagrams.

Channel
Alice \( P \)  \( \xrightarrow{\text{noisy channel}} \) \( N \)
Bob

Classical-Quantum
A \( P \)  \( \xrightarrow{\text{M}} \) \( m \)
B \( P \)
breaks entanglement

Quantum Instrument
\( P \)  \( \xrightarrow{\epsilon_j} \) \( j \leftarrow \text{partial meas. outcome} \)
\( \epsilon_j(P)/p_j \leftarrow \text{renormalized} \)

Conditional Q Channel
This could be used to simultaneously encode classical + Q into.
A \( P \)
B \( m \)
\( \epsilon_m(P) \)