Quantum Information Summer School, Karachi
Lectures on Quantum Processes
Review material and notation

Fabio Costa

1 States, measurements, transformations

1.1 Closed systems

Pure states in quantum theory are represented as elements of a complex Hilbert space, denoted as "kets": $|\psi\rangle \in \mathcal{H}$. A Hilbert spaces is a vector space endowed with a scalar product. We will restrict to finite-dimensional Hilbert space for simplicity, although almost everything extends to infinite dimension, modulo appropriate definitions.

Given a Hilbert space $\mathcal{H}$, its **dual space** $\mathcal{H}^*$ is defined as the space of complex-valued linear functionals over $\mathcal{H}$. A functional is denoted as a "bra": $\langle\phi| \in \mathcal{H}^*$; its action on a ket is denoted as $\langle\phi|\psi\rangle \in \mathbb{C}$. The scalar product in $\mathcal{H}$ induces an isomorphism between $\mathcal{H}$ and $\mathcal{H}^*$: $\langle\phi| \equiv |\phi\rangle^\dagger$. In terms of an orthonormal basis $|e_j\rangle$, and the corresponding dual basis $\langle e_i|e_j\rangle = \delta_{ij}$, a ket $|\phi\rangle = \sum_j \phi^j |e_j\rangle$ corresponds to the bra $\langle\phi| = \sum_j (\phi^j)^* \langle e_j|$. The action of a bra on a ket is, in terms of components, $\langle\phi|\psi\rangle = \sum_j (\phi^j)^* \psi^j$, which corresponds to the scalar product between states $|\psi\rangle$ and $|\phi\rangle$.

> **Notation.** ψ^j denotes a complex number, not a vector. It represents the j-th component of a ket in a given basis, $\psi^j = \langle e_j|\psi\rangle$ (similarly for bra components). Vectors in Hilbert space (or dual space) are always denoted as kets $|\psi\rangle$ (or bras $\langle\phi|$).

Physically, a "bra" represents a measurement that yields a particular outcome. The isomorphism between bras and kets means that, for each state $|\phi\rangle$, there exists a measurement that can be interpreted as the question "is the system in state $|\phi\rangle$?". The probability for this question to have a positive answer, given the system was in state $|\psi\rangle$ when measured, is given by

$$P(\phi|\psi) = \frac{|\langle\phi|\psi\rangle|^2}{\langle\phi|\phi\rangle\langle\psi|\psi\rangle}. \tag{1}$$

We usually work with normalised states, $\langle\phi|\phi\rangle = \langle\psi|\psi\rangle = 1$, so that the denominator can be omitted. In the following, we will assume states to be normalised unless otherwise stated. The complex number $\langle\phi|\psi\rangle$ is called **probability amplitude**.

The presentation above takes "ket" states as the starting point and derives "bra" measurements as their duals. The reverse direction is possible: take as a starting point the space of elementary yes/no questions, making up the Hilbert space of bras, and define ket states as linear functionals over this space. This duality between "how something can be" and "what questions can be asked" will be a recurring theme in these lectures.

Multiple systems. The composition of systems in quantum mechanics is described by the tensor product. Given two Hilbert spaces $\mathcal{H}^A$, $\mathcal{H}^B$, describing the possible states of two different systems (e.g., the internal states of two atoms), all possible states of the pair of systems live in the product Hilbert space $\mathcal{H}^A \otimes \mathcal{H}^B$. If $\{|e_j\rangle\}_{j=1}^{d_A}$, $\{|f_k\rangle\}_{k=1}^{d_B}$ are orthonormal bases of the two spaces, $\{|e_j\rangle \otimes |f_k\rangle\}_{j=1}^{d_A}{}_{k=1}^{d_B}$ is a basis of the joint space. States of the form $|\psi_1\rangle^A \otimes |\psi_2\rangle^B$ are called *product states*; states that are not product, $|\psi\rangle^{AB} = \sum_{jk} \psi^{jk} |e_j\rangle \otimes |f_k\rangle$, $\psi^{jk} \neq \psi_1^j \psi_2^k$, are called *entangled.*

> **Notation.** We often simplify the notation for basis states as $|e_j\rangle \equiv |j\rangle$.

The dual space of (and the scalar product in) the composite space is defined component-wise:

$$\left(\langle\phi_1|^A \otimes \langle\phi_2|^B\right) \left(|\psi_1\rangle^A \otimes |\psi_2\rangle^B\right) := \langle\phi_1|\psi_1\rangle\langle\phi_2|\psi_2\rangle.$$

This extends to entangled states by linearity.

> **Notation.** The superscripts $A, B, \ldots$ typically refer to the space in which the object (here, bra or ket) is defined, while it does not distinguish objects from each other. This means that an expression like $|\psi\rangle^A \otimes |\psi\rangle^B$ represents the tensor product of two copies of *the same* state $|\psi\rangle$. The tensor product symbol and the system labels are sometimes omitted, when they can be inferred from context: $|\psi\rangle^A \otimes |\phi\rangle^B \equiv |\psi\rangle^A |\phi\rangle^B \equiv |\psi\rangle|\phi\rangle$.

We will also use the *partial scalar product*:

$$\langle\phi|^B \left(|\psi_1\rangle^A \otimes |\psi_2\rangle^B\right) \equiv \left(\mathbb{1}^A \otimes \langle\phi|^B\right) \left(|\psi_1\rangle^A \otimes |\psi_2\rangle^B\right) = |\psi_1\rangle^A \langle\phi|\psi_2\rangle,$$

where $\mathbb{1}$ denotes the identity transformation, $\mathbb{1}|\psi\rangle = |\psi\rangle$.

A **linear transformation** (or map, or operator) $L : \mathcal{H}^A \to \mathcal{H}^B$ takes elements from a Hilbert space $\mathcal{H}^A$ to another Hilbert space $\mathcal{H}^B$, with the property

$$L\left(\alpha|\psi_1\rangle + |\psi_2\rangle\right) = \alpha L|\psi_1\rangle + L|\psi_2\rangle$$

for every $\alpha \in \mathbb{C}$ and $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}^A$. We use the same symbol to denote an opertor L and its dual $L : \mathcal{H}^{B^*} \to \mathcal{H}^{A^*}$, defined as $L\left(\langle\phi|^B\right)|\psi\rangle^A := \langle\phi|^B \left(L|\psi\rangle^A\right) \equiv \langle\phi|L|\psi\rangle$ (i.e., it acts on the bras "from the right"). *Matrix elements* are the components $L^{ij} = \langle i|L|j\rangle$ for some chosen bases $\left\{|i\rangle^B\right\}_i$, $\left\{|j\rangle^A\right\}_j$. The *adjoint* $L^\dagger : \mathcal{H}^B \to \mathcal{H}^A$ is defined as $L^\dagger|\phi\rangle := (\langle\phi|L)^\dagger$ and has matrix elements equal to the conjugate transpose of the original operator: $(L^\dagger)^{ji} = \left(L^{ij}\right)^*$.

There is a natural isomorphism between the space of linear operators, denoted $\mathcal{L}\left(\mathcal{H}^A, \mathcal{H}^B\right)$ and the product of the dual of the input space with the output space, $\mathcal{H}^B \otimes \mathcal{H}^{A^*}$. This is implicit in the bra-ket notation; for example, the linear transformation $L_{\phi_1 \to \phi_2} |\psi\rangle^A := |\phi_2\rangle^B \langle \phi_1|\psi\rangle$ corresponds to $|\phi_2\rangle^A \otimes \langle \phi_1|^B \equiv |\phi_2\rangle^A \langle \phi_1|^B$. The space of operators of a space onto itself is denoted $\mathcal{L}(\mathcal{H}) \equiv \mathcal{L}(\mathcal{H}, \mathcal{H}) \equiv \mathcal{H} \otimes \mathcal{H}^*$.

Here are some relevant definitions and examples of linear operators.

- Every pure state $|\psi\rangle \in \mathcal{H}$ can be seen as a map $\mathbb{C} \to \mathcal{H}$, acting as $\mathbb{C} \ni \alpha \mapsto \alpha|\psi\rangle$.
- Bras are, by definition, linear maps $\mathcal{H} \to \mathbb{C}$.
- An operator U is **unitary** if $U^\dagger U = UU^\dagger = \mathbb{1}$. (These operators are invertible and preserve the scalar product). The dimension of input and output spaces of a unitary operator must be equal.
- V is an **isometry** if $V^\dagger V = \mathbb{1}$. (These are the operators that preserve the scalar product, without necessarily being invertible, in the sense that they might not have a right inverse.) The output dimension of an isometry must be larger or equal to the input dimension. Unitaries and ket states are particular examples of isometries, while bras are not.
- A **projector** is a *self-adjoint* operator, $\Pi^\dagger = \Pi \in \mathcal{L}(\mathcal{H})$, such that $\Pi^2 = \Pi$. As a special case, $|\psi\rangle\langle\psi|$ is a *rank-1* projector for any normalised state $\langle\psi|\psi\rangle = 1$. We use the short-hand notation $[\psi] \equiv |\psi\rangle\langle\psi|$.
- An operator $L \in \mathcal{L}(\mathcal{H})$ is called **positive semidefinite** if $\langle\psi|L|\psi\rangle \geq 0$ for every $|\psi\rangle \in \mathcal{H}$. We express this property as $L \geq 0$. L is *positive definite* if $\langle\psi|L|\psi\rangle > 0$ for all $|\psi\rangle$, in which case we write $L > 0$. Projectors are positive semidefinite operators; the only positive definite projector is $\mathbb{1}$.
- The **trace** of an operator L with equal input and output dimensions is the sum of its diagonal elements, $\operatorname{tr} L := \sum_j \langle j|L|j\rangle$. The *partial trace* of an operator on a composite system is defined in terms of the partial scalar product: $\operatorname{tr}_B L^{AB} := \sum_j \langle j|^B L^{AB} |j\rangle^B$.

Physically, isometries represent the most general way to transform a system so that pure states are mapped to pure states, and such that the transformation can be performed with probability 1. Unitaries are the particualar case where the inverse, $U^\dagger$ is also an isometry (and thus also a deterministic transformation). The time evolution of a closed system, as well as symmetry transformations (such as rotations or translations), are typically described as unitaries. An example of an isometry that is not a unitary is a transformation where a system is transformed in a reversible why, while some extra system is appended to it:

$$V^{AB} = U^A \otimes |\psi\rangle^B.$$

Notation. A single superscript on a linear operator usually denotes its output space. Bras are an exception, as we label them with the Hilbert space they act on. This notation can lead to some ambiguity when input and output spaces are different. Context and appropriate descriptions should clarify such ambiguities. The notation $L^{A\to B}$ can be used if extra clarity is needed to specify input and output space.

A **projection-valued measure (POM)** is a complete set of projectors, $\{\Pi_j\}_j$ with $\sum_j \Pi_j = \mathbb{1}$. Physically, it represents a measurement procedure, with j labelling the measurement outcome. Given a system prepared in a pure state $|\psi\rangle$, the probability to observe the outcome j is

$$P(j|\psi) = \langle\psi|\Pi_j|\psi\rangle. \tag{2}$$

A particular case is a binary measure $\{[\psi], \mathbb{1} - [\psi]\}$, for which one recovers Eq.(1).

> **Jargon.** A *measure* typically refers to a mathematical construction, while *measurement* refers to a physical procedure.

Note that the formulas for probabilities (1), (2) make no reference to what happens to the state after the measurement. As we will see below, this depends on a more complete description of the measurement procedure and is not uniquely defined by the POM. Nonetheless, "projective measurement" usually refers to a procedure where, upon observing an outcome corresponding to a projector Π, the state is transformed as

$$|\psi\rangle \mapsto \Pi|\psi\rangle. \tag{3}$$

Several textbooks introduce Eq. (4) as a postulate of quantum mechanics. However, it turns out that state update can be derived from a quantum description of measurements. Furthermore, many sources renormalise the post-measurement state: $|\psi\rangle \mapsto \frac{\Pi|\psi\rangle}{\langle\psi|\Pi|\psi\rangle}$, making state update a non-linear transformation. We will avoid this here, and interpret the sub-normalised state $\Pi|\psi\rangle$ as the result of an operation that does not happen with unit probability. The norm of the post-measurement state, $\|\Pi|\psi\rangle\| = \langle\psi|\Pi|\psi\rangle$, represents the probability for the transformation to succeed. This is convenient when describing multiple measurements: upon observing subsequent outcomes i, j, the system is transformed as

$$|\psi\rangle \mapsto \Pi_j\Pi_i|\psi\rangle. \tag{4}$$

The probability for this two happen, i.e., the joint probability for observing i, j, is given by

$$P(i,j|\psi) = \|\Pi_j\Pi_i|\psi\rangle\| = \langle\psi|\Pi_i\Pi_j\Pi_i|\psi\rangle. \tag{5}$$

1.2 Open systems

The space of linear operators $\mathcal{L}(\mathcal{H})$ inherits a scalar product from the scalar product on $\mathcal{H}$. This is defined as

$$(A, B)_{\mathrm{HS}} := \operatorname{tr} A^\dagger B, \tag{6}$$

ans is called the *Hilbert-Schimdt* product. We will refer to a space endowed with this product as *Hilbert-Schimdt space*. We will use a system's label to refer to the corresponding Hilbert-Schmidt space, $\mathcal{L}\left(\mathcal{H}^A\right) \equiv A$.

Hilbert-Schmidt spaces are used to represent systems of which we have partial information. The structure of states-measurements-transformations based on Hilbert-Schmidt space is similar to that on Hilbert space.

A **state** of is a positive semidefinite operator $\rho \geq 0$ (ρ is often called *density matrix*, or *density operator*). A state is called *pure* if $\rho = [\psi]$ for some ket $|\psi\rangle$. It is called *mixed* otherwise. Mixed states can arise in two different ways. If we have a system that can be in one of several possible pure states $|\psi_1\rangle, \ldots, |\psi_n\rangle$ with probabilities $p_1, \ldots, p_n$, we describe it as the **convex mixture**

$$\rho = \sum_j p_j \, [\psi_j] \, ,$$

with $p_j \geq 0$, $\sum_j p_j = 1$. As a second possibility, given a composite system in a pure state $|\psi\rangle^{AB}$, the **reduced state** is defined as

$$\rho^A := \mathrm{tr}_B \, [\psi]^{AB} \, .$$

Unless otherwise stated, our density matrices will be normalised, $\mathrm{tr}\, \rho = 1$. A sub-normalised density matrix physically describes a procedure that prepares the state with probability $\mathrm{tr}\, \rho$.

An **effect** e on a system A is a real-valued linear functional, with $0 \leq e(\rho) \leq \mathrm{tr}\, \rho$. It describes an outcome of a measurement on A, where $e(\rho)$ is the probability for the outcome to be observed. Effects live in the dual linear space to A, $A^* \equiv \mathcal{L}(\mathcal{H})^*$. The Hilbert-Schmidt scalar product induces an isomorphism between A and A^*: every effect $e \in A^*$ corresponds to a positive semidefinite operator with $0 \leq E \leq \mathbb{1}$ such that

$$e(\rho) = \mathrm{tr}\, E\rho. \tag{7}$$

Note that every positive operator is also self-adjoint, so that $\mathrm{tr}\, E\rho = \mathrm{tr}\, E^\dagger \rho = (E, \rho)_{\mathrm{HS}}$.

As with pure states, it is possible to start with measurements as the primary concept and derive states. One can start from the space of effects, represented as positive, sub-unity operators $0 \leq E \leq \mathbb{1}$, and introduce states as linear functionals on effects, $E \mapsto \rho(E)$, with $0 \leq \rho(E) \leq 1$. Coming from this direction, effect operators represent the possible outcomes that can be obtained when measuring a system, while states are the possible ways to assign probabilities to measurement outcomes. We will typically identify any effect e with the corresponding operator E.

A **Positive-Operator-Valued Measure (POVM)**, is a set of positive semi-definite operators $\{E_j\}_j$, $E_j \geq 0$, with $\sum_j E_j = \mathbb{1}$. This represents a measurement procedure with outcomes labelled by j. The probability to observe outcome j, given that the system was initially in state ρ, is

$$P(j|\rho) = \mathrm{tr}\, E_j \rho. \tag{8}$$

The condition $\sum_j E_j = \mathbb{1}$ ensures that at least one outcome is observed, i.e., $\sum_j P(j|\rho) = 1$. Note that, for a pure state $\rho = [\psi]$ and a POM $E_j = \Pi_j$, the probability rule (8) reduces to (2): $\mathrm{tr}\, \Pi_j \, [\psi] = \langle\psi|\Pi_j|\psi\rangle$. We will refer to the probability rule (8) as the **Born rule**, although sometimes the name is associated with the particular cases (2) or (1).

Just as for pure states, composite systems are described by tensor products. A **product state** $\rho_1^A \otimes \rho_2^B$ describes two uncorrelated systems, while a product effect $E_i^A \otimes E_j^B$ describes two independent measurements. Indeed, uncorrelated states give uncorrelated outcomes when measured independently:

$$P(i,j|\rho_1^A \otimes \rho_2^B) = (\operatorname{tr} E_i\rho_1)\,(\operatorname{tr} E_j\rho_2) = P(i|\rho_1)P(j|\rho_2). \tag{9}$$

For non-product states, we distinguish *separable states*, which can be written as convex mixtures of product states, $\rho^{AB} = \sum p_j \rho_j^A \otimes \sigma_j^B$, $p_j \geq 0$, $\sum_j p_j = 1$, and *entangled states*, which cannot be decomposed in this way.

Linear transformations $\mathcal{M} : A \to B$ on operator spaces are sometimes called **superoperators**, or more simply **maps**. Linearity means

$$\mathcal{M}\left(\alpha\rho + \sigma\right) = \alpha\mathcal{M}\left(\rho\right) + \mathcal{M}\left(\sigma\right)$$

for any number α and operators ρ, σ. Maps can also be tensored:

$$\left(\mathcal{M}^A \otimes \mathcal{M}^B\right)\left(\rho \otimes \sigma\right) := \mathcal{M}^A\left(\rho\right) \otimes \mathcal{M}^B\left(\sigma\right),$$

with the action on non-product states defined by linearity.

Notation. As with Hilbert-space transformations, a superscript on a map typically refers to the output space. The notation $\mathcal{M}^{A\to B}$ can be used for extra clarity. However, we often omit superscripts to avoid cluttering.

The composition of maps is denoted as $\mathcal{M}_1 \circ \mathcal{M}_2\left(\rho\right) := \mathcal{M}_1\left(\mathcal{M}_2\left(\rho\right)\right)$. Some examples and relevant definitions of maps follow.

- Every state ρ^A defines a map $\mathbb{R} \to A$, defined as $q \mapsto q\rho$ for any $q \in \mathbb{R}$.
- Effects are by definition linear maps $A \to \mathbb{R}$. In terms of effect operators, $\rho^A \mapsto \operatorname{tr} E\rho$ for any $\rho \in A$. An example is the trace operator, dual to the identity operator, $\rho \mapsto \operatorname{tr}\rho = \operatorname{tr} \mathbb{1}\rho$.
- Any linear transformation on Hilbert space, $V : \mathcal{H}^A \to \mathcal{H}^B$, induces a map $\mathcal{V} : A \to B$, defined as $\mathcal{V}(\rho) = V\rho V^\dagger$. The maps induced by projectors, unitaries, isometries, etc., retain their names when taken as superperators. Context will tell whether they should be thought of as acting on Hilbert space or on operators.
- A map is called **positive** if it preserves positive semidefiniteness: $\mathcal{M}(\rho) \geq 0$ for all $\rho \geq 0$.
- A map $\mathcal{M}^A$ is called **completely positive (CP)** if $\mathcal{M}^A \otimes \mathcal{I}^B$ is positive when extended to an arbitrary system B, where $\mathcal{I}$ is the identity map, $\mathcal{I}(\rho) = \rho$.
- A map $\mathcal{T}$ is **trace preserving** if $\operatorname{tr}\mathcal{T}(\rho) = \operatorname{tr}\rho$. Completely positive and trace preserving (**CPTP**) maps are also called **quantum channels**.

Physical transformation of mixed states are described by superoperators. CPTP maps (quantum channels) are the most general transformations that can be performed with unit probability. Isomtetries and unitaries are special examples of quantum channels. More general transformations include

- **partial knowledge** of the transformation performed: if one of a set $\{V_j\}_j$ takes place with probability p_j, the corresponding quantum channel is

$$\mathcal{T}(\rho) = \sum_j p_j V_j \rho V_j^\dagger;$$

- **discarding** a part of a system:

$$\mathcal{T}^A\left(\rho^{AB}\right) = \operatorname{tr}_B\left[U^{AB}\rho^{AB}(U^{AB})^\dagger\right],$$

 where U is a joint unitary acting on A and B;

- **interaction** of the system with an environment,

$$\mathcal{T}^A\left(\rho^{A}\right) = \operatorname{tr}_B\left[U^{AB}\rho^{A}\otimes\sigma^B(U^{AB})^\dagger\right],$$

 where σ is a fixed initial state of the environment.

Non-deterministic transformations, which might take place with less-than-unit probability, are described by CP, *trace non-increasing* maps, also called **quantum operations**[1]. These describe the transformation of a system resulting from a particular outcome of a measurement procedure. A particular example is a projective measurement,

$$\rho \mapsto \Pi\rho\Pi.$$

Other sources renormalise the post-measurement state, $\rho \mapsto \frac{\mathcal{M}(\rho)}{\operatorname{tr}\mathcal{M}(\rho)}$. Instead, as for pure states, here we keep sub-normalised states, with $\operatorname{tr}\mathcal{M}(\rho) \leq 1$ denoting the probability for the transformation to take place.

Pictures. We usually work within the *Schrödinger picture*, where states are transformed but effects stay fixed. Equivalently, one could keep states fixed and transform effects, in the *Heisenberg picture*. Heisenberg and Schrödinger's picture maps are each other's adjoints relative to the Hilbert-Schmidt scalar product: $\operatorname{tr}[E\,\mathcal{M}(\rho)] = \operatorname{tr}\left[\mathcal{M}^\dagger(E)\,\rho\right]$, ensuring that probabilities are independent of the picture used. For a map induced by a Hilbert-space transformation, $\mathcal{M}(\rho) = M\rho M^\dagger$ in the Schrödinger picture, the Heisenberg-picture adjoint is $\mathcal{M}^\dagger(E) = M^\dagger E M$.

A **quantum instrument** describes a general measurement procedure, which also specifies how the system is transformed after the measurement. An instrument is defined as a collection of CP maps that sum up to a CPTP map: $\mathcal{J} = \{\mathcal{M}_j\}_j$ such that $\operatorname{tr}\sum_j \mathcal{M}_j(\rho) = \operatorname{tr}\rho$. The probability to observe outcome j, given a system in state ρ, is

$$P(j|\rho) = \operatorname{tr}\mathcal{M}_j(\rho).$$

We see that each instrument defines a POVM, with the individual effects defined as $e_j(\rho) = \operatorname{tr}\mathcal{M}_j(\rho)$. However, a POVM does not correspond to a unique instrument, as measurment procedures might different in state transformation but

[1] We will often just call quantum operations CP maps, implicitly assuming the trace non-increasing condition, $\operatorname{tr}\mathcal{M}(\rho) \leq \operatorname{tr}\rho$.

have equal outcome probabilities. Indeed, given a CPTP map $\mathcal{T}$, the instruments $\mathcal{J} = \{\mathcal{M}_j\}_j$ and $\mathcal{J}' = \{\mathcal{T} \circ \mathcal{M}_j\}_j$ have the same outcome probabilities for any state, $\operatorname{tr} \mathcal{M}_j(\rho) = \operatorname{tr} \mathcal{T} \circ \mathcal{M}_j(\rho)$.

When more measurements are performed in a sequence, the joint probabilities for successive outcomes i, j are given by

$$P(i,j|\rho) = \operatorname{tr} \mathcal{M}_j \circ \mathcal{M}_i(\rho) \,. \tag{10}$$

2 More on channels and operations

2.1 Kraus representation of quantum opertations

There is a useful way to represent CP maps, wihch enables us to better link transformations on pure and mixed states. The following result holds:

A map $\mathcal{M} : A \to B$ is completely positive if and only if there exists a set of operators $K_k : \mathcal{H}^A \to \mathcal{H}^B$, $k = 1, \ldots, r$, such that

$$\mathcal{M}(\rho) = \sum_{k=1}^{r} K_k \rho K_k^\dagger. \tag{11}$$

The smallest r from which this is possible is called the Kraus rank *(or* rank*, for short) of the map $\mathcal{M}$.*

Given two maps $\mathcal{M}$, $\mathcal{N}$, with respective Kraus operators $\{K_k\}_k$, $\{H_h\}_h$, the composed map $\mathcal{M} \circ \mathcal{N}$ has Kraus operators $\{\tilde{K}_{k,h}\}_{k,h} = \{K_k H_h\}_{k,h}$. This is useful, for example, if we apply the operation on an initially pure state $[\psi]$. The joint probability for obtaining first $\mathcal{N}$, then $\mathcal{M}$ as two sequential operations is given by

$$\begin{aligned}\operatorname{tr} \mathcal{M} \circ \mathcal{N}([\psi]) &= \sum_{kh} \operatorname{tr} K_k H_h |\psi\rangle\langle\psi| H_h^\dagger K_k^\dagger \\ &= \sum_{khj} \langle j| K_k H_h |\psi\rangle\langle\psi| H_h^\dagger K_k^\dagger |j\rangle = \sum_{khj} \left|\langle j| K_k H_h |\psi\rangle\right|^2 ,\end{aligned}$$

where $\{|j\rangle\}_j$ is a basis of the final Hilbert space. Therefore we can interpret the complex numbers $\langle j|K_k H_h|\psi\rangle$ as probability amplitudes corresponding to (unobserved) measurements h, k, j. This means that we can go through the calculation in Hilbert space, instead of the larger (and often more cumbersome) Hilbert-Schmidt space, to obtain probability amplitudes, and only turn to probabilties at the very end, by taking the modulus square of the amplitudes and summing over unobserved outcomes. Note that this interpretation extends the standard notion of probability amplitudes, which is defined only for projective operators.

A CP map is **trace preserving** if its Kraus operators satisfy

$$\sum_k K_k^\dagger K_k = \mathbb{1}. \tag{12}$$

Therefore, an instrument $\mathcal{J} = \{\mathcal{M}_j\}_j$ is equivalently defined as a set of Kraus operators $\{K_{jk}\}_{jk}$ such that

$$\sum_{jk} K_{jk}^\dagger K_{jk} = \mathbb{1},$$

with the probability for outcome j given by

$$P(j|\rho) = \sum_k \operatorname{tr} K_{jk} \rho K_{jk}^\dagger.$$

This expression also gives us a direct link between an instrument and the corresponding POVM. For an instrument with Kraus operators $\{K_{jk}\}_{jk}$, the corresponding POVM elements are $E_j = \sum_k K_{jk}^\dagger K_{jk}$.

2.2 The church of the larger Hilbert space

It is intuitive that taking a part of a closed system, while ignoring the rest, results in an open system. An important result in quantum theory is that the opposite is also true: every open system or process can be seen as the restriction to a part of a larger system.

State purification. For every density matrix $\rho \in A$, there is a Hilbert space $\mathcal{H}^B$ and a pure state $|\psi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ such that

$$\rho^A = \operatorname{tr}_B [\psi]^{AB}.$$

The state $|\psi\rangle$ is called a *purification* of ρ. Each desitiy matrix has multiple distinct purifications. A canonical one can be constructed in the following way: write ρ in its diagonal basis,

$$\rho^A = \sum_j r^j \, [j]^A,$$

with $0 \leq r^j \leq 1$. Take a Hilbert space $\mathcal{H}^B$ with the same dimension of $\mathcal{H}^A$ and an arbitrary basis in it. It is easy to check that

$$|\psi\rangle^{AB} = \sum_j \sqrt{r^j} |j\rangle^A |j\rangle^B$$

is a purification of ρ^A. This construction tells us that an additional Hilbert space $\mathcal{H}^B$ of dimension at most equal to $\mathcal{H}^A$ is sufficient to find a purification of any state. Depending on the state, a smaller dimension of $\mathcal{H}^B$ might work as well. A purification using a larger Hilbert space is also always possible, although it would not be minimal.

Steinspring's dilation. Every CPTP map $\mathcal{T} : A \to B$ can be expressed as the interaction of the system with an environment, originally in a pure state $|\psi_0\rangle$, followed by discarding part of the system.

$$\mathcal{T}(\rho) = \operatorname{tr}_{E'} \left[U^{AE \to BE'} \rho^A \otimes [\psi_0]^E \, (U^{AE \to BE'})^\dagger \right],$$

where U is unitary. Note that A and B can have arbitrary, different dimensions. The only constraint is that the dimension of the initial combined system AE is equal to that of the final system BE'.

A similar construction works for CP maps and more generally for instruments. Physically, it means that every measurement procedure can be implemented by preparing an environment in a pure state, let it interact reversiby with the system of interest, and then performing a projective measurement on the environment. Formally, any CP map $\mathcal{M} : A \to B$ can be represented as

$$\mathcal{M}(\rho) = \mathrm{tr}_{E'} \left[\Pi^{E'} U^{AE \to BE'} \rho^A \otimes [\psi_0]^E \, (U^{AE \to BE'})^\dagger \right],$$

where U is unitary and Π is a projector. Note that Π needs not to be rank one; the limit case where $\Pi = \mathbb{1}$ corresponds to just discarding, rather than measuring, the environment. Often one considers an environment divided in two subsystems, $E' = E_1 \otimes E_2$, one of which is measured and the other discarded, $\Pi^{E'} = \tilde{\Pi}^{E_1} \otimes \mathbb{1}^{E_2}$.

We can connect Stinespring's dilation with the Kraus representation of CP maps. By expanding the trace over E' using a basis $\{|k\rangle\}_k$, Stinespring's dilation reads

$$\begin{aligned}
\mathcal{M}(\rho) &= \sum_k \langle k|^{E'} \left[\Pi^{E'} U^{AE \to BE'} \rho^A \otimes [\psi_0]^E \, (U^{AE \to BE'})^\dagger \right] |k\rangle^{E'} \\
&= \sum_k \langle k|^{E'} \Pi^{E'} \left[U^{AE \to BE'} \rho^A \otimes |\psi_0\rangle\langle\psi_0|^E (U^{AE \to BE'})^\dagger \right] \Pi^{E'} |k\rangle^{E'}.
\end{aligned}$$

We recognise that this is a Kraus representation of $\mathcal{M}$, with operators

$$K_k = \langle k|^{E'} \Pi^{E'} U^{AE \to BE'} |\psi_0\rangle^E.$$

The Kraus decomposition for a CPTP map corresponds to the particular case $\Pi^{E'} = \mathbb{1}^{E'}$, giving

$$\begin{aligned}
\sum_k K_k^\dagger K_k &= \sum_k \langle\psi_0|^E (U^{AE \to BE'})^\dagger |k\rangle\langle k|^{E'} U^{AE \to BE'} |\psi_0\rangle^E \\
&= \langle\psi_0|^E (U^{AE \to BE'})^\dagger U^{AE \to BE'} |\psi_0\rangle^E = \mathbb{1}.
\end{aligned}$$

2.3 State-channel duality

The last ingredient of this short overview is a convenient way to jump from states to transformations and back, which also serves as a useful representation of processes and operations. Stated in short, each transformation from a system A to a system B corresponds to a bipartite state on the joint system $A \otimes B$ and vice versa. There are several ways and different conventions to implement this duality, each with its own pros and cons. We will adhere to the definitions used in most of the literature concerned with indefinite causal order.

Let us look at **pure states** first. Take a linear transformation $L : \mathcal{H}^A \to \mathcal{H}^B$ and fix a basis $\{|j\rangle\}_j \subset \mathcal{H}^A$. L is mapped to the (non-normalised) state $|L\rangle\!\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$, defined as

$$|L\rangle\!\rangle^{AB} := \sum_j |j\rangle^A \otimes L^B |j\rangle.$$

Conversely, a state $|\psi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ defines a linear transformation $L_\psi : \mathcal{H}^A \to \mathcal{H}^B$, through the partial scalar product

$$L_\psi^B |\phi\rangle := \langle\phi^*|^A |\psi\rangle^{AB},$$

where the complex conjugate $\langle\phi^*|^A$ is defined, in the chosen basis, as

$$\langle\phi^*|^A = \sum_j \langle j|^A \langle j|\phi\rangle = \sum_j \langle j|^A \phi^j.$$

We will call $|L\rangle\rangle$ the "Choi vector" of the operator L, although this is not standard terminology. It refers to the "Choi-Jamiołkowski isomorphism", which we will see later.

Looking at the matrix elements of $L = \sum_{ij} L^{ij}|i\rangle\langle j|$, $L^{ij} = \langle i|L|j\rangle$, we see that the corresponding vector is

$$|L\rangle\rangle^{AB} := \sum_{ij} L^{ij}|j\rangle^A|i\rangle^B.$$

In other words, the vector corresponding to $|L\rangle\rangle$ is obtained by aligning the rows of the matrix corresponding to L into a single row. We can also interpret the isomorphism as arising from "flipping" the basis elements of the dual space $(\mathcal{H}^*)^A$, $\langle j|^A \mapsto |j\rangle^A$. Note that this is different from the canonical isomorphism $|\psi\rangle = \langle\psi|^\dagger$, which not only flips the basis elements but also takes the complex conjugates of the components. Note also that the isomorphism $|\psi\rangle \leftrightarrow \langle\psi|$ is basis independent, but not linear, while the correspondence $L \leftrightarrow |L\rangle\rangle$ is basis-dependent, but linear.

To get a more physical interpretation, consider a maximally entangled state on two copies of system A: $|\phi+\rangle := \frac{1}{\sqrt{d^A}}\sum_{j=1}^{d^A}|j\rangle|j\rangle \equiv \frac{1}{\sqrt{d^A}}|\mathbb{1}\rangle\rangle$, where d^A is the dimension of $\mathcal{H}^A$. If we apply the transformation L on one side, leaving the other unchanged, we obtain the Choi vector, up to normalisation:

$$\mathbb{1} \otimes L|\phi+\rangle = \frac{1}{\sqrt{d^A}}|L\rangle\rangle.$$

The correspondence more conventionally known as **Choi-Jamiołkowski (CJ) isomorphism** relates CP maps to bipartite operators. It is defined as a function $\mathfrak{C} : \mathcal{L}(A, B) \mapsto A \otimes B$, which send a CP map $\mathcal{M} : A \to B$ into

$$\mathfrak{C}(\mathcal{M})^{AB} := \sum_{ij} |i\rangle\langle j|^A \otimes \mathcal{M}^B(|i\rangle\langle j|) = \mathcal{I}^A \otimes \mathcal{M}^B([[\mathbb{1}]]), \qquad (13)$$

with

$$[[\mathbb{1}]] = |\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}| = \sum_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j|.$$

The operator $\mathfrak{C}(\mathcal{M}) \in A \otimes B$ defined in this way is called the *Choi operator*, or *Choi matrix*, of the map $\mathcal{M}$. The reverse direction of the isomorphism relates any operator $M \in A \otimes B$ to the map

$$\mathcal{M}^B\left(\rho^A\right) = \mathrm{tr}_A\left[\left(\rho^{A^T} \otimes \mathbb{1}^B\right) M^{AB}\right], \qquad (14)$$

where ^{T} denotes transposition in the chosen basis.

The important property of this isomorphism is that a map $\mathcal{M}$ is completely positive if and only if $\mathfrak{C}(\mathcal{M})$ is positive semidefinite. Furthermore, trace preserving maps are characterised as those whose Choi operator satisfies

$$\mathrm{tr}_B\,\mathfrak{C}(\mathcal{M})^{AB} = \mathbb{1}^A. \qquad (15)$$

Generalised Born rule. To have a glimpse of how the CJ isomorphism can be used, consider a bipartite communication scenario. Let us call Alice the first party, who acts on some fixed initial state ρ and sends the output of her operation to the second party, Bob, through a fixed channel $\mathcal{T}$. Each of them implements a quantum instrument and we want to know the probability for Alice's instrument to yield an outcome corresponding to operation $\mathcal{M}$, and Bob's instruments to yield $\mathcal{N}$. As we have seen earlier, this probability is given by the formula

$$P(\mathcal{M},\mathcal{N}|\mathcal{T},\rho) = \operatorname{tr}\mathcal{N}\circ\mathcal{T}\circ\mathcal{M}\,(\rho)\,. \tag{16}$$

Let us now introduce the operators

$$\begin{aligned} M &= \mathfrak{C}(M)^T, \\ N &= \mathfrak{C}(N)^T, \\ T &= \mathfrak{C}(T). \end{aligned}$$

Note that we transposed the Choi operators corresponding to the experimenter-controlled maps. Let us further denote A_I, A_O the input and output spaces of Alice's operation, and B_I, B_O input and output for Bob. With a bit of work, one can show that Eq. (16) is equivalent to

$$\begin{aligned} P(\mathcal{M},\mathcal{N}|\mathcal{T},\rho) &= \operatorname{tr}\left[\left(M^{A_IA_O}\otimes N^{B_IB_O}\right)W^{A_IA_OB_IB_O}\right], \\ W^{A_IA_OB_IB_O} &= \rho^{A_I}\otimes T^{A_OB_I}\otimes \mathbb{1}^{B_O}. \end{aligned} \tag{17}$$

We see that the probability formula (17) is formally analogous to the Born rule (8), where the role of the POVM element E_j is replaces by the (transpose) Choi operators $M^{A_IA_O}$, $N^{B_IB_O}$ of the local operation, while the role of the density matrix ρ is replaced by the operator W, which we call **process operator**, or **process matrix**.

The advantage of expression (17) is that Alice and Bob's operations appear in a symmetric position, with their causal relation encoded in the process matrix W. A scenario where Bob acted before Alice would be described in the same way, except for the different process matrix

$$W^{A_IA_OB_IB_O} = T^{B_OA_I}\otimes \mathbb{1}^{A_O}\otimes\rho^{B_I}. \tag{18}$$

The "generalised Born rule" (17) can be seen as the starting point for the study of quantum causal structures.