

Quantum Information Theory

Lectures 17–20

Quantum dynamics. III–VI. Superoperators and completely positive maps.

2002 November 5; modified 2005 September 30 and 2009 February 3 and 2014 October 3

1 Measurement models and quantum operations

Measurements on quantum systems are made by coupling the system to be measured to a measuring apparatus. The measurement is completed by observing the apparatus and inferring some property of the system.

To formalize this, consider a primary quantum system Q , initially in state ρ , which is brought into contact with an *ancilla* A (which can also be thought of as an *environment* or a *measuring apparatus*), initially in state $\sigma = \sum_l \lambda_l |e_l\rangle\langle e_l|$, where the states $|e_k\rangle$ are the eigenstates of σ . The two systems interact for a time, the interaction described by a unitary operator U acting on the joint system QA . Then the ancilla is subjected to a von Neumann measurement, described by orthogonal projectors

$$P_\alpha = \sum_j |f_{\alpha j}\rangle\langle f_{\alpha j}|, \quad (1)$$

where the states $|f_{\alpha j}\rangle$ make up an orthonormal basis for the ancilla, satisfying the completeness relation

$$I_A = \sum_{\alpha,j} |f_{\alpha j}\rangle\langle f_{\alpha j}| = \sum_{\alpha} P_\alpha. \quad (2)$$

The Greek index α is used to label the outcome of the measurement. If the result of the measurement is α , i.e., the ancilla is observed to be in the subspace S_α , the unnormalized state of the system after the measurement is obtained, from the standard rules for a von Neumann measurement, by projecting the joint QA state into the subspace S_α and then tracing out the ancilla,

$$\text{tr}_A(P_\alpha U \rho \otimes \sigma U^\dagger P_\alpha) = \text{tr}_A(P_\alpha U \rho \otimes \sigma U^\dagger) \equiv \mathcal{A}_\alpha(\rho), \quad (3)$$

where \mathcal{A}_α is a linear map on system density operators. Any linear map on operators is called a superoperator, and the particular kind of superoperator defined by Eq. (3) is called a *quantum operation*. This method of defining a set of quantum operation in terms of interaction with an initially uncorrelated ancilla, followed by measurement on the ancilla, is called a *measurement model*.

Inserting the forms of the ancilla projector P_α and the initial ancilla state σ leads to a form for the operation \mathcal{A}_α that only involves operators on the primary system Q :

$$\mathcal{A}_\alpha(\rho) = \sum_{j,k} \sqrt{\lambda_k} \langle f_{\alpha j} | U | e_k \rangle \rho \langle e_k | U^\dagger | f_{\alpha j} \rangle \sqrt{\lambda_k} = \sum_{j,k} A_{\alpha j k} \rho A_{\alpha j k}^\dagger. \quad (4)$$

The system operators

$$A_{\alpha j k} \equiv \sqrt{\lambda_k} \langle f_{\alpha j} | U | e_k \rangle \quad (5)$$

are said to provide a *Kraus decomposition* (or *operator-sum representation*) of the operation \mathcal{A}_α and thus are called *Kraus operators* or *operation elements*. The Kraus operators (5) are defined by the relative states in a decomposition of $U|\psi\rangle \otimes |e_k\rangle$ relative to the ancilla basis $|f_{\alpha j}\rangle$:

$$U|\psi\rangle \otimes |e_k\rangle = \sum_{\alpha,j} \langle f_{\alpha j} | U | e_k \rangle |\psi\rangle \otimes |f_{\alpha j}\rangle = \sum_{\alpha,j} \frac{1}{\sqrt{\lambda_k}} A_{\alpha j k} |\psi\rangle \otimes |f_{\alpha j}\rangle . \quad (6)$$

The Kraus operators satisfy a completeness relation:

$$\sum_{\alpha,j,k} A_{\alpha j k}^\dagger A_{\alpha j k} = \sum_{\alpha,j,k} \lambda_k \langle e_k | U^\dagger | f_{\alpha j} \rangle \langle f_{\alpha j} | U | e_k \rangle = \text{tr}_A(U^\dagger U \sigma) = \text{tr}_A(I \otimes \sigma) = I . \quad (7)$$

The probability to obtain result α in the measurement on the ancilla is, from the standard rules for an OP measurement (measurement described by orthogonal projectors),

$$p_\alpha = \text{tr}(P_\alpha U \rho \otimes \sigma U^\dagger) = \text{tr}(\mathcal{A}_\alpha(\rho)) = \text{tr}\left(\rho \sum_{j,k} A_{\alpha j k}^\dagger A_{\alpha j k}\right) = \text{tr}(\rho E_\alpha) . \quad (8)$$

The operators

$$E_\alpha \equiv \sum_{j,k} A_{\alpha j k}^\dagger A_{\alpha j k} = \sum_{j,k} \lambda_k \langle e_k | U^\dagger | f_{\alpha j} \rangle \langle f_{\alpha j} | U | e_k \rangle = \text{tr}_A(U^\dagger P_\alpha U \sigma) \quad (9)$$

are clearly positive and satisfy a completeness relation because of Eq. (7). Any measurement model thus gives rise to a POVM that describes the measurement statistics. The *normalized* post-measurement system state, conditioned on result α , is

$$\rho_\alpha = \frac{\mathcal{A}_\alpha(\rho)}{\text{tr}(\mathcal{A}_\alpha(\rho))} = \frac{\mathcal{A}_\alpha(\rho)}{p_\alpha} = \frac{1}{p_\alpha} \sum_{j,k} A_{\alpha j k} \rho A_{\alpha j k}^\dagger . \quad (10)$$

If we don't know the result of the measurement on the ancilla, the post-measurement state is obtained by averaging over the possible measurement results:

$$\rho' = \sum_{\alpha} p_\alpha \rho_\alpha = \sum_{\alpha} \mathcal{A}_\alpha(\rho) = \text{tr}_A(U \rho \otimes \sigma U^\dagger) = \sum_{\alpha,j,k} A_{\alpha j k} \rho A_{\alpha j k}^\dagger \equiv \mathcal{A}(\rho) . \quad (11)$$

This is the dynamics we would get for the system in the absence of any measurement on the ancilla or, formally, for a completely uninformative measurement on the ancilla, i.e., one that has a single result with $P_\alpha = I_A$.

A primary quantum system that is exposed to an initially uncorrelated environment *always* has dynamics described by a quantum operation, whether or not a measurement is made on the environment. If we do make a measurement on the environment, the system state after the dynamics is conditioned on the result of the measurement through the projection operator $P_\alpha \neq I_A$ in Eq. (3); the corresponding quantum operation \mathcal{A}_α is said to be *trace decreasing* because the trace of the output is generally smaller than the trace of the input, the reduction factor being the probability for the measurement result α . If we do not make a measurement on the environment, we have an open-system dynamics described by the operation \mathcal{A} of Eq. (11), which is said to be *trace preserving* because the trace of the output is the same as the trace of the input. Formally \mathcal{A} is the quantum operation for a completely uninformative measurement on the ancilla, which has one outcome $P_\alpha = I_A$. Moreover, we can think of any trace-preserving open-system dynamics as coming

from an environment that “monitors” the system, even though we acquire none of the monitored information; this monitoring destroys quantum coherence in Q , a process called *decoherence*.

The way we treat the measurement result α , averaging over it if we don’t know the result, suggests how we should think about the indices j and k that are summed over in the Kraus decomposition of \mathcal{A}_α . They represent information that is potentially available to us, but that we don’t have. Indeed, we can always imagine that there is a more capable agent than ourselves, who has two kinds of fine-grained information that we don’t have: before the interaction between system and ancilla, this agent knows which eigenstate $|e_k\rangle$ of σ applies to the ancilla, but only reports to us that the ancilla occupies one of the eigenstates $|e_k\rangle$ with probabilities given by the eigenvalues of σ ; after the interaction, this agent also knows the result αj of a fine-grained measurement in the basis $|f_{\alpha j}\rangle$, but only reports to us the subspace S_α corresponding to the result α . After the measurement, the agent attributes a post-measurement state

$$\rho_{\alpha jk} = \frac{A_{\alpha jk} \rho A_{\alpha jk}^\dagger}{p_{\alpha jk}} = \mathcal{A}_{\alpha jk}(\rho) \quad (12)$$

to Q , where

$$p_{\alpha jk} = \text{tr}(\mathcal{A}_{\alpha jk}(\rho)) = \text{tr}(\rho A_{\alpha jk}^\dagger A_{\alpha jk}) \quad (13)$$

is the probability associated with initial state k and result αj . Not knowing the values of j and k , but knowing α , we assign a post-measurement state that averages over these unknown values, given α :

$$\rho_\alpha = \sum_{j,k} p_{jk|\alpha} \rho_{\alpha jk} . \quad (14)$$

Using Bayes’s theorem, we can write $p_{jk|\alpha} = p_{\alpha jk}/p_\alpha$, thus obtaining the result of Eq. (10),

$$\rho_\alpha = \frac{1}{p_\alpha} \sum_{j,k} p_{\alpha jk} \rho_{\alpha jk} = \frac{1}{p_\alpha} \sum_{j,k} A_{\alpha jk} \rho A_{\alpha jk}^\dagger = \frac{1}{p_\alpha} \mathcal{A}_\alpha(\rho) , \quad (15)$$

where

$$p_\alpha = \sum_{j,k} p_{\alpha jk} = \text{tr}\left(\rho \sum_{j,k} A_{\alpha jk}^\dagger A_{\alpha jk}\right) \quad (16)$$

is the probability for result α . Notice that even though we have to worry about conditional and joint probabilities when relating the various kinds of post-measurement states, everything about the probabilities disappears from the operations: *to construct the operation corresponding to the coarse-grained data, we simply sum the fine-grained operations over the fine-grained data we don’t have:*

$$\mathcal{A}_\alpha = \sum_{j,k} \mathcal{A}_{\alpha jk} . \quad (17)$$

Although there is a physical difference between the two kinds of potentially available fine-grained information symbolized by the indices j and k , there is no mathematical difference between them, so we combine them into a single index in the following.

Von Neumann measurements. A *von Neumann* measurement on the system Q is a measurement whose Kraus operators are a set of orthogonal projection operators Π_α . Since orthogonal projectors commute, they can be simultaneously diagonalized in an orthonormal system basis $|e_{\alpha j}\rangle$, i.e.,

$$\Pi_\alpha = \sum_j |e_{\alpha j}\rangle \langle e_{\alpha j}| . \quad (18)$$

The POVM elements are the same projection operators, i.e., $E_\alpha = \Pi_\alpha^\dagger \Pi_\alpha = \Pi_\alpha$ (this is an OP measurement). The probability for outcome α is $p_\alpha = \text{tr}(\rho \Pi_\alpha)$, and the post-measurement state, given outcome α , is

$$\rho_\alpha = \frac{\Pi_\alpha \rho \Pi_\alpha}{p_\alpha} = \frac{1}{p_\alpha} \sum_{j,k} |e_{\alpha j}\rangle \langle e_{\alpha j} | \rho | e_{\alpha k}\rangle \langle e_{\alpha k} | . \quad (19)$$

If we don't know the result of the measurement, the post-measurement state becomes

$$\rho' = \sum_\alpha p_\alpha \rho_\alpha = \sum_\alpha \sum_{j,k} |e_{\alpha j}\rangle \langle e_{\alpha j} | \rho | e_{\alpha k}\rangle \langle e_{\alpha k} | . \quad (20)$$

This measurement destroys the coherence between subspaces corresponding to different outcomes, but retains the coherence within each outcome subspace.

Suppose instead that the Kraus operators corresponding to outcome α are the one-dimensional projectors $\Pi_{\alpha j} = |e_{\alpha j}\rangle \langle e_{\alpha j}|$. These give rise to the same OP POVM elements, $E_\alpha = \Pi_\alpha$, since

$$\sum_j \Pi_{\alpha j}^\dagger \Pi_{\alpha j} = \sum_j \Pi_{\alpha j} = \Pi_\alpha = E_\alpha . \quad (21)$$

This means that the measurement statistics for these fine-grained Kraus operators are the same as for the coarse-grained Kraus operators Π_α , but the post-measurement states are quite different. For the one-dimensional projectors, the post-measurement state corresponding to outcome α is

$$\rho_\alpha = \frac{1}{p_\alpha} \sum_j \Pi_{\alpha j} \rho \Pi_{\alpha j} = \frac{1}{p_\alpha} \sum_j |e_j\rangle \langle e_j | \rho | e_j\rangle \langle e_j | . \quad (22)$$

This measurement removes all the coherence in the $|e_j\rangle$ basis, leaving the density operator diagonal in this basis.

There are two lessons here. The first is that generally a POVM element corresponds to many different quantum operations: *measurement statistics do not specify the post-measurement state*. The second is that the finer-grained the Kraus operators underneath a given a POVM element, the more decoherent the measurement is.

The above considerations show that any measurement model can be summarized by a set of system operators $A_{\alpha j}$ that satisfy the completeness relation (7). We would like to know the converse, i.e., that any set of superoperators \mathcal{A}_α with Kraus operators that satisfy a completeness relation can be realized by a measurement model. The converse is stated formally as the Kraus representation theorem.

Kraus representation theorem. Given a set of superoperators with Kraus decompositions,

$$\mathcal{A}_\alpha(\rho) = \sum_j A_{\alpha j} \rho A_{\alpha j}^\dagger , \quad (23)$$

where the Kraus operators satisfy the completeness relation

$$\sum_{\alpha,j} A_{\alpha j}^\dagger A_{\alpha j} = I , \quad (24)$$

there exists an ancilla A with (pure) initial state $|e_0\rangle \langle e_0|$, a joint unitary operator U on QA , and orthogonal projectors P_α on A such that

$$\mathcal{A}_\alpha(\rho) = \text{tr}_A(P_\alpha U \rho \otimes |e_0\rangle \langle e_0| U^\dagger) . \quad (25)$$

Proof: The proof is simply a matter of reversing the steps that led from a measurement model to Kraus operators, making sure that the operator U is unitary.

Pick an ancilla whose Hilbert space has as many dimensions as the number of values of the pair αj . Notice that in constructing the measurement model, we need one ancilla dimension for each Kraus operator. Take *any* ancilla pure state $|e_0\rangle\langle e_0|$ and *any* ancilla orthonormal basis $|f_{\alpha j}\rangle$. Partially define a joint QA operator U by

$$U|\psi\rangle \otimes |e_0\rangle = \sum_{\alpha,j} A_{\alpha j} |\psi\rangle \otimes |f_{\alpha j}\rangle \quad \Longleftrightarrow \quad \langle f_{\alpha j} | U | e_0 \rangle = A_{\alpha j} . \quad (26)$$

The operator U is defined on the D -dimensional subspace $\mathcal{H}_Q \otimes R_0$, where R_0 is the one-dimensional ancilla subspace spanned by $|e_0\rangle$. This partial definition preserves inner products,

$$(\langle \phi | \otimes \langle e_0 | U^\dagger)(U |\psi\rangle \otimes |e_0\rangle) = \sum_{\substack{\alpha,j \\ \beta,k}} \langle \phi | A_{\beta k}^\dagger A_{\alpha j} |\psi\rangle \underbrace{\langle f_{\beta k} | f_{\alpha j} \rangle}_{= \delta_{\alpha\beta} \delta_{jk}} = \left\langle \phi \left| \underbrace{\sum_{\alpha,j} A_{\alpha j}^\dagger A_{\alpha j}}_{= I} \right| \psi \right\rangle = \langle \phi | \psi \rangle , \quad (27)$$

which means that U maps the subspace $\mathcal{H}_Q \otimes R_0$ unitarily to a D -dimensional subspace S_0 of $\mathcal{H}_Q \otimes \mathcal{H}_A$. We can extend U to be a unitary operator on all of $\mathcal{H}_Q \otimes \mathcal{H}_A$ by defining it to map the subspace $\mathcal{H}_Q \otimes R_\perp$, where R_\perp is the ancilla subspace orthogonal to R_0 , unitarily to the subspace orthogonal to S_0 .

With this definition of U , we have

$$\mathcal{A}_\alpha(\rho) = \sum_j A_{\alpha j} \rho A_{\alpha j}^\dagger = \sum_j \langle f_{\alpha j} | U | e_0 \rangle \rho \langle e_0 | U^\dagger | f_{\alpha j} \rangle = \text{tr}_A \left(\left(\sum_j |f_{\alpha j}\rangle \langle f_{\alpha j}| \right) U \rho \otimes |e_0\rangle \langle e_0| U^\dagger \right) . \quad (28)$$

Defining a complete set of ancilla orthogonal projectors,

$$P_\alpha = \sum_j |f_{\alpha j}\rangle \langle f_{\alpha j}| , \quad (29)$$

we have

$$\mathcal{A}_\alpha(\rho) = \text{tr}_A \left(P_\alpha U \rho \otimes |e_0\rangle \langle e_0| U^\dagger \right) , \quad (30)$$

as required. **QED**

Comments on the Kraus representation theorem.

1. The Kraus representation theorem guarantees that any single superoperator \mathcal{A} with a Kraus decomposition,

$$\mathcal{A}(\rho) = \sum_j A_j \rho A_j^\dagger , \quad (31)$$

where the Kraus operators satisfy

$$E \equiv \sum_j A_j^\dagger A_j \leq I , \quad (32)$$

can be realized by a measurement model. The reason is that we can always consider two superoperators defined by $\mathcal{A}_1 = \mathcal{A}$ and $\mathcal{A}_2(\rho) = \sqrt{I - E} \rho \sqrt{I - E}$. Together these two superoperators satisfy the completeness relation of the representation theorem, and thus they can be realized by a measurement model. In constructing this measurement model, the number of ancilla dimensions we need is given by the number of Kraus operators in the Kraus

decomposition of \mathcal{A} plus one more dimension to accommodate the one Kraus operator in \mathcal{A}_2 (if \mathcal{A} is trace-preserving, we don't need \mathcal{A}_2 , so we don't need that one extra dimension). We conclude that we can characterize a quantum operation as any superoperator that has a Kraus decomposition satisfying Eq. (32).

2. The Kraus representation theorem says that any quantum operation can be realized by a measurement model in which the ancilla's initial state is pure. It is clear why we need only consider initial pure states for the ancilla: if we find a measurement model with the ancilla initially in a mixed state σ , we can always purify σ into an even larger ancilla. More important is that the theorem only holds for initial ancilla pure states: it is not guaranteed that a quantum operation has a measurement model with an initial mixed-state ancilla. Indeed, the question of whether a quantum operation has a nontrivial mixed-state model is exactly the question of which operations are extreme points in the convex set of operations, as is shown in Appendix C.
3. Now that we have generalized from von Neumann measurements to quantum operations, we should ask if we would get some even more general kind of dynamics if we allowed measurement models in which the measurement on the ancilla was described by operations instead of orthogonal projectors. The representation theorem allows us to answer this question in the negative, because the quantum operations on the ancilla could always be modeled by projection operators on a yet larger ancilla.

What we have shown in this section is that a quantum operation defined by a measurement model, as in Eq. (3), can equivalently be specified by a set of Kraus operators whose corresponding POVM element satisfies Eq. (32). In the next section we develop yet a third way of characterizing a quantum operation.

2 Completely positive maps

We now want to formulate an abstract set of properties that characterize any dynamics allowed by quantum mechanics and then show that such a dynamics must be described in terms of a quantum operation. This will give us an abstract characterization of quantum operations, akin to our previous characterization of measurement statistics in terms of POVMs.

What we want to describe is a general quantum dynamics that has as input a quantum system Q in input state ρ and that can produce one or more outcomes, which we will label by an index α . Given input state ρ , outcome α occurs with probability $p_{\alpha|\rho}$, and the state of the system Q , conditioned on outcome α , is the density operator ρ_α . Let us define a map, \mathcal{A}_α , not yet assumed to be linear, that takes in the input state ρ and outputs a positive operator that encodes both $p_{\alpha|\rho}$ and ρ_α in the way we are familiar with, i.e.,

$$p_{\alpha|\rho} = \text{tr}(\mathcal{A}_\alpha(\rho)) \quad \text{and} \quad \rho_\alpha = \mathcal{A}_\alpha(\rho)/p_{\alpha|\rho} . \quad (33)$$

This kind of map is *trace decreasing*, because the outcome probabilities are between 0 and 1, inclusive, and are generally less than 1. If there is only one outcome, which therefore occurs with probability one, we omit the index α and write the output state as

$$\rho' = \mathcal{A}(\rho) . \quad (34)$$

This kind of map is *trace preserving*.

We will now argue that \mathcal{A}_α should be *convex linear*, i.e.,

$$\mathcal{A}_\alpha(\lambda\rho_1 + (1-\lambda)\sigma) = \lambda\mathcal{A}_\alpha(\rho) + (1-\lambda)\mathcal{A}_\alpha(\sigma), \quad 0 \leq \lambda \leq 1. \quad (35)$$

The argument proceeds as follows. Suppose we know that the input state is either ρ_1 , occurring with probability p_1 , or ρ_2 , occurring with probability p_2 . Thus the input state is the mixture $\rho = p_1\rho_1 + p_2\rho_2$. The probability for outcome α can be written in two ways, the first being part of the definition of our map,

$$p_{\alpha|\rho} = \text{tr}(\mathcal{A}_\alpha(\rho)) = \text{tr}(\mathcal{A}_\alpha(p_1\rho_1 + p_2\rho_2)), \quad (36)$$

and the second coming from the rules of probability theory,

$$p_{\alpha|\rho} = p_{\alpha|\rho_1}p_1 + p_{\alpha|\rho_2}p_2 = p_1\text{tr}(\mathcal{A}_\alpha(\rho_1)) + p_2\text{tr}(\mathcal{A}_\alpha(\rho_2)) = \text{tr}(p_1\mathcal{A}_\alpha(\rho_1) + p_2\mathcal{A}_\alpha(\rho_2)). \quad (37)$$

Thus, by using elementary probability theory, we get that the trace of both sides of Eq. (35) should be the same.

We get to the stronger conclusion of convex linearity by a similar argument that writes the output state ρ_α in two ways. The first of these ways is just the statement that, not knowing whether the input state is ρ_1 or ρ_2 , but knowing only the probabilities for these two inputs, the input state is $\rho = p_1\rho_1 + p_2\rho_2$, so we apply the dynamics to ρ :

$$\rho_\alpha = \frac{\mathcal{A}_\alpha(\rho)}{p_{\alpha|\rho}} = \frac{\mathcal{A}_\alpha(p_1\rho_1 + p_2\rho_2)}{p_{\alpha|\rho}}. \quad (38)$$

This way can be summarized as mixing followed by dynamics.

The second way comes from arguing that we should get the same result from dynamics followed by mixing. We should be able to get to the same ρ_α by applying the dynamics separately to the two possible inputs, ρ_1 and ρ_2 , yielding output states $\text{tr}(\mathcal{A}_\alpha(\rho_1))/p_{\alpha|\rho_1}$ and $\text{tr}(\mathcal{A}_\alpha(\rho_2))/p_{\alpha|\rho_2}$, and then mixing these two output states to reflect our lack of knowledge of which applies at the output. But what probabilities should we use for the mixing at the output? Not the original probabilities, p_1 and p_2 , because once we know the outcome α , we know something more about the inputs, and we revise our probabilities for the two inputs to reflect this knowledge. What we should do is to mix the two output states with the updated probabilities for the two inputs, i.e., the probabilities $p_{\rho_1|\alpha}$ and $p_{\rho_2|\alpha}$. This gives the second way of writing ρ_α :

$$\rho_\alpha = p_{\rho_1|\alpha} \frac{\mathcal{A}_\alpha(\rho_1)}{p_{\alpha|\rho_1}} + p_{\rho_2|\alpha} \frac{\mathcal{A}_\alpha(\rho_2)}{p_{\alpha|\rho_2}}. \quad (39)$$

The updated probabilities come from Bayes's theorem:

$$p_{\rho_1|\alpha} = \frac{p_{\alpha|\rho_1}p_1}{p_{\alpha|\rho}} \quad \text{and} \quad p_{\rho_2|\alpha} = \frac{p_{\alpha|\rho_2}p_2}{p_{\alpha|\rho}}. \quad (40)$$

Plugging these updated probabilities into Eq. (39) gives

$$\rho_\alpha = \frac{1}{p_{\alpha|\rho}} (p_1\mathcal{A}_\alpha(\rho_1) + p_2\mathcal{A}_\alpha(\rho_2)), \quad (41)$$

and equating the right-hand sides of Eqs. (38) and (39) gives us the desired convex linearity.

It turns out that a convex-linear map from trace-one positive operators to positive operators can always be extended to a linear map on all operators, i.e., a *superoperator*. Though this extension

is trivial, it is annoyingly tedious to show that it works, so the demonstration is relegated to Appendix A.

We now summarize the properties that we have so far found desirable for a map to describe a general quantum dynamics. In proceeding, the index α just gets in the way, so we omit it, simply calling the map \mathcal{A} .

- *Condition 1.* \mathcal{A} is a map from trace-one positive operators (density operators) to positive operators.
- *Condition 2.* \mathcal{A} is *trace decreasing*, i.e., $\text{tr}(\mathcal{A}(\rho)) \leq 1$ for all density operators ρ . Trace-preserving dynamics is the special case where $\text{tr}(\mathcal{A}(\rho)) = 1$ for all density operators ρ .
- *Condition 3.* \mathcal{A} is convex linear, i.e.,

$$\mathcal{A}(\lambda\rho_1 + (1 - \lambda)\sigma) = \lambda\mathcal{A}(\rho) + (1 - \lambda)\mathcal{A}(\sigma), \quad 0 \leq \lambda \leq 1, \quad (42)$$

and thus can be extended to be a superoperator, i.e., a linear map on operators.

Conditions 1 and 2 are immediate consequences of the way we set up our description of quantum dynamics. Condition 3 is less firm, since we had to argue for it based on ideas of how the dynamics handles a mixture of density operators, but the argument certainly seemed reasonable.

In Appendix B we develop the theory and a useful notation for superoperators. We use the resulting notation and terminology freely in the following, so it would be a good idea to be mastering Appendix B as you proceed through the discussion. Using the terminology formulated in Appendix B, we can summarize conditions 1–3 by the concise statement that \mathcal{A} is a positive, trace-decreasing superoperator.

It is easy to verify that any quantum operation, defined by a Kraus decomposition (31) satisfying (32), satisfies conditions 1–3. It is tempting to think that these three conditions are sufficient to characterize a quantum operation, but they are not. The transposition superoperator \mathcal{T} , which takes a density operator $\rho = \sum_{j,k} \rho_{jk} |e_j\rangle\langle e_k|$ to its *transpose* relative to the orthonormal basis $|e_j\rangle$, i.e.,

$$\mathcal{T}(\rho) = \sum_{j,k} \rho_{kj} |e_j\rangle\langle e_k| \equiv \rho^T \quad \Longleftrightarrow \quad \mathcal{T} = \sum_{j,k} |e_j\rangle\langle e_k| \odot |e_j\rangle\langle e_k|, \quad (43)$$

clearly satisfies conditions 1–3, but as we show below, it is not a quantum operation because it cannot be written in terms of any Kraus decomposition. In Eq. (43) we introduce the symbol \odot , called “o-dot,” so that we can write a superoperator in an abstract form that does not include the operator that it acts on; the \odot can be regarded as a place-holder, to be replaced by the operator on which \mathcal{S} acts.

The transposition superoperator is important in physics because, since ρ is Hermitian, it is the same as complex conjugating the matrix elements of ρ in the basis $|e_j\rangle$. If, for example, in ordinary quantum mechanics, the standard basis is taken to be the position eigenbasis, then the transposition superoperator performs a time reversal.

We thus need some additional condition on a map \mathcal{A} to be a suitable quantum dynamics. The additional condition can be motivated physically in the following way. Suppose that R is a “reference system” that, though it does not itself take part in the dynamics, cannot be neglected because the initial state ρ of Q is the marginal density operator of a joint state ρ_{RQ} . This certainly being one of the ways to get a density operator for Q , we can’t avoid thinking about this situation. We want the map $\mathcal{I}_R \otimes \mathcal{A}$, where \mathcal{I}_R is the unit superoperator acting on R , to be a suitable quantum dynamics, which means that it must take joint states ρ_{RQ} to positive operators, which

can be normalized to be output density operators. This requirement holds trivially for a quantum operation, because the operators $I_R \otimes A_j$ are a Kraus decomposition for the extended operation $\mathcal{I}_R \otimes \mathcal{A}$, i.e.,

$$(\mathcal{I}_R \otimes \mathcal{A})(\rho_{RQ}) = \sum_j (I_R \otimes A_j) \rho_{RQ} (I_R \otimes A_j^\dagger) \geq 0, \quad (44)$$

Thus we can add an additional requirement, which strengthens condition 1, to our list of conditions for a map to describe a quantum dynamics:

- *Condition 4.* $(\mathcal{I}_R \otimes \mathcal{A})(\rho_{RQ}) \geq 0$ for all joint density operators ρ_{RQ} of Q and reference systems R of arbitrary dimension. Such a map is said to be *completely positive*.

Our objective now is to show that any map that satisfies conditions 1–4 is a quantum operation. Before turning to that task, however, let's first see what goes wrong with an apparently satisfactory map like the transposition superoperator. We want to consider the map $\mathcal{I}_R \otimes \mathcal{T}$, which is called the *partial transposition superoperator*, because it transposes matrix elements in system Q only. Let's suppose R has the same dimension as Q , and let's apply the partial transposition superoperator to an (unnormalized) maximally entangled state

$$|\Psi\rangle \equiv \sum_j |f_j\rangle \otimes |e_j\rangle = \sum_j |f_j, e_j\rangle, \quad (45)$$

where the vectors $|f_j\rangle$ comprise an orthonormal basis for R , and the vectors $|e_j\rangle$ make up an orthonormal basis for Q . Partially transposing $|\Psi\rangle$ gives

$$\begin{aligned} \mathcal{I}_R \otimes \mathcal{T}(|\Psi\rangle\langle\Psi|) &= \mathcal{I}_R \otimes \mathcal{T}\left(\sum_{j,k} |f_j\rangle\langle f_k| \otimes |e_j\rangle\langle e_k|\right) \\ &= \sum_{j,k} \mathcal{I}_R(|f_j\rangle\langle f_k|) \otimes \mathcal{T}(|e_j\rangle\langle e_k|) \\ &= \sum_{j,k} |f_j\rangle\langle f_k| \otimes |e_k\rangle\langle e_j| \\ &= \sum_{j,k} |f_j, e_k\rangle\langle f_k, e_j|. \end{aligned} \quad (46)$$

It is easy to see that the normalized eigenvectors of this operator are the states $|f_j, e_j\rangle$ for $j = 1, \dots, D$ and the states $(|f_j, e_k\rangle \pm |f_k, e_j\rangle)/\sqrt{2}$, for all pairs of indices. The states $|f_j, e_j\rangle$ have eigenvalue 1, but the states $(|f_j, e_k\rangle \pm |f_k, e_j\rangle)/\sqrt{2}$ have eigenvalue ± 1 , showing that the operator (46) is not a positive operator (it is, in fact, a unitary operator). This shows that \mathcal{T} cannot be given a Kraus decomposition.

This example suggests that the problem with superoperators that are positive, but not completely positive is that, when extended to R , they don't map all entangled states to positive operators, as we would like them to. Indeed, as we now show, the general requirements for complete positivity follow from considering only one kind of reference system, one whose dimension is the same as the dimension of Q , and only one kind of joint state, the maximally entangled state (45). All we need to consider is how the extended map $\mathcal{I}_R \otimes \mathcal{A}$ acts on $|\Psi\rangle\langle\Psi|$, i.e., the following operator on RQ :

$$\mathcal{I}_R \otimes \mathcal{A}(|\Psi\rangle\langle\Psi|). \quad (47)$$

The demonstration is really just a rewrite for general superoperators of what we have just done with the transposition superoperator, but to make it more formal, we introduce some new terminology, which is fleshed out further in Appendix B.

One way of thinking about Eq. (47) is that it is a map that takes in a superoperator \mathcal{A} on Q and spits out an operator on RQ . We have already had some experience with a similar map, the VEC map, which takes in an operator A on Q and spits out the following vector on RQ ,

$$|\Phi_A\rangle \equiv I_R \otimes A|\Psi\rangle = \sum_j |f_j\rangle \otimes A|e_j\rangle = \sum_{j,k} |f_j\rangle \otimes |e_k\rangle \underbrace{\langle e_k|A|e_j\rangle}_{=A_{kj}}. \quad (48)$$

The VEC map is a one-to-one, linear map from the operators on Q to \mathcal{H}_{RQ} ; we recover A from $|\Phi_A\rangle$ via

$$\langle f_j, e_k | \Phi_A \rangle = \langle e_k | A | e_j \rangle = A_{kj}. \quad (49)$$

It is easy to see that VEC preserves inner products,

$$\langle \Phi_A | \Phi_B \rangle = \text{tr}(A^\dagger B) = (A|B). \quad (50)$$

Here we introduce operator “bras” and “kets” with rounded brackets, $(A| = A^\dagger$ and $|A) = A$, so that we can use Dirac-style notation for operators and their inner products when it is convenient to do so. We can summarize the VEC map in the following way:

$$A = |A) \leftrightarrow |\Phi_A\rangle = I_R \otimes A|\Psi\rangle \quad \text{and} \quad A^\dagger = (A| \leftrightarrow \langle \Phi_A| = \langle \Psi| I_R \otimes A^\dagger. \quad (51)$$

One further aspect of VEC deserves mention. Given a density operator ρ for Q , applying VEC to $\sqrt{\rho}$,

$$|\Phi_{\sqrt{\rho}}\rangle = I_R \otimes \sqrt{\rho}|\Psi\rangle = \sum_j |f_j\rangle \otimes \sqrt{\rho}|e_j\rangle, \quad (52)$$

generates a purification of ρ , i.e.,

$$\text{tr}_R(|\Phi_{\sqrt{\rho}}\rangle\langle\Phi_{\sqrt{\rho}}|) = \rho. \quad (53)$$

The analogous OP map (47) is a one-to-one, linear map from superoperators on Q to operators on RQ . It takes a superoperator \mathcal{S} on Q to the operator

$$\mathcal{I}_R \otimes \mathcal{S}(|\Psi\rangle\langle\Psi|) = \sum_{j,k} |f_j\rangle\langle f_k| \otimes \mathcal{S}(|e_j\rangle\langle e_k|), \quad (54)$$

and we recover \mathcal{S} via its “matrix elements”

$$\langle f_j, e_l | \mathcal{I}_R \otimes \mathcal{S}(|\Psi\rangle\langle\Psi|) | f_k, e_m \rangle = \langle e_l | \mathcal{S}(|e_j\rangle\langle e_k|) | e_m \rangle \equiv \mathcal{S}_{lj,mk}. \quad (55)$$

These matrix elements clearly specify \mathcal{S} . To see explicitly how, we write \mathcal{S} abstractly in the following way:

$$\mathcal{S} = \sum_{l,j,m,k} \mathcal{S}_{lj,mk} \underbrace{|e_l\rangle\langle e_j|}_{=\tau_\alpha} \odot \underbrace{|e_k\rangle\langle e_m|}_{=\tau_\beta^\dagger} = \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta} \tau_\alpha \odot \tau_\beta^\dagger = \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta} |\tau_\alpha\rangle\langle\tau_\beta|. \quad (56)$$

In the third and fourth forms, we let a single Greek index stand for a pair of Latin indices. The fourth form uses our operator bra-ket notation (see Appendix B for more details). To see that the form (56) is consistent with the matrix elements in Eq. (55), we calculate

$$\langle e_l | \mathcal{S}(|e_j\rangle\langle e_k|) | e_m \rangle = \left\langle e_l \left| \sum_{n,p,q,r} \mathcal{S}_{np,qr} |e_n\rangle\langle e_p| (|e_j\rangle\langle e_k|) |e_r\rangle\langle e_q| \right| e_m \right\rangle = \mathcal{S}_{lj,mk}. \quad (57)$$

Equation (56) is written in terms of an outer-product operator basis $\tau_\alpha = \tau_{jk} = |e_j\rangle\langle e_k|$, but the third and fourth forms actually hold for any orthonormal operator basis. Just to drive home the connection between a superoperator and its OP, we write Eq. (54) in yet another form,

$$\mathcal{I}_R \otimes \mathcal{S}(|\Psi\rangle\langle\Psi|) = \sum_{j,k} \mathcal{S}_{lj,mk} |f_j\rangle\langle f_k| \otimes |e_l\rangle\langle e_m|, \quad (58)$$

which should be compared with the abstract form of \mathcal{S} in Eq. (56).

These tools in hand, we look at the action of OP in a slightly different way,

$$\mathcal{I}_R \otimes \mathcal{S}(|\Psi\rangle\langle\Psi|) = \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta} I_R \otimes \tau_\alpha |\Psi\rangle\langle\Psi| I_R \otimes \tau_\beta^\dagger = \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta} |\Phi_{\tau_\alpha}\rangle\langle\Phi_{\tau_\beta}|. \quad (59)$$

What the OP map does is to transform \mathcal{S} represented in an orthonormal operator basis τ_α to an equivalent operator represented in the orthonormal RQ basis $|\Phi_{\tau_\alpha}\rangle$. Indeed, we find that

$$\langle\Phi_A|\mathcal{I}_R \otimes \mathcal{S}(|\Psi\rangle\langle\Psi|)|\Phi_B\rangle = \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta} \langle\Phi_A|\Phi_{\tau_\alpha}\rangle\langle\Phi_{\tau_\beta}|\Phi_B\rangle = \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta} (A|\tau_\alpha)(\tau_\beta|B) = (A|\mathcal{S}|B). \quad (60)$$

Thus the OP of a superoperator \mathcal{S} operates in the same way as \mathcal{S} would operate to the right and left if we just did Dirac algebra using operator bra-ket notation. This *left-right action* of \mathcal{S} is quite distinct from the way we usually use \mathcal{S} , called the *ordinary action*, where we put an operator in place of the \odot symbol. The left-right action of \mathcal{S} receives its physical significance in terms of the OP'ed RQ operator $\mathcal{I}_R \otimes \mathcal{S}(|\Psi\rangle\langle\Psi|)$. The OP isomorphism between superoperators in their left-right action and operators acting on RQ is called the *Choi-Jamiolkowski isomorphism*.

We now have the tools to deal summarily with complete positivity. We are trying to show that any completely positive superoperator \mathcal{A} has a Kraus decomposition. The key point is that complete positivity requires that the OP of \mathcal{A} , i.e., $\mathcal{I}_R \otimes \mathcal{A}(|\Psi\rangle\langle\Psi|)$, be a positive operator, but Eq. (60) now shows this to be equivalent to the requirement that \mathcal{A} be positive relative to its left-right action, which we write as $\mathcal{A} \geq 0$. Any such left-right positive superoperator has many Kraus decompositions, including its orthogonal decomposition, just like a positive operator has many decompositions.

Indeed, if we have an ensemble decomposition of $\mathcal{I}_R \otimes \mathcal{A}(|\Psi\rangle\langle\Psi|)$, we can simply unVEC the unnormalized states in the decomposition to get a Kraus decomposition of \mathcal{A} . Explicitly, given an ensemble decomposition,

$$\mathcal{I}_R \otimes \mathcal{A}(|\Psi\rangle\langle\Psi|) = \sum_{\alpha} |\Phi_{\alpha}\rangle\langle\Phi_{\alpha}|, \quad (61)$$

\mathcal{A} has the Kraus decomposition

$$\mathcal{A} = \sum_{\alpha} |A_{\alpha}\rangle\langle A_{\alpha}| = \sum_{\alpha} A_{\alpha} \odot A_{\alpha}^{\dagger}, \quad (62)$$

where the Kraus operators A_{α} are recovered from the ensemble vectors $|\Phi_{\alpha}\rangle$ using Eq. (49), i.e., $(A_{\alpha})_{kj} = \langle f_j, e_k | \Phi_{\alpha} \rangle$.

It is instructive to see how this works out for the transposition superoperator (43):

$$\mathcal{T} = \sum_{j,k} \tau_{jk} \odot \tau_{jk} = \sum_{j,k} |\tau_{jk}\rangle\langle\tau_{jk}^{\dagger}| = \sum_{j,k} |\tau_{jk}\rangle\langle\tau_{kj}|. \quad (63)$$

The left-right eigenoperators of \mathcal{T} are clearly the operators τ_{jj} for $j = 1, \dots, D$ —these operators have eigenvalue 1—and the operators $(\tau_{jk} \pm \tau_{kj})/\sqrt{2}$, for all pairs of indices—these operators

have eigenvalues ± 1 . The negative eigenvalues shows that \mathcal{T} is not completely positive. This demonstration is clearly just a rewrite of the previous one, which used the OP'ed state directly, but it illustrates how the left-right action tells us about the OP of a superoperator.

Now we get for free two important results. First, a completely positive superoperator has a left-right eigendecomposition with no more than D^2 eigenoperators with nonzero eigenvalue. This means that any quantum operation has a Kraus decomposition with no more than D^2 Kraus operators. Via our proof of the Kraus representation theorem, we can construct a measurement model in which the ancilla has no more than D^2 dimensions if the operation is trace-preserving and no more than $D^2 + 1$ dimensions if the operation is strictly trace-decreasing. Second, by generalizing the HJW theorem for decompositions of positive operators, we get a similar theorem for Kraus decompositions: two sets of Kraus operators, $\{A_\alpha\}$ and $\{B_\alpha\}$, give rise to the same completely positive superoperator if and only if they are related by a unitary matrix $V_{\beta\alpha}$, i.e.,

$$B_\beta = \sum_\alpha V_{\beta\alpha} A_\alpha \quad (64)$$

(in the standard way, if one Kraus decomposition has a smaller number of operators, it is extended by appending zero operators).

An example will show the importance of this decomposition theorem. Consider a von Neumann measurement in the basis $|e_j\rangle$. If we forget the result of the measurement, the trace-preserving operation that describes the process is

$$\mathcal{A} = \sum_{j=1}^D |e_j\rangle\langle e_j| \odot |e_j\rangle\langle e_j| = \sum_j P_j \odot P_j. \quad (65)$$

This operation corresponds to writing the input density operator in the basis $|e_j\rangle$ and then setting all the off-diagonal terms to zero. Physically, it is the ultimate decoherence process: it wipes out all the coherence in the basis $|e_j\rangle$ and replaces the input state with the corresponding incoherent mixture of the basis states $|e_j\rangle\langle e_j|$. Though it would seem to have nothing to do with unitary evolutions, we can nonetheless write this operation as a mixture (convex combination) of unitary operators. If we transform the projectors P_j using the unitary matrix

$$V_{kj} = \frac{1}{\sqrt{D}} e^{2\pi i k j / D}, \quad (66)$$

we get new Kraus operators

$$\frac{1}{\sqrt{D}} U_k = \sum_j V_{kj} P_j = \frac{1}{\sqrt{D}} \sum_j e^{2\pi i k j / D} |e_j\rangle\langle e_j|, \quad (67)$$

The operators U_k are clearly unitary operators—they're written in their eigendecomposition with eigenvalues that are phases—in terms of which the operation becomes

$$\mathcal{A} = \frac{1}{D} \sum_{k=1}^D U_k \odot U_k^\dagger. \quad (68)$$

Thinking in terms of this Kraus decomposition, \mathcal{A} describes a process where one chooses one of the D unitaries out of a hat and applies it to the system, not knowing which unitary has been chosen—all have equal probability $1/D$.

What we have shown in this section is that a superoperator is completely positive if and only if it is a positive superoperator relative to the left-right action. For a quantum operation we must add the trace-decreasing condition, which now can be put in the compact form

$$\mathcal{A}^*(I) = \sum_{\alpha} A_{\alpha}^{\dagger} A_{\alpha} \leq I , \quad (69)$$

with equality if and only if the operation is trace preserving (see Appendix B for notation).

We can now summarize the three equivalent ways we have developed for describing a quantum operation:

- A quantum operation is a superoperator that can be realized by a measurement model.
- A quantum operation is a superoperator with a Kraus decomposition that satisfies the trace-decreasing condition (32).
- A quantum operation is a trace-decreasing, completely positive superoperator.

Additional topics to be added at some future time:

1. Section on POVMs, including the Neumark extension (do qubit cases like trine and cardinal directions) and Neumark interferometers.
3. Section on convex structures: POVM elements, POVMs, and operations.

Appendix A. Convex linearity implies linearity

Consider a convex linear map \mathcal{A} from density operators (trace-one positive operators) to positive operators, as in Eq. (35).

We first extend the action of \mathcal{A} to all positive operators G in the obvious way, i.e.,

$$\mathcal{A}(G) \equiv \text{tr}(G) \mathcal{A}\left(\frac{G}{\text{tr}(G)}\right) . \quad (70)$$

This extension is clearly linear under (i) scalar multiplication by positive numbers,

$$\mathcal{A}(aG) = \text{tr}(aG) \mathcal{A}\left(\frac{aG}{\text{tr}(aG)}\right) = a \text{tr}(G) \mathcal{A}\left(\frac{G}{\text{tr}(G)}\right) = a \mathcal{A}(G) , \quad (71)$$

and (ii) addition,

$$\begin{aligned} \mathcal{A}(G + F) &= \text{tr}(G + F) \mathcal{A}\left(\frac{G + F}{\text{tr}(G + F)}\right) && \text{(definition of extension)} \\ &= \text{tr}(G + F) \mathcal{A}\left(\frac{\text{tr}(G)}{\text{tr}(G + F)} \frac{G}{\text{tr}(G)} + \frac{\text{tr}(F)}{\text{tr}(G + F)} \frac{F}{\text{tr}(F)}\right) && \text{(rewrite)} \\ &= \text{tr}(G) \mathcal{A}\left(\frac{G}{\text{tr}(G)}\right) + \text{tr}(F) \mathcal{A}\left(\frac{F}{\text{tr}(F)}\right) && \text{(convex linearity)} \\ &= \mathcal{A}(G) + \mathcal{A}(F) && \text{(definition of extension)} . \end{aligned} \quad (72)$$

The third equality is the crucial one; it follows from convex linearity on unit-trace positive operators.

We now extend the action of \mathcal{A} to all Hermitian operators in the following way. We first extend the action of \mathcal{A} to the difference between two positive operators by defining

$$\mathcal{A}(G - F) \equiv \mathcal{A}(G) - \mathcal{A}(F) . \quad (73)$$

Of course, since the difference can be written in many ways, we have to check that we get the same output no matter how the difference is written; i.e., we have to check that $\mathcal{A}(G_1) - \mathcal{A}(F_1) = \mathcal{A}(G_2) - \mathcal{A}(F_2)$ when $G_1 - F_1 = G_2 - F_2$. Since $G_1 + F_2 = G_2 + F_1$, we have $\mathcal{A}(G_1) + \mathcal{A}(F_2) = \mathcal{A}(G_1 + F_2) = \mathcal{A}(G_2 + F_1) = \mathcal{A}(G_2) + \mathcal{A}(F_1)$, which establishes the result we want. The eigendecomposition of a Hermitian operator,

$$H = \sum_{\lambda} \lambda P_{\lambda} , \quad (74)$$

allows us to write H as a difference between two positive operators,

$$H = \underbrace{\sum_{\lambda \geq 0} \lambda P_{\lambda}}_{\equiv H_+} + \underbrace{\sum_{\lambda < 0} \lambda P_{\lambda}}_{\equiv -H_-} . \quad (75)$$

Thus the action of the extended map on a Hermitian operator H can be defined as

$$\mathcal{A}(H) = \mathcal{A}(H_+) - \mathcal{A}(H_-) . \quad (76)$$

The extension is linear under (i) scalar multiplication by real numbers,

$$\mathcal{A}(aH) = \frac{a}{|a|} \left(\mathcal{A}(|a|H_+) - \mathcal{A}(|a|H_-) \right) = a \left(\mathcal{A}(H_+) - \mathcal{A}(H_-) \right) = a\mathcal{A}(H) , \quad (77)$$

and (ii) addition,

$$\begin{aligned} \mathcal{A}(H + K) &= \mathcal{A}(H_+ - H_- + K_+ - K_-) \\ &= \mathcal{A}\left((H_+ + K_+) - (H_- + K_-)\right) \\ &= \mathcal{A}(H_+ + K_+) - \mathcal{A}(H_- + K_-) \quad [\text{Eq. (73)}] \\ &= \mathcal{A}(H_+) + \mathcal{A}(K_+) - \mathcal{A}(H_-) - \mathcal{A}(K_-) \quad [\text{Eq. (72)}] \\ &= \mathcal{A}(H_+ - H_-) + \mathcal{A}(K_+ - K_-) \quad [\text{Eq. (73)}] \\ &= \mathcal{A}(H) + \mathcal{A}(K) . \end{aligned} \quad (78)$$

This establishes that the extended map acts linearly on the real vector space of Hermitian operators. Though this is really all we need, the map can be extended to a linear map on the complex vector space of all operators by the standard complexification.

Appendix B. Superoperators

The space of linear operators acting on a D -dimensional Hilbert space \mathcal{H} is a D^2 -dimensional complex vector space. We introduce operator “kets” $|A\rangle = A$ and “bras” $\langle A| = A^\dagger$, distinguished from vector bras and kets by the use of smooth brackets. The natural inner product on the space of operators can be written as $\langle A|B\rangle = \text{tr}(A^\dagger B)$. An orthonormal basis $|e_j\rangle$ induces an orthonormal operator basis

$$|e_j\rangle\langle e_k| = \tau_{jk} \equiv \tau_{\alpha} , \quad (79)$$

where the Greek index is an abbreviation for two Roman indices. Not all orthonormal operator bases are of this outer-product form. In the following, τ_α denotes the operators in a general orthonormal operator basis, which can be specialized to an outer-product basis.

The space of superoperators on \mathcal{H} , i.e., linear maps on operators, is a D^4 -dimensional complex vector space. A superoperator \mathcal{S} is specified by its “matrix elements”

$$\mathcal{S}_{lj,mk} \equiv \langle e_l | \mathcal{S}(|e_j\rangle\langle e_k|) | e_m \rangle , \quad (80)$$

for the superoperator can be written in terms of its matrix elements as

$$\mathcal{S} = \sum_{l,j,m,k} \mathcal{S}_{lj,mk} \underbrace{|e_l\rangle\langle e_j|}_{=\tau_\alpha} \odot \underbrace{|e_k\rangle\langle e_m|}_{=\tau_\beta^\dagger} = \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta} \tau_\alpha \odot \tau_\beta^\dagger = \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta} |\tau_\alpha\rangle\langle\tau_\beta| . \quad (81)$$

The *ordinary action* of \mathcal{S} on an operator A is obtained by dropping an operator A into the center of this representation of \mathcal{S} , in place of the \odot (read “o-dot”), i.e.,

$$\mathcal{A}(A) = \sum_{\alpha,\beta} \mathcal{A}_{\alpha\beta} \tau_\alpha A \tau_\beta^\dagger . \quad (82)$$

The ordinary action is used above to generate the matrix elements, as we see from

$$\langle e_l | \mathcal{S}(|e_j\rangle\langle e_k|) | e_m \rangle = \left\langle e_l \left| \sum_{n,p,q,r} \mathcal{S}_{np,qr} |e_n\rangle\langle e_p| (|e_j\rangle\langle e_k|) |e_r\rangle\langle e_q| \right| e_m \right\rangle = \mathcal{S}_{lj,mk} . \quad (83)$$

The symbol \odot is used in the abstract representation (81) of a superoperator, because this abstract representation really does involve a tensor product of operators; since it is not a tensor product between different subsystems, however, we denote it by a symbol other than \otimes .

There is clearly another way that \mathcal{S} can act on an operator A , the *left-right action*,

$$\mathcal{S}|A \rangle \equiv \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta} |\tau_\alpha\rangle\langle\tau_\beta| (A) = \sum_{\alpha,\beta} \tau_\alpha \mathcal{S}_{\alpha\beta} \text{tr}(\tau_\beta^\dagger A) , \quad (84)$$

in terms of which the matrix elements are

$$\mathcal{S}_{\alpha\beta} = (\tau_\alpha | \mathcal{S} | \tau_\beta) = \left(|e_l\rangle\langle e_j| \left| \mathcal{S} \right| |e_m\rangle\langle e_k| \right) = \langle e_l | \mathcal{S}(|e_j\rangle\langle e_k|) | e_m \rangle = \mathcal{S}_{lj,mk} . \quad (85)$$

This expression, valid only in an outer-product basis, provides the fundamental connection between the two actions of a superoperator.

With respect to the left-right action, superoperator algebra works just like Dirac operator algebra. Multiplication of superoperators \mathcal{R} and \mathcal{S} is given by

$$\mathcal{R}\mathcal{S} = \sum_{\alpha,\beta,\gamma} \mathcal{R}_{\alpha\gamma} \mathcal{S}_{\gamma\beta} |\tau_\alpha\rangle\langle\tau_\beta| , \quad (86)$$

and the “left-right” adjoint, defined by

$$(A | \mathcal{S}^\dagger | B) = (B | \mathcal{S} | A)^* , \quad (87)$$

is given by

$$\mathcal{S}^\dagger = \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta}^* |\tau_\beta\rangle\langle\tau_\alpha| = \sum_{\alpha,\beta} \mathcal{S}_{\beta\alpha}^* |\tau_\alpha\rangle\langle\tau_\beta| = \sum_{\alpha,\beta} \mathcal{S}_{\alpha\beta}^* \tau_\beta \odot \tau_\alpha^\dagger . \quad (88)$$

The question is whether we can give a physical interpretation to the left-right action, given that the ordinary action of a superoperator is the one that is used by a quantum operation to map input states to unnormalized output states. The chief answer is provided by the proof of complete positivity, which shows that under its left-right action, a superoperator is equivalent to the operator obtained by applying $\mathcal{I}_R \otimes \mathcal{S}$ to an (unnormalized) maximally entangled state.

With respect to the ordinary action, superoperator multiplication, denoted as a composition $\mathcal{R} \circ \mathcal{S}$, is given by

$$\mathcal{R} \circ \mathcal{S} = \sum_{\alpha, \beta, \gamma, \delta} \mathcal{R}_{\gamma\delta} \mathcal{S}_{\alpha\beta} \tau_\gamma \tau_\alpha \odot \tau_\beta^\dagger \tau_\delta^\dagger. \quad (89)$$

The adjoint with respect to the ordinary action, denoted by \mathcal{S}^* , is defined by

$$\text{tr}([\mathcal{S}^*(B)]^\dagger A) = \text{tr}(B^\dagger \mathcal{S}(A)). \quad (90)$$

In terms of a representation in an operator basis, this “star” adjoint becomes

$$\mathcal{S}^* = \sum_{\alpha, \beta} \mathcal{S}_{\alpha\beta}^* \tau_\alpha^\dagger \odot \tau_\beta. \quad (91)$$

Notice that

$$(\mathcal{R} \circ \mathcal{S})^\dagger = \mathcal{R}^\dagger \circ \mathcal{S}^\dagger \quad \text{and} \quad (\mathcal{R}\mathcal{S})^* = \mathcal{R}^* \mathcal{S}^*. \quad (92)$$

We can formalize the connection between the two kinds of action by defining an operation, called “sharp,” which exchanges the two:

$$\mathcal{S}^\# |A\rangle \equiv \mathcal{S}(A). \quad (93)$$

Simple consequences of the definition are that

$$(\mathcal{S}^\#)^\dagger = (\mathcal{S}^*)^\#, \quad (94)$$

$$(\mathcal{R} \circ \mathcal{S})^\# = \mathcal{R}^\# \mathcal{S}^\#. \quad (95)$$

The matrix elements of $\mathcal{S}^\#$ are given by

$$\begin{aligned} \mathcal{S}_{lj, mk}^\# &= \left(|e_l\rangle\langle e_j| \mathcal{S}^\# |e_m\rangle\langle e_k| \right) \\ &= \text{tr}(|e_j\rangle\langle e_l| \mathcal{S}(|e_m\rangle\langle e_k|)) \\ &= \langle e_l | \mathcal{S}(|e_m\rangle\langle e_k|) | e_j \rangle \\ &= \mathcal{S}_{lm, jk}, \end{aligned} \quad (96)$$

which implies that

$$\mathcal{S}^\# = \sum_{lj, mk} \mathcal{S}_{lj, mk} |e_l\rangle\langle e_m| \odot |e_k\rangle\langle e_j|. \quad (97)$$

We now see that there are two important matrices that represent a superoperator in an orthonormal operator basis τ_α : $\mathcal{S}_{\alpha\beta} = (\tau_\alpha | \mathcal{S} | \tau_\beta)$, which for a quantum operation is called the *process matrix*, and

$$\mathcal{S}_{\alpha\beta}^\# = (\tau_\alpha | \mathcal{S} | \tau_\beta) = \text{tr}(\tau_\alpha^\dagger \mathcal{S}^\# | \tau_\beta) = \text{tr}(\tau_\alpha^\dagger \mathcal{S}(\tau_\beta)) = \sum_{\gamma, \delta} \mathcal{S}_{\gamma\delta} \text{tr}(\tau_\alpha^\dagger \tau_\gamma \tau_\beta \tau_\delta^\dagger), \quad (98)$$

which for a quantum operation is called the *Choi matrix*. The Choi matrix is the one that maps an input operator, represented in the operator basis τ_α , to the output operator represented in

that basis. It is useful to be able to map between the two representations, and we do that using the final form in Eq. (98), which is called the *Choi transformation*. In the outer-product basis $\tau_\alpha = \tau_{jk} = |e_j\rangle\langle e_k|$, the Choi transformation is just a transposition of two indices as in Eq. (96).

Two important superoperators are the identity operators with respect to the two kinds of action. The identity superoperator with respect to the ordinary action is

$$\mathcal{I} = I \odot I = \sum_{j,k} |e_j\rangle\langle e_j| \odot |e_k\rangle\langle e_k|. \quad (99)$$

This superoperator is Hermitian in both senses, i.e., $\mathcal{I} = \mathcal{I}^\dagger = \mathcal{I}^*$. It is the identity superoperator relative to the ordinary action because $\mathcal{I}(A) = A$ for all operators A , but its left-right action gives $\mathcal{I}|A\rangle = \text{tr}(A)I$. A Kraus decomposition for \mathcal{I} consists of the unit operator I .

The identity superoperator with respect to the left-right action is

$$\mathbf{I} = \sum_{\alpha} |\tau_{\alpha}\rangle\langle\tau_{\alpha}| = \sum_{j,k} |e_j\rangle\langle e_k| \odot |e_k\rangle\langle e_j|. \quad (100)$$

This superoperator is also Hermitian in both senses, i.e., $\mathbf{I} = \mathbf{I}^\dagger = \mathbf{I}^*$. It is the identity superoperator relative to the left-right action because $\mathbf{I}|A\rangle = A$ for all operators A , but its ordinary action gives $\mathbf{I}(A) = \text{tr}(A)I$, which means that \mathbf{I}/D is the completely depolarizing operation that takes all input states to the maximally mixed state. Since sharpening exchanges the two kinds of action, it is clear that $\mathcal{I}^\# = \mathbf{I}$. Any orthonormal set of operators τ_α provides a Kraus decomposition of $\mathbf{I} = \sum_{\alpha} |\tau_{\alpha}\rangle\langle\tau_{\alpha}| = \sum_{\alpha} \tau_{\alpha} \odot \tau_{\alpha}^\dagger$.

Another important superoperator is the one that transposes operators in a particular orthonormal basis $|e_j\rangle$. The ordinary action of this transposition superoperator is given by

$$\mathcal{T}(|e_j\rangle\langle e_k|) = |e_k\rangle\langle e_j| \quad \Longleftrightarrow \quad \mathcal{T}(A) = \sum_{j,k} A_{kj} |e_j\rangle\langle e_k|, \quad (101)$$

so the superoperator has the abstract form

$$\mathcal{T} = \sum_{j,k} |e_j\rangle\langle e_k| \odot |e_j\rangle\langle e_k| = \sum_{j,k} \tau_{jk} \odot \tau_{jk} = \sum_{j,k} |\tau_{jk}\rangle\langle\tau_{jk}|. \quad (102)$$

The transposition superoperator is Hermitian in both senses and is unchanged by sharpening, i.e., $\mathcal{T} = \mathcal{T}^\dagger = \mathcal{T}^* = \mathcal{T}^\#$. In addition to satisfying $\mathcal{T} \circ \mathcal{T} = \mathcal{I}$, the transposition superoperator has the property that

$$\mathbf{I} \circ \mathcal{T} = \mathbf{I}, \quad (103)$$

which in view of Eq. (95), is equivalent to $\mathcal{IT} = \mathcal{I}$.

We now turn to cataloguing some general properties of superoperators that characterize the most interesting classes of superoperators.

A superoperator is left-right Hermitian, i.e., $\mathcal{A}^\dagger = \mathcal{A}$, if and only if it has an eigendecomposition relative to the left-right action,

$$\mathcal{A} = \sum_{\alpha} \mu_{\alpha} |\tau_{\alpha}\rangle\langle\tau_{\alpha}| = \sum_{\alpha} \mu_{\alpha} \tau_{\alpha} \odot \tau_{\alpha}^\dagger, \quad (104)$$

where the μ_{α} are real (left-right) eigenvalues and the operators τ_{α} are orthonormal (left-right) eigenoperators. Notice that for a left-right Hermitian operator, the star-adjoint is given by

$$\mathcal{A}^* = \sum_{\alpha} \mu_{\alpha} \tau_{\alpha}^\dagger \odot \tau_{\alpha} = \sum_{\alpha} \mu_{\alpha} |\tau_{\alpha}^\dagger\rangle\langle\tau_{\alpha}^\dagger|, \quad (105)$$

which means that \mathcal{A}^* is also left-right Hermitian, with the same left-right eigenvalues as \mathcal{A} , but with the corresponding eigenoperators being τ_α^\dagger .

There is another useful way to characterize left-right Hermiticity: a superoperator is left-right Hermitian if and only if it maps all Hermitian operators to Hermitian operators under the ordinary action. This is the first hint that the left-right action has important consequences for the more physical ordinary action. Before proving this result, however, we need to do a little preliminary work. Let \mathcal{S} be a superoperator, and let $|e_j\rangle$ be an orthonormal basis, which induces an orthonormal operator basis $|e_j\rangle\langle e_k|$. Notice that

$$\begin{aligned}\langle e_l|\mathcal{S}^\dagger(|e_j\rangle\langle e_k|)|e_m\rangle &= \langle e_l|\langle e_j|\mathcal{S}^\dagger|e_m\rangle\langle e_k| \\ &= \langle e_m\rangle\langle e_k|\mathcal{S}|e_l\rangle\langle e_j|)^* \\ &= \langle e_m|\mathcal{S}(|e_k\rangle\langle e_j|)|e_l\rangle^* \\ &= \langle e_l|[\mathcal{S}(|e_k\rangle\langle e_j|)]^\dagger|e_m\rangle.\end{aligned}\tag{106}$$

Here the first and third equalities follow from relating the ordinary action of a superoperator to its left-right action [Eq. (85)], the second equality follows from the definition of the left-right adjoint of \mathcal{S} [Eq. (87)], and the fourth equality follows from the definition of the operator adjoint. Equation (106) gives the relation between the operator adjoint and the left-right superoperator adjoint:

$$\mathcal{S}^\dagger(|e_j\rangle\langle e_k|) = [\mathcal{S}(|e_k\rangle\langle e_j|)]^\dagger.\tag{107}$$

Thus we have that $\mathcal{S} = \mathcal{S}^\dagger$, i.e., \mathcal{S} is left-right Hermitian, if and only if

$$\mathcal{S}(|e_j\rangle\langle e_k|) = [\mathcal{S}(|e_k\rangle\langle e_j|)]^\dagger\tag{108}$$

for all j and k . This result allows us to prove the desired theorem easily.

Theorem. A superoperator \mathcal{A} is left-right Hermitian if and only if it maps all Hermitian operators to Hermitian operators.

Proof: First suppose \mathcal{A} is left-right Hermitian, i.e., $\mathcal{A} = \mathcal{A}^\dagger$. This implies that \mathcal{A} has a complete, orthonormal set of eigenoperators τ_α , with real eigenvalues μ_α . Using the eigendecomposition (104), we have for any Hermitian operator H ,

$$\mathcal{A}(H) = \sum_\alpha \mu_\alpha \tau_\alpha H \tau_\alpha^\dagger = \mathcal{A}(H)^\dagger.\tag{109}$$

Now suppose \mathcal{A} maps all Hermitian operators to Hermitian operators. Letting $\tau_{jk} = |e_j\rangle\langle e_k|$, it follows that

$$\begin{aligned}\mathcal{A}(\tau_{jk}) &= \mathcal{A}\left(\frac{1}{2}(\tau_{jk} + \tau_{kj}) + i\frac{-i}{2}(\tau_{jk} - \tau_{kj})\right) \\ &= \mathcal{A}\left(\frac{1}{2}(\tau_{jk} + \tau_{kj})\right) + i\mathcal{A}\left(\frac{-i}{2}(\tau_{jk} - \tau_{kj})\right) \quad (\text{linearity of } \mathcal{A}) \\ &= \left[\mathcal{A}\left(\frac{1}{2}(\tau_{jk} + \tau_{kj})\right)\right]^\dagger + i\left[\mathcal{A}\left(\frac{-i}{2}(\tau_{jk} - \tau_{kj})\right)\right]^\dagger \quad (\text{the assumption}) \\ &= \left[\mathcal{A}\left(\frac{1}{2}(\tau_{jk} + \tau_{kj})\right) - i\mathcal{A}\left(\frac{-i}{2}(\tau_{jk} - \tau_{kj})\right)\right]^\dagger \quad (\text{antilinearity of operator adjoint}) \\ &= \left[\mathcal{A}\left(\frac{1}{2}(\tau_{jk} + \tau_{kj}) - i\frac{-i}{2}(\tau_{jk} - \tau_{kj})\right)\right]^\dagger \quad (\text{linearity of } \mathcal{A}) \\ &= [\mathcal{A}(\tau_{kj})]^\dagger.\end{aligned}\tag{110}$$

Equation (108) then implies that $\mathcal{A} = \mathcal{A}^\dagger$. **QED**

As noted above, if \mathcal{A} is left-right Hermitian, \mathcal{A}^* is also left-right Hermitian and thus maps Hermitian operators to Hermitian operators. In particular, we have that $\mathcal{A}^*(I)$ is a Hermitian operator.

A superoperator is *trace preserving* if, under the ordinary action, it leaves the trace unchanged, i.e., if $\text{tr}(A) = \text{tr}(\mathcal{A}(A)) = \text{tr}([\mathcal{A}^*(I)]^\dagger A)$ for all operators A . Thus \mathcal{A} is trace preserving if and only if $\mathcal{A}^*(I) = I$.

A superoperator is said to be *positive* if it maps positive operators to positive operators under the ordinary action. Since any Hermitian operator can be written as the difference between two positive operators, it follows that a positive superoperator maps Hermitian operators to Hermitian operators and thus, by the above theorem, that a positive operator is left-right Hermitian.

A positive superoperator is *trace decreasing* if it does not increase the trace of positive operators, i.e., if $\text{tr}(G) \geq \text{tr}(\mathcal{A}(G)) = \text{tr}([\mathcal{A}^*(I)]^\dagger G) = \text{tr}(\mathcal{A}^*(I)G)$ for all positive operators G . Since any one-dimensional projector is a positive operator, a positive superoperator must satisfy $\langle \psi | \mathcal{A}^*(I) | \psi \rangle \leq 1$ for all $|\psi\rangle$, which implies that $\mathcal{A}^*(I) \leq I$. We can see that this is a sufficient condition for \mathcal{A} to be positive by writing G in its eigendecomposition:

$$\text{tr}(\mathcal{A}^*(I)G) = \sum_j \lambda_j \langle e_j | \mathcal{A}^*(I) | e_j \rangle \leq \sum_j \lambda_j = \text{tr}(G). \quad (111)$$

Our conclusion is that a positive superoperator \mathcal{A} is trace decreasing if and only if $\mathcal{A}^*(I) \leq I$. Notice that $\mathcal{A}^*(I) = E \leq I$ is the POVM element associated with operation \mathcal{A} .

A superoperator \mathcal{A} is *completely positive* if it and all its extensions $\mathcal{I} \otimes \mathcal{A}$ to tensor-product spaces, where \mathcal{I} is the unit superoperator on the appended space, are positive. We show in the main text that \mathcal{A} is completely positive if and only if it is positive relative to the left-right action, i.e., $(A|\mathcal{A}|A) \geq 0$ for all operators A , which is equivalent to saying that \mathcal{A} is left-right Hermitian with nonnegative left-right eigenvalues. This is the reason why the left-right action has physical significance. A *quantum operation* is a trace-decreasing, completely positive superoperator.

Since a superoperator is left-right Hermitian if and only if it has an eigendecomposition as in Eq. (104), we can conclude, by grouping together positive and negative eigenvalues, that being left-right Hermitian is equivalent to being the difference between two completely positive superoperators. Using the above theorem, we have that a superoperator takes all Hermitian operators to Hermitian operators if and only if it is the difference between two completely positive superoperators. In particular, a positive superoperator is the difference between two completely positive superoperators. A positive operator that is not completely positive has one or more negative left-right eigenvalues.

A completely positive superoperator has a left-right eigendecomposition,

$$\mathcal{A} = \sum_\alpha \mu_\alpha |\tau_\alpha\rangle\langle\tau_\alpha| = \sum_\alpha \sqrt{\mu_\alpha} \tau_\alpha \odot \sqrt{\mu_\alpha} \tau_\alpha^\dagger = \sum_\alpha |\sqrt{\mu_\alpha} \tau_\alpha\rangle\langle\sqrt{\mu_\alpha} \tau_\alpha|, \quad (112)$$

in which all the eigenvalues μ_α are nonnegative. The operators $\sqrt{\mu_\alpha} \tau_\alpha$ are special Kraus operators, special because they are orthogonal. We can immediately get two useful results. First, notice that there are at most D^2 Kraus operators $\sqrt{\mu_\alpha} \tau_\alpha$ in the Kraus decomposition associated with the eigendecomposition. If we use this Kraus decomposition to construct a measurement model for a quantum operation, we will need an ancilla with at most D^2 dimensions if \mathcal{A} is trace-preserving and at most $D^2 + 1$ dimensions if \mathcal{A} is strictly trace-decreasing. Thus we conclude that any quantum operation can be realized by a measurement model where the ancilla has at most these numbers of dimensions. Second, we can generalize the HJW theorem for decompositions of positive operators

to a similar statement about Kraus decompositions of complete positive superoperators: two sets of operators, $\{A_\alpha\}$ and $\{B_\beta\}$, are Kraus operators for the same completely positive superoperator if and only if there is a unitary matrix $V_{\beta\alpha}$ such that

$$A_\alpha = \sum_\beta V_{\alpha\beta} B_\beta . \quad (113)$$

Appendix C. Mixed-state measurement models and extreme operations

Suppose that an operation \mathcal{A} has a mixed-state measurement model, i.e.,

$$\mathcal{A}(\rho) = \text{tr}_A(P U \rho \otimes \sigma U^\dagger) = \sum_k \lambda_k \text{tr}_A(P U \rho \otimes |e_k\rangle\langle e_k| U^\dagger) = \sum_k \lambda_k \mathcal{A}_k(\rho) , \quad (114)$$

where the ancilla projector P has eigendecomposition

$$P = \sum_j |f_j\rangle\langle f_j| . \quad (115)$$

Clearly \mathcal{A} is a convex combination of the quantum operations \mathcal{A}_k that arise from the eigenstates of the ancilla's initial density operator σ , i.e.,

$$\mathcal{A}_k(\rho) = \text{tr}_A(P U \rho \otimes |e_k\rangle\langle e_k| U^\dagger) = \sum_j A_{kj} \rho A_{kj}^\dagger , \quad (116)$$

where the Kraus operators are defined by

$$A_{kj} \equiv \langle f_j|U|e_k\rangle . \quad (117)$$

By a nontrivial mixed-state model, we mean that the operations \mathcal{A}_k are not all identical, thus making \mathcal{A} a proper convex combination of the operations \mathcal{A}_k . We can conclude immediately that if \mathcal{A} is an extreme point, it cannot be given a nontrivial mixed-state measurement model.

We now show that any decomposition of an operation into a proper convex combination has a corresponding nontrivial mixed-state measurement model. Suppose that a quantum operation,

$$\mathcal{A} = \sum_k \lambda_k \mathcal{A}_k , \quad (118)$$

is a convex combination of K operations, \mathcal{A}_k , which are defined by Kraus decompositions

$$\mathcal{A}_k = \sum_j A_{kj} \odot A_{kj}^\dagger . \quad (119)$$

For convenience we make all the Kraus decompositions have the same number of Kraus operators, say N , by appending zero operators where necessary. We can generalize the Kraus representation theorem by constructing a pure-state measurement model for each of the \mathcal{A}_k and then putting them together to give a (nontrivial) mixed-state measurement model for \mathcal{A} .

First define the operator

$$E_k \equiv \sum_{j=1}^N A_{kj}^\dagger A_{kj} \leq I \quad (120)$$

for each k , and associate one more operator,

$$A_{k0} \equiv \sqrt{I - E_k} , \quad (121)$$

with each value of k . This allows us to write a completeness relation for each k ,

$$\sum_{j=0}^N A_{kj}^\dagger A_{kj} = I . \quad (122)$$

Now pick an ancilla whose Hilbert-space dimension is $K(N+1)$. It is convenient to regard the ancilla as being the tensor product of two systems, the first, call it B , having dimension K and the second, call it C , having dimension $N+1$. Now we reprise the steps in the proof of the Kraus representation theorem in this more general context.

Choose any product orthonormal basis $|e_k\rangle \otimes |f_j\rangle = |e_k, f_j\rangle$ for the ancilla, where $k = 1, \dots, K$ and $j = 0, \dots, N$. Partially define a joint QBC operator U by

$$U|\psi\rangle \otimes |e_k, f_0\rangle = \sum_{j=0}^N A_{kj}|\psi\rangle \otimes |e_k, f_j\rangle \quad \Longleftrightarrow \quad \langle e_l, f_j|U|e_k, f_0\rangle = \delta_{kl}A_{kj} . \quad (123)$$

The operator U is defined on the $(D \times K)$ -dimensional subspace $\mathcal{H}_Q \otimes \mathcal{H}_B \otimes R_0$, where R_0 is the one-dimensional subspace of \mathcal{H}_C spanned by $|f_0\rangle$. This partial definition preserves inner products,

$$\begin{aligned} \left(\langle \phi| \otimes \langle e_l, f_0|U^\dagger \right) \left(U|\psi\rangle \otimes |e_k, f_0\rangle \right) &= \sum_{j,j'=0}^N \langle \phi|A_{lj'}^\dagger A_{kj}|\psi\rangle \underbrace{\langle e_l|e_k\rangle \langle f_{j'}|f_j\rangle}_{=\delta_{kl}\delta_{jj'}} \\ &= \delta_{kl} \left\langle \phi \left| \underbrace{\sum_{j=0}^N A_{kj}^\dagger A_{kj}}_{=I} \right| \psi \right\rangle \\ &= \delta_{kl} \langle \phi|\psi\rangle , \end{aligned} \quad (124)$$

which means that U maps the subspace $\mathcal{H}_Q \otimes \mathcal{H}_B \otimes R_0$ unitarily to a $(D \times K)$ -dimensional subspace S_K of $\mathcal{H}_Q \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. We can extend U to a unitary operator on all of $\mathcal{H}_Q \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ by defining it to map the subspace $\mathcal{H}_Q \otimes \mathcal{H}_B \otimes R_\perp$, where R_\perp is the subspace of \mathcal{H}_C spanned by the vectors $|f_j\rangle$, $j = 1, \dots, N$, unitarily to the subspace orthogonal to S_K .

With this definition of U , we have

$$\begin{aligned} \mathcal{A}(\rho) &= \sum_k \lambda_k \mathcal{A}_k(\rho) \\ &= \sum_k \lambda_k \sum_{j=1}^N A_{kj} \rho A_{kj}^\dagger \\ &= \sum_k \lambda_k \sum_l \sum_{j=1}^N \langle e_l, f_j|U|e_k, f_0\rangle \rho \langle e_k, f_0|U^\dagger|e_l, f_j\rangle \\ &= \text{tr}_A \left(\left(\sum_l \sum_{j=1}^N |e_l, f_j\rangle \langle e_l, f_j| \right) U \left[\rho \otimes \left(\sum_k \lambda_k |e_k\rangle \langle e_k| \right) \otimes |f_0\rangle \langle f_0| \right] U^\dagger \right) \\ &= \text{tr}_A (P U \rho \otimes \sigma U^\dagger) , \end{aligned} \quad (125)$$

where

$$P \equiv \sum_l \sum_{j=1}^N |e_l, f_j\rangle \langle e_l, f_j| \quad (126)$$

is an ancilla projector and

$$\sigma = \left(\sum_k \lambda_k |e_k\rangle \langle e_k| \right) \otimes |f_0\rangle \langle f_0| \quad (127)$$

is the initial ancilla density operator. Thus we have constructed the required mixed-state measurement model for \mathcal{A} . Notice that the way the construction works is simply to construct a separate measurement model for each of the operations \mathcal{A}_k , using system C as the ancilla, and then to stitch these models together by appending to each one of the orthonormal vectors $|e_k\rangle$ from system B , thus keeping the various models from interfering with one another.

What we have shown is that an operation is not an extreme point, i.e., can be written as a convex combination of other operations, if and only if it can be realized by a nontrivial mixed-state measurement model. Alternatively, we can say that an operation is an extreme point if and only if it cannot be realized by a nontrivial mixed-state measurement model.

It is easy to see how to construct trivial mixed-state measurement models for any operation. Simply construct multiple models for a particular operation and then stitch them together with the technique used above.