2

Basic notions of quantum information

This chapter introduces the most basic objects and notions of quantum information theory, including *registers*, *states*, *channels*, and *measurements*, and investigates some of their elementary properties.

2.1 Registers and states

This first section of the chapter concerns *registers* and *states*. A register is an abstraction of a physical device in which quantum information may be stored, and the state of a register represents a description of its contents at a particular instant.

2.1.1 Registers and classical state sets

The term *register* is intended to be suggestive of a computer component in which some finite amount of data can be stored and manipulated. While this is a reasonable picture to keep in mind, it should be understood that any physical system in which a finite amount of data may be stored, and whose state may change over time, could be modeled as a register. For example, a register could represent a medium used to transmit information from a sender to a receiver. At an intuitive level, what is most important is that registers represent mathematical abstractions of physical objects, or parts of physical objects, that store information.

Definition of registers

The following formal definition of a register is intended to capture a basic but nevertheless important idea, which is that multiple registers may be viewed collectively as forming a single register. It is natural to choose an inductive definition for this reason.

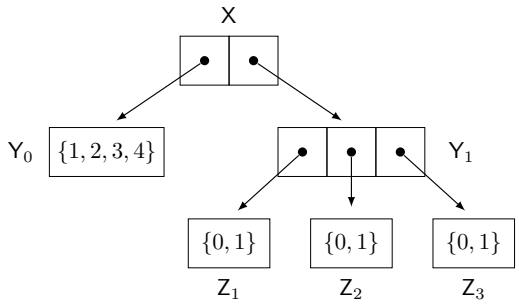

Figure 2.1 The tree associated with the registers described in Example 2.2.

Definition 2.1 A *register* X is either one of the following two objects:

1. An alphabet Σ.
2. An n-tuple $\mathsf{X} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$, where n is a positive integer and $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$ are registers.

Registers of the first type are called *simple registers* and registers of the second type are called *compound registers* when it is helpful to distinguish them.

In the case of a simple register $\mathsf{X} = \Sigma$, the alphabet Σ represents the set of *classical states* that the register may store. The classical state set associated with a compound register will be specified shortly. As is suggested by the definition, registers will be denoted by capital letters in a *sans serif* font, such as X, Y, and Z. Sometimes registers will be subscripted, such as $\mathsf{X}_1, \ldots, \mathsf{X}_n$, when it is necessary to refer to a variable number of registers or convenient to name them in this way for some other reason.

Based on Definition 2.1, one may naturally identify a tree structure with a given register, with each leaf node corresponding to a simple register. A register Y is said to be a *subregister* of X if the tree associated with Y is a subtree of the tree associated with X.

Example 2.2 Define registers X, Y_0, Y_1, Z_1, Z_2, and Z_3, as follows:

$$\begin{aligned} &\mathsf{X} = (\mathsf{Y}_0, \mathsf{Y}_1), \qquad && \mathsf{Y}_0 = \{1,2,3,4\}, && \mathsf{Z}_1 = \{0,1\}, \\ & && \mathsf{Y}_1 = (\mathsf{Z}_1, \mathsf{Z}_2, \mathsf{Z}_3), && \mathsf{Z}_2 = \{0,1\}, \\ & && && \mathsf{Z}_3 = \{0,1\}. \end{aligned} \tag{2.1}$$

The tree associated with the register X is illustrated in Figure 2.1. The subregisters of X include Y_0, Y_1, Z_1, Z_2, Z_3, and (trivially) X itself.

The classical state set of a register

Every register has associated with it a *classical state set*, as specified by the following definition.

Definition 2.3 The *classical state set* of a register X is determined as follows:

1. If $\mathsf{X} = \Sigma$ is a simple register, the classical state set of X is Σ.
2. If $\mathsf{X} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$ is a compound register, the classical state set of X is the Cartesian product

$$\Sigma = \Gamma_1 \times \cdots \times \Gamma_n, \tag{2.2}$$

where Γ_k denotes the classical state set associated with the register Y_k for each $k \in \{1, \ldots, n\}$.

Elements of a register's classical state set are called *classical states* of that register.

The term *classical state* is intended to be suggestive of the classical notion of a state in computer science. Intuitively speaking, a classical state of a register can be recognized unambiguously, like the values 0 and 1 stored by a single bit memory component. The term *classical state* should not be confused with the term *state*, which by default will mean *quantum state* rather than *classical state* throughout this book.

A register is said to be *trivial* if its classical state set contains just a single element. While trivial registers are useless from the viewpoint of information processing, it is mathematically convenient to allow for this possibility. The reader will note, however, that registers with *empty* classical state sets are disallowed by the definition. This is consistent with the idea that registers represent physical systems; while it is possible that a physical system could have just one possible classical state, it is nonsensical for a system to have no states whatsoever.

Reductions of classical states

There is a straightforward way in which each classical state of a register uniquely determines a classical state for each of its subregisters. To be more precise, suppose that

$$\mathsf{X} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n) \tag{2.3}$$

is a compound register. Let $\Gamma_1, \ldots, \Gamma_n$ denote the classical state sets of the registers $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$, respectively, so that the classical state set of X is equal to $\Sigma = \Gamma_1 \times \cdots \times \Gamma_n$. A given classical state $a = (b_1, \ldots, b_n)$ of X then

determines that the classical state of Y_k is $b_k \in \Gamma_k$, for each $k \in \{1, \ldots, n\}$. By applying this definition recursively, one defines a unique classical state of each subregister of X.

Conversely, the classical state of any register is uniquely determined by the classical states of its simple subregisters. Every classical state of a given register X therefore uniquely determines a classical state of any register whose simple subregisters form a subset of those of X. For instance, if X takes the form (2.3), then one may wish to consider a new register

$$\mathsf{Z} = (\mathsf{Y}_{k_1}, \ldots, \mathsf{Y}_{k_m}) \tag{2.4}$$

for some choice of indices $1 \leq k_1 < \cdots < k_m \leq n$. If $a = (b_1, \ldots, b_n)$ is the classical state of X at a particular moment, then the corresponding state of Z is $(b_{k_1}, \ldots, b_{k_m})$.

2.1.2 Quantum states of registers

Quantum states, as they will be presented in this book, may be viewed as being analogous to probabilistic states, with which the reader is assumed to have some familiarity.

Probabilistic states of registers

A *probabilistic state* of a register X refers to a probability distribution, or random mixture, over the classical states of that register. Assuming the classical state set of X is Σ, a probabilistic state of X is identified with a probability vector $p \in \mathcal{P}(\Sigma)$; the value $p(a)$ represents the probability associated with a given classical state $a \in \Sigma$. It is typical that one views a probabilistic state as being a mathematical representation of the contents of a register, or of a hypothetical individual's knowledge of the contents of a register, at a particular moment.

The difference between probabilistic states and quantum states is that, whereas probabilistic states are represented by probability vectors, quantum states are represented by *density operators* (q.v. Section 1.1.2). Unlike the notion of a probabilistic state, which has a relatively clear and intuitive meaning, the notion of a quantum state can seem non-intuitive. While it is both natural and interesting to seek an understanding of why Nature appears to be well-modeled by quantum states in certain regimes, this book will not attempt to provide such an understanding: quantum states will be considered as mathematical objects and nothing more.

The complex Euclidean space associated with a register

It is helpful to introduce the following terminology to discuss quantum states in mathematical terms.

Definition 2.4 The complex Euclidean space associated with a register X is defined to be $\mathbb{C}^{\Sigma}$, for Σ being the classical state set of X.

The complex Euclidean space associated with a given register will be denoted by the same letter as the register itself, but with a *scripted* font rather than a *sans serif* font. For example, the complex Euclidean space associated with a register X will be denoted $\mathcal{X}$, and the spaces associated with registers $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$ will be denoted $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$.

The reader will note that the complex Euclidean space $\mathcal{X}$ associated with a compound register $\mathsf{X} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$ is given by the tensor product

$$\mathcal{X} = \mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n. \tag{2.5}$$

This fact follows directly from the definition stating that the classical state set of X is given by $\Sigma = \Gamma_1 \times \cdots \times \Gamma_n$, assuming that the classical state sets of $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$ are $\Gamma_1, \ldots, \Gamma_n$, respectively; one has that the complex Euclidean space associated with X is

$$\mathcal{X} = \mathbb{C}^{\Sigma} = \mathbb{C}^{\Gamma_1 \times \cdots \times \Gamma_n} = \mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n \tag{2.6}$$

for $\mathcal{Y}_1 = \mathbb{C}^{\Gamma_1}$, …, $\mathcal{Y}_n = \mathbb{C}^{\Gamma_n}$.

Definition of quantum states

As stated above, quantum states are represented by density operators. The following definition makes this precise.

Definition 2.5 A *quantum state* is a density operator of the form $\rho \in \mathrm{D}(\mathcal{X})$ for some choice of a complex Euclidean space $\mathcal{X}$.

When one refers to a quantum state of a register X, it is to be understood that the state in question takes the form $\rho \in \mathrm{D}(\mathcal{X})$ for $\mathcal{X}$ being the complex Euclidean space associated with X. It is common that the term *state* is used in place of *quantum state* in the setting of quantum information, because it is the default assumption that one is primarily concerned with quantum states (as opposed to classical states and probabilistic states) in this setting.

Convex combinations of quantum states

For every complex Euclidean space $\mathcal{X}$, the set $\mathrm{D}(\mathcal{X})$ is a convex set. For any choice of an alphabet Γ, a collection

$$\{\rho_a : a \in \Gamma\} \subseteq \mathrm{D}(\mathcal{X}) \tag{2.7}$$

of quantum states, and a probability vector $p \in \mathcal{P}(\Gamma)$, it therefore holds that the convex combination

$$\rho = \sum_{a \in \Gamma} p(a)\rho_a \tag{2.8}$$

is an element of $\mathrm{D}(\mathcal{X})$. The state ρ defined by the equation (2.8) is said to be a *mixture* of the states $\{\rho_a : a \in \Gamma\}$ according to the probability vector p.

Suppose that X is a register whose associated complex Euclidean space is $\mathcal{X}$. It is taken as an axiom that a random selection of $a \in \Gamma$ according to the probability vector p, followed by a preparation of X in the state ρ_a, results in X being in the state ρ defined in (2.8). More succinctly, random selections of quantum states are assumed to be represented by convex combinations of density operators.

Ensembles of quantum states

The notion of a probability distribution over a finite set of quantum states arises frequently in the theory of quantum information. A distribution of the form described above may be succinctly represented by a function

$$\eta : \Gamma \to \mathrm{Pos}(\mathcal{X}) \tag{2.9}$$

satisfying the constraint

$$\mathrm{Tr}\left(\sum_{a \in \Gamma} \eta(a)\right) = 1. \tag{2.10}$$

A function η of this sort is called an *ensemble* of states. The interpretation of an ensemble of states $\eta : \Gamma \to \mathrm{Pos}(\mathcal{X})$ is that, for each element $a \in \Gamma$, the operator $\eta(a)$ represents a state together with the probability associated with that state: the probability is $\mathrm{Tr}(\eta(a))$, while the state is

$$\rho_a = \frac{\eta(a)}{\mathrm{Tr}(\eta(a))}. \tag{2.11}$$

(The operator ρ_a is, of course, determined only when $\eta(a) \neq 0$. In the case that $\eta(a) = 0$ for some choice of a, one does not generally need to specify a specific density operator ρ_a, as it corresponds to a discrete event that occurs with probability zero.)

Pure states

A quantum state $\rho \in \mathrm{D}(\mathcal{X})$ is said to be a *pure state* if it has rank equal to 1. Equivalently, ρ is a pure state if there exists a unit vector $u \in \mathcal{X}$ such that

$$\rho = uu^*. \tag{2.12}$$

It follows from the spectral theorem (Corollary 1.4) that every quantum state is a mixture of pure quantum states, and moreover that a state $\rho \in \mathrm{D}(\mathcal{X})$ is pure if and only if it is an extreme point of the set $\mathrm{D}(\mathcal{X})$.

It is common that one refers to the pure state (2.12) simply as u, rather than uu^*. There is an ambiguity that arises in following this convention: if one considers two unit vectors u and $v = \alpha u$, for any choice of $\alpha \in \mathbb{C}$ with $|\alpha| = 1$, then their corresponding pure states uu^* and vv^* are equal, as

$$vv^* = |\alpha|^2 uu^* = uu^*. \tag{2.13}$$

Fortunately, this convention does not generally cause confusion; it must simply be kept in mind that every pure state corresponds to an equivalence class of unit vectors, where u and v are equivalent if and only if $v = \alpha u$ for some choice of $\alpha \in \mathbb{C}$ with $|\alpha| = 1$, and that any particular unit vector may be viewed as being a representative of a pure state from this equivalence class.

Flat states

A quantum state $\rho \in \mathrm{D}(\mathcal{X})$ is said to be a *flat state* if it holds that

$$\rho = \frac{\Pi}{\mathrm{Tr}(\Pi)} \tag{2.14}$$

for a nonzero projection operator $\Pi \in \mathrm{Proj}(\mathcal{X})$. The symbol ω will often be used to denote a flat state, and the notation

$$\omega_{\mathcal{V}} = \frac{\Pi_{\mathcal{V}}}{\mathrm{Tr}(\Pi_{\mathcal{V}})} \tag{2.15}$$

is sometimes used to denote the flat state proportional to the projection $\Pi_{\mathcal{V}}$ onto a nonzero subspace $\mathcal{V} \subseteq \mathcal{X}$. Specific examples of flat states include pure states, which correspond to the case that Π is a rank-one projection, and the *completely mixed state*

$$\omega = \frac{\mathbb{1}_{\mathcal{X}}}{\dim(\mathcal{X})}. \tag{2.16}$$

Intuitively speaking, the completely mixed state represents the state of complete ignorance, analogous to a uniform probabilistic state.

Classical states and probabilistic states as quantum states

Suppose X is a register and Σ is the classical state set of X, so that the complex Euclidean space associated with X is $\mathcal{X} = \mathbb{C}^{\Sigma}$. Within the set $\mathrm{D}(\mathcal{X})$ of states of X, one may represent the possible classical states of X in the following simple way: the operator $E_{a,a} \in \mathrm{D}(\mathcal{X})$ is taken as a representation of the register X being in the classical state a, for each $a \in \Sigma$. Through this association, probabilistic states of registers correspond to diagonal density operators, with each probabilistic state $p \in \mathcal{P}(\Sigma)$ being represented by the density operator

$$\sum_{a \in \Sigma} p(a) E_{a,a} = \mathrm{Diag}(p). \tag{2.17}$$

In this way, the set of probabilistic states of a given register form a subset of the set of all quantum states of that register (with the containment being proper unless the register is trivial).[1]

Within some contexts, it may be necessary or appropriate to specify that one or more registers are *classical registers*. Informally speaking, a classical register is one whose states are restricted to being diagonal density operators, corresponding to a classical (probabilistic) states as just described. A more formal and precise meaning of this terminology must be postponed until the section on quantum channels following this one.

Product states

Suppose $\mathsf{X} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$ is a compound register. A state $\rho \in \mathrm{D}(\mathcal{X})$ is said to be a *product state* of X if it takes the form

$$\rho = \sigma_1 \otimes \cdots \otimes \sigma_n \tag{2.18}$$

for $\sigma_1 \in \mathrm{D}(\mathcal{Y}_1)$, $\ldots$, $\sigma_n \in \mathrm{D}(\mathcal{Y}_n)$ being states of $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$, respectively. Product states represent independence among the states of registers, and when the compound register $\mathsf{X} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$ is in a product state ρ of the form (2.18), the registers $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$ are said to be *independent*. When it is not the case that $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$ are independent, they are said to be *correlated*.

Example 2.6 Consider a compound register of the form $\mathsf{X} = (\mathsf{Y}, \mathsf{Z})$, for Y and Z being registers sharing the classical state set $\{0, 1\}$. (Registers having the classical state set $\{0, 1\}$ are typically called *qubits*, which is short for *quantum bits*.)

[1] The other basic notions of quantum information to be discussed in this chapter have a similar character of admitting analogous probabilistic notions as special cases. In general, the theory of quantum information may be seen as an extension of classical information theory, including the study of random processes, protocols, and computations.

The state $\rho \in \mathrm{D}(\mathcal{Y} \otimes \mathcal{Z})$ defined as

$$\rho = \frac{1}{4}E_{0,0} \otimes E_{0,0} + \frac{1}{4}E_{0,0} \otimes E_{1,1} + \frac{1}{4}E_{1,1} \otimes E_{0,0} + \frac{1}{4}E_{1,1} \otimes E_{1,1} \tag{2.19}$$

is an example of a product state, as one may write

$$\rho = \left(\frac{1}{2}E_{0,0} + \frac{1}{2}E_{1,1}\right) \otimes \left(\frac{1}{2}E_{0,0} + \frac{1}{2}E_{1,1}\right). \tag{2.20}$$

Equivalently, in matrix form, one has

$$\rho = \begin{pmatrix} \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \otimes \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}. \tag{2.21}$$

The states $\sigma, \tau \in \mathrm{D}(\mathcal{Y} \otimes \mathcal{Z})$ defined as

$$\sigma = \frac{1}{2}E_{0,0} \otimes E_{0,0} + \frac{1}{2}E_{1,1} \otimes E_{1,1} \tag{2.22}$$

and

$$\tau = \frac{1}{2}E_{0,0} \otimes E_{0,0} + \frac{1}{2}E_{0,1} \otimes E_{0,1} + \frac{1}{2}E_{1,0} \otimes E_{1,0} + \frac{1}{2}E_{1,1} \otimes E_{1,1} \tag{2.23}$$

are examples of states that are not product states, as they cannot be written as tensor products, and therefore represent correlations between the registers Y and Z. In matrix form, these states are as follows:

$$\sigma = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad \tau = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}. \tag{2.24}$$

The states ρ and σ are diagonal, so they correspond to probabilistic states; ρ represents the situation in which Y and Z store independent random bits, while σ represents the situation in which Y and Z store perfectly correlated random bits. The state τ does not represent a probabilistic state, and more specifically is an example of an *entangled* state. Entanglement is a particular type of correlation having great significance in quantum information theory, and is the primary focus of Chapter 6.

Bases of density operators

It is an elementary fact, but nevertheless a useful one, that for every complex Euclidean space $\mathcal{X}$ there exist spanning sets of the space $\mathrm{L}(\mathcal{X})$ consisting only of density operators. One implication of this fact is that every linear mapping of the form

$$\phi : \mathrm{L}(\mathcal{X}) \to \mathbb{C} \tag{2.25}$$

is uniquely determined by its action on the elements of $\mathrm{D}(\mathcal{X})$. This implies, for instance, that channels and measurements are uniquely determined by their actions on density operators. The following example describes one way of constructing such a spanning set.

Example 2.7 Let Σ be an alphabet, and assume that a total ordering has been defined on Σ. For every pair $(a, b) \in \Sigma \times \Sigma$, define a density operator $\rho_{a,b} \in \mathrm{D}(\mathbb{C}^{\Sigma})$ as follows:

$$\rho_{a,b} = \begin{cases} E_{a,a} & \text{if } a = b \\ \frac{1}{2}(e_a + e_b)(e_a + e_b)^* & \text{if } a < b \\ \frac{1}{2}(e_a + ie_b)(e_a + ie_b)^* & \text{if } a > b. \end{cases} \tag{2.26}$$

For each pair $(a, b) \in \Sigma \times \Sigma$ with $a < b$, one has

$$\begin{aligned} \left(\rho_{a,b} - \frac{1}{2}\rho_{a,a} - \frac{1}{2}\rho_{b,b}\right) - i\left(\rho_{b,a} - \frac{1}{2}\rho_{a,a} - \frac{1}{2}\rho_{b,b}\right) &= E_{a,b}, \\ \left(\rho_{a,b} - \frac{1}{2}\rho_{a,a} - \frac{1}{2}\rho_{b,b}\right) + i\left(\rho_{b,a} - \frac{1}{2}\rho_{a,a} - \frac{1}{2}\rho_{b,b}\right) &= E_{b,a}, \end{aligned} \tag{2.27}$$

and from these equations it follows that

$$\operatorname{span}\{\rho_{a,b} : (a, b) \in \Sigma \times \Sigma\} = \mathrm{L}(\mathbb{C}^{\Sigma}). \tag{2.28}$$

2.1.3 Reductions and purifications of quantum states

One may consider a register obtained by removing one or more subregisters from a given compound register. The quantum state of any register that results from this process, viewed in isolation from the subregisters that were removed, is uniquely determined by the state of the original compound register. This section explains how such states are determined. The special case in which the original compound register is in a pure state is particularly important, and is discussed in detail.

The partial trace and reductions of quantum states

Let $\mathsf{X} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$ be a compound register, for $n \geq 2$. For any choice of $k \in \{1, \ldots, n\}$, one may form a new register

$$(\mathsf{Y}_1, \ldots, \mathsf{Y}_{k-1}, \mathsf{Y}_{k+1}, \ldots, \mathsf{Y}_n) \tag{2.29}$$

by removing the register Y_k from X and leaving the remaining registers untouched. For every state $\rho \in \mathrm{D}(\mathcal{X})$ of X, the state of the register (2.29) that is determined by this process is called the *reduction* of ρ to the register (2.29), and is denoted $\rho[\mathsf{Y}_1, \ldots, \mathsf{Y}_{k-1}, \mathsf{Y}_{k+1}, \ldots, \mathsf{Y}_n]$. This state is defined as

$$\rho[\mathsf{Y}_1, \ldots, \mathsf{Y}_{k-1}, \mathsf{Y}_{k+1}, \ldots, \mathsf{Y}_n] = \mathrm{Tr}_{\mathcal{Y}_k}(\rho), \tag{2.30}$$

where

$$\mathrm{Tr}_{\mathcal{Y}_k} \in \mathrm{T}(\mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n, \mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_{k-1} \otimes \mathcal{Y}_{k+1} \otimes \cdots \otimes \mathcal{Y}_n) \tag{2.31}$$

denotes the *partial trace* mapping (q.v. Section 1.1.2).[2] This is the unique linear mapping that satisfies the equation

$$\mathrm{Tr}_{\mathcal{Y}_k}(Y_1 \otimes \cdots \otimes Y_n) = \mathrm{Tr}(Y_k)\, Y_1 \otimes \cdots \otimes Y_{k-1} \otimes Y_{k+1} \otimes \cdots \otimes Y_n \tag{2.32}$$

for all operators $Y_1 \in \mathrm{L}(\mathcal{Y}_1), \ldots, Y_n \in \mathrm{L}(\mathcal{Y}_n)$. Alternately, one may define

$$\mathrm{Tr}_{\mathcal{Y}_k} = \mathbb{1}_{\mathrm{L}(\mathcal{Y}_1)} \otimes \cdots \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y}_{k-1})} \otimes \mathrm{Tr} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y}_{k+1})} \otimes \cdots \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y}_n)}, \tag{2.33}$$

where it is to be understood that the trace mapping on the right-hand side of this equation acts on $\mathrm{L}(\mathcal{Y}_k)$.

If the classical state sets of $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$ are $\Gamma_1, \ldots, \Gamma_n$, respectively, one may write the $((a_1, \ldots, a_{k-1}, a_{k+1}, \ldots, a_n), (b_1, \ldots, b_{k-1}, b_{k+1}, \ldots, b_n))$ entry of the state $\rho[\mathsf{Y}_1, \ldots, \mathsf{Y}_{k-1}, \mathsf{Y}_{k+1}, \ldots, \mathsf{Y}_n]$ explicitly as

$$\sum_{c \in \Gamma_k} \rho\big((a_1, \ldots, a_{k-1}, c, a_{k+1}, \ldots, a_n), (b_1, \ldots, b_{k-1}, c, b_{k+1}, \ldots, b_n)\big) \tag{2.34}$$

for each choice of $a_j, b_j \in \Gamma_j$ and j ranging over the set $\{1, \ldots, n\} \backslash \{k\}$.

Example 2.8 Let Y and Z be registers, both having the classical state set Σ, let $\mathsf{X} = (\mathsf{Y}, \mathsf{Z})$, and let $u \in \mathcal{X} = \mathcal{Y} \otimes \mathcal{Z}$ be defined as

$$u = \frac{1}{\sqrt{|\Sigma|}} \sum_{a \in \Sigma} e_a \otimes e_a, \tag{2.35}$$

so that

$$uu^* = \frac{1}{|\Sigma|} \sum_{a,b \in \Sigma} E_{a,b} \otimes E_{a,b}. \tag{2.36}$$

It holds that

$$(uu^*)[\mathsf{Y}] = \frac{1}{|\Sigma|} \sum_{a,b \in \Sigma} \mathrm{Tr}(E_{a,b}) E_{a,b} = \frac{1}{|\Sigma|} \mathbb{1}_{\mathcal{Y}}. \tag{2.37}$$

[2] It should be noted that reductions of states are determined in this way, by means of the partial trace, by necessity—no other choice is consistent with the basic notions concerning channels and measurements to be discussed in the sections following this one.

The state uu^* is the canonical example of a *maximally entangled* state of two registers sharing the classical state set Σ.

By applying this definition iteratively, one finds that each state ρ of the register $(\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$ uniquely determines the state of

$$(\mathsf{Y}_{k_1}, \ldots, \mathsf{Y}_{k_m}), \tag{2.38}$$

for $k_1, \ldots, k_m$ being any choice of indices satisfying $1 \leq k_1 < \cdots < k_m \leq n$. The state determined by this process is denoted $\rho[\mathsf{Y}_{k_1}, \ldots, \mathsf{Y}_{k_m}]$ and again is called the reduction of ρ to $(\mathsf{Y}_{k_1}, \ldots, \mathsf{Y}_{k_m})$.

The definition above may be generalized in a natural way so that it allows one to specify the states that result from removing an arbitrary collection of subregisters from a given compound register, assuming that this removal results in a valid register. For the registers described in Example 2.2, for instance, removing the subregister Z_3 from X while it is in the state ρ would leave the resulting register in the state

$$(\mathbb{1}_{\mathrm{L}(\mathcal{Y}_0)} \otimes (\mathbb{1}_{\mathrm{L}(\mathcal{Z}_1)} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z}_2)} \otimes \mathrm{Tr}\,))(\rho), \tag{2.39}$$

with the understanding that the trace mapping is defined with respect to $\mathcal{Z}_3$. The pattern represented by this example, in which identity mappings and trace mappings are tensored in accordance with the structure of the register under consideration, is generalized in the most straightforward way to other examples. While it is possible to formalize this definition in complete generality, there is little point in doing so for the purposes of this book: all of the instances of state reductions to be encountered are either cases where the reductions take the form $\rho[\mathsf{Y}_{k_1}, \ldots, \mathsf{Y}_{k_m}]$, as discussed above, or are easily specified explicitly as in the case of the example (2.39) just mentioned.

Purifications of states and operators

In a variety of situations that arise in quantum information theory, wherein a given register X is being considered, it is useful to assume (or simply to imagine) that X is a subregister of a compound register (X, Y), and to view a given state $\rho \in \mathrm{D}(\mathcal{X})$ of X as having been obtained as a reduction

$$\rho = \sigma[\mathsf{X}] = \mathrm{Tr}_{\mathcal{Y}}(\sigma) \tag{2.40}$$

of some state σ of (X, Y). Such a state σ is called an *extension* of ρ. It is particularly useful to consider the case in which σ is a pure state, and to ask what the possible states of X are that can arise from a pure state of (X, Y) in this way. This question has a simple answer to be justified shortly: a state $\rho \in \mathrm{D}(\mathcal{X})$ of X can arise in this way if and only if the rank of ρ does not

exceed the number of classical states of the register Y removed from (X, Y) to obtain X.

The following definition is representative of the situation just described. The notion of a *purification* that it defines is used extensively throughout the remainder of the book.

Definition 2.9 Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $P \in \mathrm{Pos}(\mathcal{X})$ be a positive semidefinite operator, and let $u \in \mathcal{X} \otimes \mathcal{Y}$ be a vector. The vector u is said to be a *purification* of P if

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = P. \tag{2.41}$$

This definition deviates slightly from the setting described above in two respects. One is that P is not required to have unit trace, and the other is that the vector u is taken to be the object that purifies P rather than the operator uu^*. Allowing P to be an arbitrary positive semidefinite operator is a useful generalization that will cause no difficulties in developing the concept of a purification (and the term *extension* is generalized in a similar way), while referring to u rather than uu^* as the purification of P is simply a matter of convenience based on the specific ways that the notion is most typically used—it is also common that the operator uu^* is the object referred to as a purification.

It is straightforward to generalize the notion of a purification. One may, for instance, consider the situation in which X is a register that is obtained by removing one or more subregisters from an arbitrary compound register Z. A purification of a given state $\rho \in \mathrm{D}(\mathcal{X})$ in this context would refer to any pure state of Z whose reduction to X is equal to ρ. The most interesting aspects of purifications are, however, represented by Definition 2.9, so the remainder of the section focuses on this specific definition of purifications for simplicity. It is to be understood, however, that the various facts concerning purifications discussed extend easily and directly to a more general notion of a purification.

Conditions for the existence of purifications

The vec mapping, defined in Section 1.1.2, is useful for understanding purifications. Given that this mapping is a linear bijection from $\mathrm{L}(\mathcal{Y}, \mathcal{X})$ to $\mathcal{X} \otimes \mathcal{Y}$, every vector $u \in \mathcal{X} \otimes \mathcal{Y}$ may be written as $u = \mathrm{vec}(A)$ for some choice of an operator $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$. By the identity (1.133), it holds that

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = \mathrm{Tr}_{\mathcal{Y}}(\mathrm{vec}(A)\,\mathrm{vec}(A)^*) = AA^*. \tag{2.42}$$

This establishes an equivalence between the following statements, for a given choice of $P \in \mathrm{Pos}(\mathcal{X})$:

1. There exists a purification $u \in \mathcal{X} \otimes \mathcal{Y}$ of P.
2. There exists an operator $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ such that $P = AA^*$.

The next theorem, whose proof is based on this observation, justifies the answer given above to the question on necessary and sufficient conditions for the existence of a purification of a given operator.

Theorem 2.10 *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, and let $P \in \mathrm{Pos}(\mathcal{X})$ be a positive semidefinite operator. There exists a vector $u \in \mathcal{X} \otimes \mathcal{Y}$ such that $\mathrm{Tr}_{\mathcal{Y}}(uu^*) = P$ if and only if $\dim(\mathcal{Y}) \geq \mathrm{rank}(P)$.*

Proof As observed above, the existence of a vector $u \in \mathcal{X} \otimes \mathcal{Y}$ for which $\mathrm{Tr}_{\mathcal{Y}}(uu^*) = P$ is equivalent to the existence of an operator $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ satisfying $P = AA^*$. Under the assumption that such an operator A exists, it must hold that $\mathrm{rank}(P) = \mathrm{rank}(A)$, and therefore $\dim(\mathcal{Y}) \geq \mathrm{rank}(P)$.

Conversely, under the assumption $\dim(\mathcal{Y}) \geq \mathrm{rank}(P)$, one may prove the existence of an operator $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ satisfying $P = AA^*$ as follows. Let $r = \mathrm{rank}(P)$ and use the spectral theorem (Corollary 1.4) to write

$$P = \sum_{k=1}^{r} \lambda_k(P) x_k x_k^* \tag{2.43}$$

for $\{x_1, \ldots, x_r\} \subset \mathcal{X}$ being an orthonormal set. For an arbitrary choice of an orthonormal set $\{y_1, \ldots, y_r\} \subset \mathcal{Y}$, which must exist by the assumption $\dim(\mathcal{Y}) \geq \mathrm{rank}(P)$, the operator

$$A = \sum_{k=1}^{r} \sqrt{\lambda_k(P)}\, x_k y_k^* \tag{2.44}$$

satisfies $AA^* = P$. □

Corollary 2.11 *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces satisfying $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. For every positive semidefinite operator $P \in \mathrm{Pos}(\mathcal{X})$, there exists a vector $u \in \mathcal{X} \otimes \mathcal{Y}$ such that $\mathrm{Tr}_{\mathcal{Y}}(uu^*) = P$.*

Unitary equivalence of purifications

Having established a simple condition under which a purification of a given positive semidefinite operator exists, it is natural to consider the possible relationships among different purifications of a given operator. The following theorem establishes a useful relationship between purifications that must always hold.

Theorem 2.12 (Unitary equivalence of purifications) *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $u, v \in \mathcal{X} \otimes \mathcal{Y}$ be vectors, and assume that*

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = \mathrm{Tr}_{\mathcal{Y}}(vv^*). \tag{2.45}$$

There exists a unitary operator $U \in \mathrm{U}(\mathcal{Y})$ such that $v = (\mathbb{1}_{\mathcal{X}} \otimes U)u$.

Proof Let $A, B \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be the unique operators satisfying $u = \mathrm{vec}(A)$ and $v = \mathrm{vec}(B)$, and let $P \in \mathrm{Pos}(\mathcal{X})$ satisfy

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = P = \mathrm{Tr}_{\mathcal{Y}}(vv^*). \tag{2.46}$$

It therefore holds that $AA^* = P = BB^*$. Letting $r = \mathrm{rank}(P)$, it follows that $\mathrm{rank}(A) = r = \mathrm{rank}(B)$.

Next, let $x_1, \ldots, x_r \in \mathcal{X}$ be any orthonormal sequence of eigenvectors of P with corresponding eigenvalues $\lambda_1(P), \ldots, \lambda_r(P)$. As $AA^* = P = BB^*$, it is possible to select singular value decompositions

$$A = \sum_{k=1}^{r} \sqrt{\lambda_k(P)}\, x_k y_k^* \quad \text{and} \quad B = \sum_{k=1}^{r} \sqrt{\lambda_k(P)}\, x_k w_k^* \tag{2.47}$$

of A and B, for some choice of orthonormal collections $\{y_1, \ldots, y_r\}$ and $\{w_1, \ldots, w_r\}$ of vectors in $\mathcal{Y}$ (as discussed in Section 1.1.3).

Finally, let $V \in \mathrm{U}(\mathcal{Y})$ be any unitary operator satisfying $V w_k = y_k$ for every $k \in \{1, \ldots, r\}$. It follows that $AV = B$, and by taking $U = V^{\mathsf{T}}$ one has

$$(\mathbb{1}_{\mathcal{X}} \otimes U)u = \left(\mathbb{1}_{\mathcal{X}} \otimes V^{\mathsf{T}}\right) \mathrm{vec}(A) = \mathrm{vec}(AV) = \mathrm{vec}(B) = v, \tag{2.48}$$

as required. □

2.2 Quantum channels

Quantum channels represent discrete changes in states of registers that are to be considered physically realizable (in an idealized sense). For example, the steps of a quantum computation, or any other processing of quantum information, as well as the effects of errors and noise on quantum registers, are modeled as quantum channels.

2.2.1 Definitions and basic notions concerning channels

In mathematical terms, a quantum channel is a linear map, from one space of square operators to another, that satisfies the two conditions of *complete positivity* and *trace preservation*.

Definition 2.13 A *quantum channel* (or simply a *channel*, for short) is a linear map

$$\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y}) \tag{2.49}$$

(i.e., an element $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$), for some choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, satisfying two properties:

1. Φ is completely positive.
2. Φ is trace preserving.

The collection of all channels of the form (2.49) is denoted $\mathrm{C}(\mathcal{X}, \mathcal{Y})$, and one writes $\mathrm{C}(\mathcal{X})$ as a shorthand for $\mathrm{C}(\mathcal{X}, \mathcal{X})$.

For a given choice of registers X and Y, one may view that a channel of the form $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ is a transformation from X into Y. That is, when such a transformation takes place, it is to be viewed that the register X ceases to exist, with Y being formed in its place. Moreover, the state of Y is obtained by applying the map Φ to the state $\rho \in \mathrm{D}(\mathcal{X})$ of X, yielding $\Phi(\rho) \in \mathrm{D}(\mathcal{Y})$. When it is the case that $\mathsf{X} = \mathsf{Y}$, one may simply view that the state of the register X has been changed according to the mapping Φ.

Example 2.14 Let $\mathcal{X}$ be a complex Euclidean space and let $U \in \mathrm{U}(\mathcal{X})$ be a unitary operator. The map $\Phi \in \mathrm{C}(\mathcal{X})$ defined by

$$\Phi(X) = UXU^* \tag{2.50}$$

for every $X \in \mathrm{L}(\mathcal{X})$ is an example of a channel. Channels of this form are called *unitary channels*. The *identity channel* $\mathbb{1}_{\mathrm{L}(\mathcal{X})}$ is one example of a unitary channel, obtained by setting $U = \mathbb{1}_{\mathcal{X}}$. Intuitively speaking, this channel represents an ideal quantum communication channel or a perfect component in a quantum computer memory, which causes no change in the state of the register X it acts upon.

Example 2.15 Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, and let $\sigma \in \mathrm{D}(\mathcal{Y})$ be a density operator. The mapping $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ defined by

$$\Phi(X) = \mathrm{Tr}(X)\sigma, \tag{2.51}$$

for every $X \in \mathrm{L}(\mathcal{X})$, is a channel. It holds that $\Phi(\rho) = \sigma$ for every $\rho \in \mathrm{D}(\mathcal{X})$; in effect, the channel Φ represents the action of discarding the register X, and replacing it with the register Y initialized to the state σ. Channels of this form will be called *replacement channels*.

The channels described in the two previous examples (along with other examples of channels) will be discussed in greater detail in Section 2.2.3. While one may prove directly that these mappings are indeed channels, these facts will follow immediately from more general results to be presented in Section 2.2.2.

Product channels

Suppose $\mathsf{X}_1, \ldots, \mathsf{X}_n$ and $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$ are registers, and recall that one denotes by $\mathcal{X}_1, \ldots, \mathcal{X}_n$ and $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$ the complex Euclidean spaces associated with these registers. A channel

$$\Phi \in \mathrm{C}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n, \mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n) \tag{2.52}$$

transforming $(\mathsf{X}_1, \ldots, \mathsf{X}_n)$ into $(\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$ is called a *product channel* if

$$\Phi = \Psi_1 \otimes \cdots \otimes \Psi_n \tag{2.53}$$

for some choice of channels $\Psi_1 \in \mathrm{C}(\mathcal{X}_1, \mathcal{Y}_1)$, $\ldots$, $\Psi_n \in \mathrm{C}(\mathcal{X}_n, \mathcal{Y}_n)$. Product channels represent an independent application of a sequence of channels to a sequence of registers, in a similar way to product states representing independence among the states of registers.

An important special case involving independent channels is the situation in which a given channel is performed on one register, while nothing at all is done to one or more other registers under consideration. (As suggested in Example 2.14, the act of doing nothing at all to a register is equivalent to performing the identity channel on that register.)

Example 2.16 Suppose that X, Y, and Z are registers, and $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ is a channel that transforms X into Y. Also suppose that the compound register (X, Z) is in some particular state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$ at some instant, and the channel Φ is applied to X, transforming it into Y. The resulting state of the pair (Y, Z) is then given by

$$(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\rho) \in \mathrm{D}(\mathcal{Y} \otimes \mathcal{Z}), \tag{2.54}$$

as one views that the identity channel $\mathbb{1}_{\mathrm{L}(\mathcal{Z})}$ has independently been applied to the register Z.

Example 2.16 illustrates the importance of the requirement that channels are complete positive. That is, it must hold that $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\rho)$ is a density operator for every choice of Z and every density operator $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$, which together with the linearity of Φ implies that Φ is completely positive (in addition to being trace preserving).

State preparations as quantum channels

As stated in Section 2.1.1, a register is *trivial* if its classical state set consists of a single element. The complex Euclidean space associated with a trivial register is therefore one-dimensional: it must take the form $\mathbb{C}^{\{a\}}$ for $\{a\}$ being the singleton classical state set of the register. No generality is lost in associating such a space with the field of complex numbers $\mathbb{C}$, and in making the identification $\mathrm{L}(\mathbb{C}) = \mathbb{C}$, one finds that the scalar 1 is the only possible state for a trivial register. As is to be expected, such a register is therefore completely useless from an information-processing viewpoint; the presence of a trivial register does nothing more than to tensor the scalar 1 to the state of any other registers under consideration.

It is instructive nevertheless to consider the properties of channels that involve trivial registers. Suppose, in particular, that X is a trivial register and Y is arbitrary, and consider a channel of the form $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ that transforms X into Y. It must hold that Φ is given by

$$\Phi(\alpha) = \alpha\rho \tag{2.55}$$

for all $\alpha \in \mathbb{C}$, for some choice of $\rho \in \mathrm{D}(\mathcal{Y})$, as Φ must be linear and it must hold that $\Phi(1)$ is positive semidefinite and has trace equal to one. The channel Φ defined by (2.55) may be viewed as the *preparation* of the quantum state ρ in a new register Y. The trivial register X can be considered as being essentially a placeholder for this preparation, which is to occur at whatever moment the channel Φ is performed. In this way, a state preparation may be seen as the application of this form of channel.

To see that every mapping of the form (2.55) is indeed a channel, for an arbitrary choice of a density operator $\rho \in \mathrm{D}(\mathcal{Y})$, one may check that the conditions of complete positivity and trace preservation hold. The mapping Φ given by (2.55) is obviously trace preserving whenever $\mathrm{Tr}(\rho) = 1$, and the complete positivity of Φ is implied by the following simple proposition.

Proposition 2.17 *Let $\mathcal{Y}$ be a complex Euclidean space and let $P \in \mathrm{Pos}(\mathcal{Y})$ be a positive semidefinite operator. The mapping $\Phi \in \mathrm{T}(\mathbb{C}, \mathcal{Y})$ defined as $\Phi(\alpha) = \alpha P$ for all $\alpha \in \mathbb{C}$ is completely positive.*

Proof Let $\mathcal{Z}$ be any complex Euclidean space. The action of the mapping $\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}$ on an operator $Z \in \mathrm{L}(\mathcal{Z}) = \mathrm{L}(\mathbb{C} \otimes \mathcal{Z})$ is given by

$$(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(Z) = P \otimes Z. \tag{2.56}$$

If Z is positive semidefinite, then $P \otimes Z$ is positive semidefinite as well, and therefore Φ is completely positive. □

The trace mapping as a channel

Another situation in which a channel Φ involves a trivial register is when this channel transforms an arbitrary register X into a trivial register Y. By identifying the complex Euclidean space $\mathcal{Y}$ with the complex numbers $\mathbb{C}$ as before, one has that the channel Φ must take the form $\Phi \in \mathrm{C}(\mathcal{X}, \mathbb{C})$.

The only mapping of this form that can possibly preserve trace is the trace mapping itself, and so it must hold that

$$\Phi(X) = \mathrm{Tr}(X) \tag{2.57}$$

for all $X \in \mathrm{L}(\mathcal{X})$. To say that a register X has been transformed into a trivial register Y is tantamount to saying that X has been destroyed, discarded, or simply ignored. This channel was, in effect, introduced in Section 2.1.3 when reductions of quantum states were defined.

In order to conclude that the trace mapping is indeed a valid channel, it is necessary to verify that it is completely positive. One way to prove this simple fact is to combine the following proposition with Proposition 2.17.

Proposition 2.18 *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a positive map, for $\mathcal{X}$ and $\mathcal{Y}$ being complex Euclidean spaces. It holds that Φ^* is positive.*

Proof By the positivity of Φ, it holds that $\Phi(P) \in \mathrm{Pos}(\mathcal{Y})$ for every positive semidefinite operator $P \in \mathrm{Pos}(\mathcal{X})$, which is equivalent to the condition that

$$\langle Q, \Phi(P) \rangle \geq 0 \tag{2.58}$$

for all $P \in \mathrm{Pos}(\mathcal{X})$ and $Q \in \mathrm{Pos}(\mathcal{Y})$. It follows that

$$\langle \Phi^*(Q), P \rangle = \langle Q, \Phi(P) \rangle \geq 0 \tag{2.59}$$

for all $P \in \mathrm{Pos}(\mathcal{X})$ and $Q \in \mathrm{Pos}(\mathcal{Y})$, which is equivalent to $\Phi^*(Q) \in \mathrm{Pos}(\mathcal{X})$ for every $Q \in \mathrm{Pos}(\mathcal{Y})$. The mapping Φ^* is therefore positive. □

Remark Proposition 2.18 implies that if $\Phi \in \mathrm{CP}(\mathcal{X}, \mathcal{Y})$ is a completely positive map, then the adjoint map Φ^* is also completely positive; for if Φ is completely positive, then $\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}$ is positive for every complex Euclidean space $\mathcal{Z}$, and therefore $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})^* = \Phi^* \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}$ is also positive.

Corollary 2.19 *The trace mapping $\mathrm{Tr} \in \mathrm{T}(\mathcal{X}, \mathbb{C})$, for any choice of a complex Euclidean space $\mathcal{X}$, is completely positive.*

Proof The adjoint of the trace is given by $\mathrm{Tr}^*(\alpha) = \alpha \mathbb{1}_{\mathcal{X}}$ for every $\alpha \in \mathbb{C}$. This map is completely positive by Proposition 2.17, therefore the trace map is completely positive by the remark above. □

2.2.2 Representations and characterizations of channels

Suppose $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ is a channel, for $\mathcal{X}$ and $\mathcal{Y}$ being complex Euclidean spaces. It may, in some situations, be sufficient to view such a channel abstractly, as a completely positive and trace-preserving linear map of the form $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$ and nothing more. In other situations, it may be useful to consider a more concrete representation of such a channel.

Four specific representations of channels (and of arbitrary maps of the form $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$) are discussed in this section. These different representations reveal interesting properties of channels, and will find uses in different situations throughout this book. The simple relationships among the representations generally allow one to convert from one representation into another, and therefore to choose the representation that is best suited to a given situation.

The natural representation

For any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, and for every linear map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, it is evident that the mapping

$$\mathrm{vec}(X) \mapsto \mathrm{vec}(\Phi(X)) \tag{2.60}$$

is linear, as it can be represented as a composition of linear mappings. There must therefore exist a linear operator $K(\Phi) \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X}, \mathcal{Y} \otimes \mathcal{Y})$ for which it holds that

$$K(\Phi)\,\mathrm{vec}(X) = \mathrm{vec}(\Phi(X)) \tag{2.61}$$

for all $X \in \mathrm{L}(\mathcal{X})$. The operator $K(\Phi)$, which is uniquely determined by the requirement that (2.61) holds for all $X \in \mathrm{L}(\mathcal{X})$, is the *natural representation* of Φ, as it directly represents the action of Φ as a linear map (with respect to the operator-vector correspondence).

It may be noted that the mapping $K : \mathrm{T}(\mathcal{X}, \mathcal{Y}) \to \mathrm{L}(\mathcal{X} \otimes \mathcal{X}, \mathcal{Y} \otimes \mathcal{Y})$ is linear:

$$K(\alpha\Phi + \beta\Psi) = \alpha K(\Phi) + \beta K(\Psi) \tag{2.62}$$

for all choices of $\alpha, \beta \in \mathbb{C}$ and $\Phi, \Psi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$. Moreover, K is a bijection, as the action of a given mapping Φ can be recovered from $K(\Phi)$; for each operator $X \in \mathrm{L}(\mathcal{X})$, one has that $Y = \Phi(X)$ is the unique operator satisfying $\mathrm{vec}(Y) = K(\Phi)\,\mathrm{vec}(X)$.

The natural representation respects the notion of adjoints, meaning that

$$K(\Phi^*) = (K(\Phi))^* \tag{2.63}$$

for every map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ (with the understanding that K refers to a

mapping from $\mathrm{T}(\mathcal{Y},\mathcal{X})$ to $\mathrm{L}(\mathcal{Y}\otimes\mathcal{Y},\mathcal{X}\otimes\mathcal{X})$ on the left-hand side of this equation, obtained by reversing the roles of $\mathcal{X}$ and $\mathcal{Y}$ in the definition above).

Despite the fact that the natural representation $K(\Phi)$ of a mapping Φ is a direct representation of the action of Φ as a linear map, this representation is the one of the four representations to be discussed in this section that is the least directly connected to the properties of complete positivity and trace preservation. As such, it will turn out to be the least useful of the four representations from the viewpoint of this book. One explanation for why this is so is that the aspects of a given map Φ that relate to the operator structure of its input and output arguments is not represented by $K(\Phi)$ in a convenient or readily accessible form. The operator-vector correspondence has the effect of ignoring this structure.

The Choi representation

For any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, one may define a mapping $J : \mathrm{T}(\mathcal{X},\mathcal{Y}) \to \mathrm{L}(\mathcal{Y}\otimes\mathcal{X})$ as

$$J(\Phi) = \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\big)\big(\mathrm{vec}(\mathbb{1}_{\mathcal{X}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*\big) \tag{2.64}$$

for each $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$. Alternatively, under the assumption that $\mathcal{X} = \mathbb{C}^{\Sigma}$, one may write

$$J(\Phi) = \sum_{a,b\in\Sigma} \Phi(E_{a,b}) \otimes E_{a,b}. \tag{2.65}$$

The operator $J(\Phi)$ is called the *Choi representation* (or the *Choi operator*) of Φ.

It is evident from the equation (2.65) that the mapping J is a linear bijection. An alternative way to prove that the mapping J is a bijection is to observe that the action of the mapping Φ can be recovered from the operator $J(\Phi)$ by means of the equation

$$\Phi(X) = \mathrm{Tr}_{\mathcal{X}}\big(J(\Phi)\big(\mathbb{1}_{\mathcal{Y}} \otimes X^{\mathsf{T}}\big)\big). \tag{2.66}$$

There is a close connection between the operator structure of $J(\Phi)$ and the aspects of Φ that relate to the operator structure of its input and output arguments. A central component of this connection is that a given map Φ is completely positive if and only if $J(\Phi)$ is positive semidefinite (as is established by Theorem 2.22 below).

For a given map $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$, the rank of its Choi representation $J(\Phi)$ is called the *Choi rank* of Φ.

Kraus representations

For any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, an alphabet Σ, and collections

$$\{A_a : a \in \Sigma\} \quad \text{and} \quad \{B_a : a \in \Sigma\} \tag{2.67}$$

of operators drawn from the space $\mathrm{L}(\mathcal{X},\mathcal{Y})$, one may define a linear map $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$ as

$$\Phi(X) = \sum_{a\in\Sigma} A_a X B_a^* \tag{2.68}$$

for every $X \in \mathrm{L}(\mathcal{X})$. The expression (2.68) is a *Kraus representation* of the map Φ. It will be established shortly that a Kraus representation exists for every map of the form $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$. Unlike the natural representation and Choi representation, however, Kraus representations are not unique.

Under the assumption that Φ is determined by the above equation (2.68), it holds that

$$\Phi^*(Y) = \sum_{a\in\Sigma} A_a^* Y B_a, \tag{2.69}$$

as follows from a calculation relying on the cyclic property of the trace:

$$\begin{aligned} \left\langle Y, \sum_{a\in\Sigma} A_a X B_a^* \right\rangle &= \sum_{a\in\Sigma} \mathrm{Tr}\big(Y^* A_a X B_a^*\big) \\ &= \sum_{a\in\Sigma} \mathrm{Tr}\big(B_a^* Y^* A_a X\big) = \left\langle \sum_{a\in\Sigma} A_a^* Y B_a, X \right\rangle \end{aligned} \tag{2.70}$$

for every $X \in \mathrm{L}(\mathcal{X})$ and $Y \in \mathrm{L}(\mathcal{Y})$.

It is common in the theory of quantum information that one encounters Kraus representations for which $A_a = B_a$ for each $a \in \Sigma$. As is established by Theorem 2.22 below, such representations exist precisely when the map being considered is completely positive.

Stinespring representations

Suppose $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are complex Euclidean spaces and $A, B \in \mathrm{L}(\mathcal{X},\mathcal{Y}\otimes\mathcal{Z})$ are operators. One may then define a map $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$ as

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(AXB^*\big) \tag{2.71}$$

for every $X \in \mathrm{L}(\mathcal{X})$. The expression (2.71) is a *Stinespring representation* of the map Φ. Similar to Kraus representations, Stinespring representations always exist for a given map Φ, and are not unique.

If a map $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$ has a Stinespring representation taking the form (2.71), then it holds that

$$\Phi^*(Y) = A^*(Y \otimes \mathbb{1}_{\mathcal{Z}})B \tag{2.72}$$

for all $Y \in \mathrm{L}(\mathcal{Y})$. This observation follows from a calculation:

$$\begin{aligned} \langle Y, \Phi(X)\rangle &= \langle Y, \mathrm{Tr}_{\mathcal{Z}}(AXB^*)\rangle = \langle Y \otimes \mathbb{1}_{\mathcal{Z}}, AXB^*\rangle \\ &= \mathrm{Tr}\big((Y \otimes \mathbb{1}_{\mathcal{Z}})^* AXB^*\big) = \mathrm{Tr}\big(B^*(Y \otimes \mathbb{1}_{\mathcal{Z}})^* AX\big) \\ &= \langle A^*(Y \otimes \mathbb{1}_{\mathcal{Z}})B, X\rangle \end{aligned} \tag{2.73}$$

for every $X \in \mathrm{L}(\mathcal{X})$ and $Y \in \mathrm{L}(\mathcal{Y})$. Expressions of the form (2.72) are also sometimes referred to as Stinespring representations (in this case of the map Φ^*), although the terminology will not be used in this way in this book.

Similar to Kraus representations, it is common in quantum information theory that one encounters Stinespring representations for which $A = B$. Also similar to Kraus representations, such representations exist if and only if Φ is completely positive.

Relationships among the representations

The following proposition relates the four representations discussed above to one another, and (implicitly) shows how any one of the representations may be converted into any other.

Proposition 2.20 *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let Σ be an alphabet, let $\{A_a : a \in \Sigma\}, \{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X},\mathcal{Y})$ be collections of operators indexed by Σ, and let $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$. The following four statements, which correspond as indicated to the four representations introduced above, are equivalent:*

1. *(Natural representation.) It holds that*
$$K(\Phi) = \sum_{a\in\Sigma} A_a \otimes \overline{B_a}. \tag{2.74}$$
2. *(Choi representation.) It holds that*
$$J(\Phi) = \sum_{a\in\Sigma} \mathrm{vec}(A_a)\,\mathrm{vec}(B_a)^*. \tag{2.75}$$
3. *(Kraus representations.) It holds that*
$$\Phi(X) = \sum_{a\in\Sigma} A_a X B_a^* \tag{2.76}$$
for all $X \in \mathrm{L}(\mathcal{X})$.

4. *(Stinespring representations.) For $\mathcal{Z} = \mathbb{C}^{\Sigma}$ and $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y}\otimes\mathcal{Z})$ defined as*
$$A = \sum_{a\in\Sigma} A_a \otimes e_a \quad \text{and} \quad B = \sum_{a\in\Sigma} B_a \otimes e_a, \tag{2.77}$$
it holds that
$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(AXB^*\big) \tag{2.78}$$
for all $X \in \mathrm{L}(\mathcal{X})$.

Proof The equivalence between statements 3 and 4 is a straightforward calculation. The equivalence between statements 1 and 3 follows from the identity

$$\mathrm{vec}(A_a X B_a^*) = \big(A_a \otimes \overline{B_a}\big)\,\mathrm{vec}(X) \tag{2.79}$$

for all choices of $a \in \Sigma$ and $X \in \mathrm{L}(\mathcal{X})$. Finally, the equivalence between statements 2 and 3 follows from the equations

$$\begin{aligned} (A_a \otimes \mathbb{1}_{\mathcal{X}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{X}}) &= \mathrm{vec}(A_a), \\ \mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*(B_a^* \otimes \mathbb{1}_{\mathcal{X}}) &= \mathrm{vec}(B_a)^*, \end{aligned} \tag{2.80}$$

which hold for every $a \in \Sigma$. □

Corollary 2.21 *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$ be a nonzero linear map, and let $r = \mathrm{rank}(J(\Phi))$ be the Choi rank of Φ. The following two facts hold:*

1. *For Σ being any alphabet with $|\Sigma| = r$, there exists a Kraus representation of Φ having the form*
$$\Phi(X) = \sum_{a\in\Sigma} A_a X B_a^*, \tag{2.81}$$
for some choice of $\{A_a : a \in \Sigma\}, \{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X},\mathcal{Y})$.
2. *For $\mathcal{Z}$ being any complex Euclidean space with $\dim(\mathcal{Z}) = r$, there exists a Stinespring representation of Φ having the form*
$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(AXB^*\big), \tag{2.82}$$
for some choice of operators $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y}\otimes\mathcal{Z})$.

Proof For Σ being any alphabet with $|\Sigma| = r$, it is possible to write

$$J(\Phi) = \sum_{a\in\Sigma} u_a v_a^* \tag{2.83}$$

for some choice of vectors

$$\{u_a : a \in \Sigma\}, \{v_a : a \in \Sigma\} \subset \mathcal{Y} \otimes \mathcal{X}. \tag{2.84}$$

In particular, one may take $\{u_a : a \in \Sigma\}$ to be any basis for the image of $J(\Phi)$, which uniquely determines a collection $\{v_a : a \in \Sigma\}$ for which (2.83) holds. Taking $\{A_a : a \in \Sigma\}$ and $\{B_a : a \in \Sigma\}$ to be operators defined by the equations

$$\operatorname{vec}(A_a) = u_a \quad \text{and} \quad \operatorname{vec}(B_a) = v_a \tag{2.85}$$

for every $a \in \Sigma$, it follows from Proposition 2.20 that the expression (2.81) is a Kraus representation of Φ. Moreover, it holds that the expression (2.82) is a Stinespring representation of Φ for $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ defined as

$$A = \sum_{a \in \Sigma} A_a \otimes e_a \quad \text{and} \quad B = \sum_{a \in \Sigma} B_a \otimes e_a, \tag{2.86}$$

which completes the proof. □

Characterizations of completely positive maps

Characterizations of completely positive maps, based on their Choi, Kraus, and Stinespring representations, will now be presented.

Theorem 2.22 *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a nonzero map, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. The following statements are equivalent:*

1. *Φ is completely positive.*
2. *$\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}$ is positive.*
3. *$J(\Phi) \in \operatorname{Pos}(\mathcal{Y} \otimes \mathcal{X})$.*
4. *There exists a collection $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Y})$, for some choice of an alphabet Σ, for which*

$$\Phi(X) = \sum_{a \in \Sigma} A_a X A_a^* \tag{2.87}$$

   for all $X \in \mathrm{L}(\mathcal{X})$.
5. *Statement 4 holds for an alphabet Σ satisfying $|\Sigma| = \operatorname{rank}(J(\Phi))$.*
6. *There exists an operator $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$, for some choice of a complex Euclidean space $\mathcal{Z}$, such that*

$$\Phi(X) = \operatorname{Tr}_{\mathcal{Z}}(AXA^*) \tag{2.88}$$

   for all $X \in \mathrm{L}(\mathcal{X})$.
7. *Statement 6 holds for $\mathcal{Z}$ having dimension equal to $\operatorname{rank}(J(\Phi))$.*

Proof The theorem will be proved by establishing the following implications among the seven statements, which are sufficient to imply their equivalence:

$$(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (5) \Rightarrow (4) \Rightarrow (1)$$
$$(5) \Rightarrow (7) \Rightarrow (6) \Rightarrow (1)$$

Note that some of these implications are immediate: statement 1 implies statement 2 by the definition of complete positivity, statement 5 trivially implies statement 4, statement 7 trivially implies statement 6, and statement 5 implies statement 7 by Proposition 2.20.

Assume $\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}$ is positive. Because

$$\operatorname{vec}(\mathbb{1}_{\mathcal{X}}) \operatorname{vec}(\mathbb{1}_{\mathcal{X}})^* \in \operatorname{Pos}(\mathcal{X} \otimes \mathcal{X}) \tag{2.89}$$

and

$$J(\Phi) = (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\operatorname{vec}(\mathbb{1}_{\mathcal{X}}) \operatorname{vec}(\mathbb{1}_{\mathcal{X}})^*), \tag{2.90}$$

it follows that $J(\Phi) \in \operatorname{Pos}(\mathcal{Y} \otimes \mathcal{X})$, so statement 2 implies statement 3.

Next, assume $J(\Phi) \in \operatorname{Pos}(\mathcal{Y} \otimes \mathcal{X})$. It follows by the spectral theorem (Corollary 1.4), together with the fact that every eigenvalue of a positive semidefinite operator is nonnegative, that one may write

$$J(\Phi) = \sum_{a \in \Sigma} u_a u_a^*, \tag{2.91}$$

for some choice of an alphabet Σ with $|\Sigma| = \operatorname{rank}(J(\Phi))$ and a collection

$$\{u_a : a \in \Sigma\} \subset \mathcal{Y} \otimes \mathcal{X} \tag{2.92}$$

of vectors. Taking $A_a \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ to be the operator defined by the equation $\operatorname{vec}(A_a) = u_a$ for each $a \in \Sigma$, one has that

$$J(\Phi) = \sum_{a \in \Sigma} \operatorname{vec}(A_a) \operatorname{vec}(A_a)^*. \tag{2.93}$$

The equation (2.87) therefore holds for every $X \in \mathrm{L}(\mathcal{X})$ by Proposition 2.20, which establishes that statement 3 implies statement 5.

Now suppose (2.87) holds for every $X \in \mathrm{L}(\mathcal{X})$, for some alphabet Σ and a collection

$$\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Y}) \tag{2.94}$$

of operators. For a complex Euclidean space $\mathcal{W}$ and a positive semidefinite operator $P \in \operatorname{Pos}(\mathcal{X} \otimes \mathcal{W})$, it is evident that

$$(A_a \otimes \mathbb{1}_{\mathcal{W}}) P (A_a \otimes \mathbb{1}_{\mathcal{W}})^* \in \operatorname{Pos}(\mathcal{Y} \otimes \mathcal{W}) \tag{2.95}$$

for each $a \in \Sigma$, and therefore

$$\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(P) \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{W}) \tag{2.96}$$

by the fact that $\mathrm{Pos}(\mathcal{Y} \otimes \mathcal{W})$ is a convex cone. It follows that Φ is completely positive, so statement 4 implies statement 1.

Finally, suppose (2.88) holds for every $X \in \mathrm{L}(\mathcal{X})$, for some complex Euclidean space $\mathcal{Z}$ and an operator $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$. For any complex Euclidean space $\mathcal{W}$ and any positive semidefinite operator $P \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{W})$, it is again evident that

$$(A \otimes \mathbb{1}_{\mathcal{W}})P(A \otimes \mathbb{1}_{\mathcal{W}})^* \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{Z} \otimes \mathcal{W}), \tag{2.97}$$

so that

$$\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(P) = \mathrm{Tr}_{\mathcal{Z}}\big((A \otimes \mathbb{1}_{\mathcal{W}})P(A \otimes \mathbb{1}_{\mathcal{W}})^*\big) \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{W}) \tag{2.98}$$

by the complete positivity of the trace (Corollary 2.19). It therefore holds that the map Φ is completely positive, so statement 6 implies statement 1, which completes the proof. □

One consequence of this theorem is the following corollary, which relates Kraus representations of a given completely positive map.

Corollary 2.23 *Let Σ be an alphabet, let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, and assume $\{A_a : a \in \Sigma\}, \{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Y})$ are collections of operators for which*

$$\sum_{a \in \Sigma} A_a X A_a^* = \sum_{a \in \Sigma} B_a X B_a^* \tag{2.99}$$

for all $X \in \mathrm{L}(\mathcal{X})$. There exists a unitary operator $U \in \mathrm{U}\big(\mathbb{C}^{\Sigma}\big)$ such that

$$B_a = \sum_{b \in \Sigma} U(a,b) A_b \tag{2.100}$$

for all $a \in \Sigma$.

Proof The maps

$$X \mapsto \sum_{a \in \Sigma} A_a X A_a^* \quad \text{and} \quad X \mapsto \sum_{a \in \Sigma} B_a X B_a^* \tag{2.101}$$

agree for all $X \in \mathrm{L}(\mathcal{X})$, and therefore their Choi representations must be equal:

$$\sum_{a \in \Sigma} \mathrm{vec}(A_a)\,\mathrm{vec}(A_a)^* = \sum_{a \in \Sigma} \mathrm{vec}(B_a)\,\mathrm{vec}(B_a)^*. \tag{2.102}$$

Let $\mathcal{Z} = \mathbb{C}^{\Sigma}$ and define vectors $u, v \in \mathcal{Y} \otimes \mathcal{X} \otimes \mathcal{Z}$ as

$$u = \sum_{a \in \Sigma} \mathrm{vec}(A_a) \otimes e_a \quad \text{and} \quad v = \sum_{a \in \Sigma} \mathrm{vec}(B_a) \otimes e_a, \tag{2.103}$$

so that

$$\begin{aligned} \mathrm{Tr}_{\mathcal{Z}}(uu^*) &= \sum_{a \in \Sigma} \mathrm{vec}(A_a)\,\mathrm{vec}(A_a)^* \\ &= \sum_{a \in \Sigma} \mathrm{vec}(B_a)\,\mathrm{vec}(B_a)^* = \mathrm{Tr}_{\mathcal{Z}}(vv^*). \end{aligned} \tag{2.104}$$

By the unitary equivalence of purifications (Theorem 2.12), there must exist a unitary operator $U \in \mathrm{U}(\mathcal{Z})$ such that

$$v = (\mathbb{1}_{\mathcal{Y} \otimes \mathcal{X}} \otimes U)u. \tag{2.105}$$

Thus, for each $a \in \Sigma$ it holds that

$$\mathrm{vec}(B_a) = (\mathbb{1}_{\mathcal{Y} \otimes \mathcal{X}} \otimes e_a^*)v = (\mathbb{1}_{\mathcal{Y} \otimes \mathcal{X}} \otimes e_a^* U)u = \sum_{b \in \Sigma} U(a,b)\,\mathrm{vec}(A_b), \tag{2.106}$$

which is equivalent to (2.100). □

Along similar lines to the previous corollary is the following one, which concerns Stinespring representations rather than Kraus representations. As the proof reveals, the two corollaries are essentially equivalent.

Corollary 2.24 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be complex Euclidean spaces and let operators $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ satisfy the equation*

$$\mathrm{Tr}_{\mathcal{Z}}(AXA^*) = \mathrm{Tr}_{\mathcal{Z}}(BXB^*) \tag{2.107}$$

for every $X \in \mathrm{L}(\mathcal{X})$. There exists a unitary operator $U \in \mathrm{U}(\mathcal{Z})$ such that

$$B = (\mathbb{1}_{\mathcal{Y}} \otimes U)A. \tag{2.108}$$

Proof Let Σ be the alphabet for which $\mathcal{Z} = \mathbb{C}^{\Sigma}$, and define two collections $\{A_a : a \in \Sigma\}, \{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Y})$ of operators as

$$A_a = (\mathbb{1}_{\mathcal{Y}} \otimes e_a^*)A \quad \text{and} \quad B_a = (\mathbb{1}_{\mathcal{Y}} \otimes e_a^*)B, \tag{2.109}$$

for each $a \in \Sigma$, so that

$$A = \sum_{a \in \Sigma} A_a \otimes e_a \quad \text{and} \quad B = \sum_{a \in \Sigma} B_a \otimes e_a. \tag{2.110}$$

The equation (2.107) is equivalent to (2.99) in Corollary 2.23. It follows from that corollary that there exists a unitary operator $U \in \mathrm{U}(\mathcal{Z})$ such that (2.100) holds, which is equivalent to $B = (\mathbb{1}_{\mathcal{Y}} \otimes U)A$. □

A map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is said to be *Hermitian preserving* if it holds that $\Phi(H) \in \mathrm{Herm}(\mathcal{Y})$ for all $H \in \mathrm{Herm}(\mathcal{X})$. The following theorem, which provides four alternative characterizations of this class of maps, is proved through the use of Theorem 2.22.

Theorem 2.25 *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. The following statements are equivalent:*

1. *Φ is Hermitian preserving.*
2. *It holds that $(\Phi(X))^* = \Phi(X^*)$ for every $X \in \mathrm{L}(\mathcal{X})$.*
3. *It holds that $J(\Phi) \in \mathrm{Herm}(\mathcal{Y} \otimes \mathcal{X})$.*
4. *There exist completely positive maps $\Phi_0, \Phi_1 \in \mathrm{CP}(\mathcal{X}, \mathcal{Y})$ for which $\Phi = \Phi_0 - \Phi_1$.*
5. *There exist positive maps $\Phi_0, \Phi_1 \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ for which $\Phi = \Phi_0 - \Phi_1$.*

Proof Assume first that Φ is a Hermitian-preserving map. For an arbitrary operator $X \in \mathrm{L}(\mathcal{X})$, one may write $X = H + iK$ for $H, K \in \mathrm{Herm}(\mathcal{X})$ being defined as

$$H = \frac{X + X^*}{2} \quad \text{and} \quad K = \frac{X - X^*}{2i}. \tag{2.111}$$

As $\Phi(H)$ and $\Phi(K)$ are both Hermitian and Φ is linear, it follows that

$$\begin{aligned} (\Phi(X))^* &= (\Phi(H) + i\Phi(K))^* \\ &= \Phi(H) - i\Phi(K) = \Phi(H - iK) = \Phi(X^*). \end{aligned} \tag{2.112}$$

Statement 1 therefore implies statement 2.

Next, assume statement 2 holds, and let Σ be the alphabet for which $\mathcal{X} = \mathbb{C}^{\Sigma}$. One then has that

$$\begin{aligned} J(\Phi)^* = \sum_{a,b \in \Sigma} \Phi(E_{a,b})^* \otimes E_{a,b}^* &= \sum_{a,b \in \Sigma} \Phi(E_{a,b}^*) \otimes E_{a,b}^* \\ &= \sum_{a,b \in \Sigma} \Phi(E_{b,a}) \otimes E_{b,a} = J(\Phi). \end{aligned} \tag{2.113}$$

It follows that $J(\Phi)$ is Hermitian, and therefore statement 3 holds.

Now assume statement 3 holds. Let $J(\Phi) = P_0 - P_1$ be the Jordan–Hahn decomposition of $J(\Phi)$, and let $\Phi_0, \Phi_1 \in \mathrm{CP}(\mathcal{X}, \mathcal{Y})$ be the maps for which $J(\Phi_0) = P_0$ and $J(\Phi_1) = P_1$. Because P_0 and P_1 are positive semidefinite, it follows from Theorem 2.22 that Φ_0 and Φ_1 are completely positive maps. By the linearity of the mapping J associated with the Choi representation, it holds that $J(\Phi) = J(\Phi_0 - \Phi_1)$, and therefore $\Phi = \Phi_0 - \Phi_1$, implying that statement 4 holds.

Statement 4 trivially implies statement 5.

Finally, assume statement 5 holds. Let $H \in \mathrm{Herm}(\mathcal{X})$ be a Hermitian operator, and let $H = P_0 - P_1$, for $P_0, P_1 \in \mathrm{Pos}(\mathcal{X})$, be the Jordan–Hahn decomposition of H. It holds that $\Phi_a(P_b) \in \mathrm{Pos}(\mathcal{Y})$, for all $a, b \in \{0, 1\}$, by the positivity of Φ_0 and Φ_1. Therefore, one has that

$$\Phi(H) = \big(\Phi_0(P_0) + \Phi_1(P_1)\big) - \big(\Phi_0(P_1) + \Phi_1(P_0)\big) \tag{2.114}$$

is the difference between two positive semidefinite operators, and is therefore Hermitian. Thus, statement 1 holds.

As the implications (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (1) among the statements have been established, the theorem is proved. □

Characterizations of trace-preserving maps

The next theorem provides multiple characterizations of the class of trace-preserving maps.

Theorem 2.26 *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. The following statements are equivalent:*

1. *Φ is a trace-preserving map.*
2. *Φ^* is a unital map.*
3. *$\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_{\mathcal{X}}$.*
4. *There exist collections $\{A_a : a \in \Sigma\}$, $\{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Y})$ of operators such that*

$$\Phi(X) = \sum_{a \in \Sigma} A_a X B_a^* \tag{2.115}$$

and

$$\sum_{a \in \Sigma} A_a^* B_a = \mathbb{1}_{\mathcal{X}}. \tag{2.116}$$

5. *For all collections $\{A_a : a \in \Sigma\}$, $\{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Y})$ of operators satisfying (2.115), the equation (2.116) must also hold.*
6. *There exist operators $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$, for some complex Euclidean space $\mathcal{Z}$, such that*

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}(AXB^*) \tag{2.117}$$

and $A^*B = \mathbb{1}_{\mathcal{X}}$.

7. *For every choice of operators $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ satisfying (2.117), it holds that $A^*B = \mathbb{1}_{\mathcal{X}}$.*

Proof Under the assumption that Φ preserves trace, it holds that

$$\langle \mathbb{1}_{\mathcal{X}}, X \rangle = \mathrm{Tr}(X) = \mathrm{Tr}(\Phi(X)) = \langle \mathbb{1}_{\mathcal{Y}}, \Phi(X) \rangle = \langle \Phi^*(\mathbb{1}_{\mathcal{Y}}), X \rangle, \tag{2.118}$$

and therefore

$$\langle \mathbb{1}_{\mathcal{X}} - \Phi^*(\mathbb{1}_{\mathcal{Y}}), X \rangle = 0, \tag{2.119}$$

for all $X \in \mathrm{L}(\mathcal{X})$. It follows that $\Phi^*(\mathbb{1}_{\mathcal{Y}}) = \mathbb{1}_{\mathcal{X}}$, and therefore Φ^* is unital. Along similar lines, the assumption that Φ^* is unital implies

$$\mathrm{Tr}(\Phi(X)) = \langle \mathbb{1}_{\mathcal{Y}}, \Phi(X) \rangle = \langle \Phi^*(\mathbb{1}_{\mathcal{Y}}), X \rangle = \langle \mathbb{1}_{\mathcal{X}}, X \rangle = \mathrm{Tr}(X) \tag{2.120}$$

for every $X \in \mathrm{L}(\mathcal{X})$, and therefore Φ preserves trace. The equivalence of statements 1 and 2 has been established.

Next, assume that $\{A_a : a \in \Sigma\}$, $\{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Y})$ satisfy

$$\Phi(X) = \sum_{a \in \Sigma} A_a X B_a^* \tag{2.121}$$

for all $X \in \mathrm{L}(\mathcal{X})$. It therefore holds that

$$\Phi^*(Y) = \sum_{a \in \Sigma} A_a^* Y B_a \tag{2.122}$$

for every $Y \in \mathrm{L}(\mathcal{Y})$, and in particular it holds that

$$\Phi^*(\mathbb{1}_{\mathcal{Y}}) = \sum_{a \in \Sigma} A_a^* B_a. \tag{2.123}$$

Thus, if Φ^* is a unital map, then

$$\sum_{a \in \Sigma} A_a^* B_a = \mathbb{1}_{\mathcal{X}}, \tag{2.124}$$

and so it has been proved that statement 2 implies statement 5. On the other hand, if (2.124) holds, then it follows that $\Phi^*(\mathbb{1}_{\mathcal{Y}}) = \mathbb{1}_{\mathcal{X}}$, so that Φ^* is unital. Therefore, statement 4 implies statement 2. As statement 5 implies statement 4, by virtue of the fact that Kraus representations exist for every map, the equivalence of statements 2, 4, and 5 has been established.

Now assume that $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ satisfy $\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}(AXB^*)$ for every $X \in \mathrm{L}(\mathcal{X})$. It follows that

$$\Phi^*(Y) = A^*(Y \otimes \mathbb{1}_{\mathcal{Z}})B \tag{2.125}$$

for all $Y \in \mathrm{L}(\mathcal{Y})$, and in particular $\Phi^*(\mathbb{1}_{\mathcal{Y}}) = A^*B$. The equivalence of statements 2, 6, and 7 follows by the same reasoning as for the case of statements 2, 4, and 5.

Finally, let Γ be the alphabet for which $\mathcal{X} = \mathbb{C}^{\Gamma}$, and consider the operator

$$\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \sum_{a,b \in \Gamma} \mathrm{Tr}(\Phi(E_{a,b}))E_{a,b}. \tag{2.126}$$

If Φ preserves trace, then it follows that

$$\mathrm{Tr}(\Phi(E_{a,b})) = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b, \end{cases} \tag{2.127}$$

and therefore

$$\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \sum_{a \in \Gamma} E_{a,a} = \mathbb{1}_{\mathcal{X}}. \tag{2.128}$$

Conversely, if $\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_{\mathcal{X}}$, then a consideration of the expression (2.126) reveals that (2.127) must hold. As the set $\{E_{a,b} : a, b \in \Gamma\}$ is a basis of $\mathrm{L}(\mathcal{X})$, one concludes by linearity that Φ preserves trace. Statements 1 and 3 are therefore equivalent, which completes the proof. □

Characterizations of channels

Theorems 2.22 and 2.26 can be combined, providing characterizations of channels based on their Choi, Kraus, and Stinespring representations.

Corollary 2.27 *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. The following statements are equivalent:*

1. *Φ is a channel.*
2. *$J(\Phi) \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$ and $\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_{\mathcal{X}}$.*
3. *There exists an alphabet Σ and a collection $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Y})$ satisfying*

   $$\sum_{a \in \Sigma} A_a^* A_a = \mathbb{1}_{\mathcal{X}} \quad \text{and} \quad \Phi(X) = \sum_{a \in \Sigma} A_a X A_a^* \tag{2.129}$$

   for all $X \in \mathrm{L}(\mathcal{X})$.
4. *Statement 3 holds for $|\Sigma| = \mathrm{rank}(J(\Phi))$.*
5. *There exists an isometry $A \in \mathrm{U}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$, for some choice of a complex Euclidean space $\mathcal{Z}$, such that*

   $$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}(AXA^*) \tag{2.130}$$

   for all $X \in \mathrm{L}(\mathcal{X})$.
6. *Statement 5 holds under the requirement $\dim(\mathcal{Z}) = \mathrm{rank}(J(\Phi))$.*

For every choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, one has that the set of channels $\mathrm{C}(\mathcal{X},\mathcal{Y})$ is compact and convex. One way to prove this fact makes use of the previous corollary.

Proposition 2.28 *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces. The set $\mathrm{C}(\mathcal{X},\mathcal{Y})$ is compact and convex.*

Proof The map $J : \mathrm{T}(\mathcal{X},\mathcal{Y}) \to \mathrm{L}(\mathcal{Y}\otimes\mathcal{X})$ defining the Choi representation is linear and invertible. By Corollary 2.27, one has $J^{-1}(\mathcal{A}) = \mathrm{C}(\mathcal{X},\mathcal{Y})$ for $\mathcal{A}$ being defined as

$$\mathcal{A} = \{X \in \mathrm{Pos}(\mathcal{Y}\otimes\mathcal{X}) \,:\, \mathrm{Tr}_{\mathcal{Y}}(X) = \mathbb{1}_{\mathcal{X}}\}. \tag{2.131}$$

It therefore suffices to prove that $\mathcal{A}$ is compact and convex. It is evident that $\mathcal{A}$ is closed and convex, as it is the intersection of the positive semidefinite cone $\mathrm{Pos}(\mathcal{Y}\otimes\mathcal{X})$ with the affine subspace

$$\{X \in \mathrm{L}(\mathcal{Y}\otimes\mathcal{X}) \,:\, \mathrm{Tr}_{\mathcal{Y}}(X) = \mathbb{1}_{\mathcal{X}}\}, \tag{2.132}$$

both of which are closed and convex. To complete the proof, it suffices to prove that $\mathcal{A}$ is bounded. For every $X \in \mathcal{A}$, one has

$$\|X\|_1 = \mathrm{Tr}(X) = \mathrm{Tr}\big(\mathrm{Tr}_{\mathcal{Y}}(X)\big) = \mathrm{Tr}\big(\mathbb{1}_{\mathcal{X}}\big) = \dim(\mathcal{X}), \tag{2.133}$$

and therefore $\mathcal{A}$ is bounded, as required. □

Corollary 2.27 will be used frequently throughout this book, sometimes implicitly. The next proposition, which builds on the unitary equivalence of purifications (Theorem 2.12) to relate a given purification of a positive semidefinite operator to any extension of that operator, is one example of an application of this corollary.

Proposition 2.29 *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be complex Euclidean spaces, and suppose that $u \in \mathcal{X}\otimes\mathcal{Y}$ and $P \in \mathrm{Pos}(\mathcal{X}\otimes\mathcal{Z})$ satisfy*

$$\mathrm{Tr}_{\mathcal{Y}}\big(uu^*\big) = \mathrm{Tr}_{\mathcal{Z}}(P). \tag{2.134}$$

There exists a channel $\Phi \in \mathrm{C}(\mathcal{Y},\mathcal{Z})$ such that

$$\big(\mathbb{1}_{\mathrm{L}(\mathcal{X})}\otimes\Phi\big)(uu^*) = P. \tag{2.135}$$

Proof Let $\mathcal{W}$ be a complex Euclidean space having dimension sufficiently large so that

$$\dim(\mathcal{W}) \geq \mathrm{rank}(P) \quad \text{and} \quad \dim(\mathcal{Z}\otimes\mathcal{W}) \geq \dim(\mathcal{Y}), \tag{2.136}$$

and let $A \in \mathrm{U}(\mathcal{Y},\mathcal{Z}\otimes\mathcal{W})$ be any isometry. Also let $v \in \mathcal{X}\otimes\mathcal{Z}\otimes\mathcal{W}$ satisfy $\mathrm{Tr}_{\mathcal{W}}\big(vv^*\big) = P$. It holds that

$$\begin{aligned}\mathrm{Tr}_{\mathcal{Z}\otimes\mathcal{W}}\big((\mathbb{1}_{\mathcal{X}}\otimes A)uu^*(\mathbb{1}_{\mathcal{X}}\otimes A)^*\big)\\ = \mathrm{Tr}_{\mathcal{Y}}\big(uu^*\big) = \mathrm{Tr}_{\mathcal{Z}}(P) = \mathrm{Tr}_{\mathcal{Z}\otimes\mathcal{W}}\big(vv^*\big).\end{aligned} \tag{2.137}$$

By Theorem 2.12 there must exist a unitary operator $U \in \mathrm{U}(\mathcal{Z}\otimes\mathcal{W})$ such that

$$\big(\mathbb{1}_{\mathcal{X}}\otimes U A\big)u = v. \tag{2.138}$$

Define $\Phi \in \mathrm{T}(\mathcal{Y},\mathcal{Z})$ as

$$\Phi(Y) = \mathrm{Tr}_{\mathcal{W}}\big((UA)Y(UA)^*\big) \tag{2.139}$$

for all $Y \in \mathrm{L}(\mathcal{Y})$. By Corollary 2.27, one has that Φ is a channel. It holds that

$$\begin{aligned}\big(\mathbb{1}_{\mathrm{L}(\mathcal{X})}\otimes\Phi\big)(uu^*) &= \mathrm{Tr}_{\mathcal{W}}\big((\mathbb{1}_{\mathcal{X}}\otimes UA)uu^*(\mathbb{1}_{\mathcal{X}}\otimes UA)^*\big)\\ &= \mathrm{Tr}_{\mathcal{W}}\big(vv^*\big) = P,\end{aligned} \tag{2.140}$$

as required. □

2.2.3 Examples of channels and other mappings

This section describes examples of channels, and other maps, along with their specifications according to the four types of representations discussed above. Many other examples and general classifications of channels and maps will be encountered throughout the book.

Isometric and unitary channels

Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $A, B \in \mathrm{L}(\mathcal{X},\mathcal{Y})$ be operators, and consider the map $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$ defined by

$$\Phi(X) = AXB^* \tag{2.141}$$

for all $X \in \mathrm{L}(\mathcal{X})$.

In the case that $A = B$, and assuming in addition that this operator is a linear isometry from $\mathcal{X}$ to $\mathcal{Y}$, it follows from Corollary 2.27 that Φ is a channel. Such a channel is said to be an *isometric channel*. If $\mathcal{Y} = \mathcal{X}$ and $A = B$ is a unitary operator, Φ is said to be a *unitary channel*. Unitary channels, and convex combinations of unitary channels, are discussed in greater detail in Chapter 4.

The natural representation of the map Φ defined by (2.141) is

$$K(\Phi) = A \otimes \overline{B} \tag{2.142}$$

and the Choi representation of Φ is

$$J(\Phi) = \operatorname{vec}(A)\operatorname{vec}(B)^*. \tag{2.143}$$

The expression (2.141) is a Kraus representation of Φ, and may also be regarded as a trivial example of a Stinespring representation if one takes $\mathcal{Z} = \mathbb{C}$ and observes that the trace acts as the identity mapping on $\mathbb{C}$.

The identity mapping $\mathbb{1}_{\mathrm{L}(\mathcal{X})}$ is a simple example of a unitary channel. The natural representation of this channel is the identity operator $\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{X}}$, while its Choi representation is given by the rank-one operator $\operatorname{vec}(\mathbb{1}_{\mathcal{X}})\operatorname{vec}(\mathbb{1}_{\mathcal{X}})^*$.

Replacement channels and the completely depolarizing channel

Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $A \in \mathrm{L}(\mathcal{X})$ and $B \in \mathrm{L}(\mathcal{Y})$ be operators, and consider the map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ defined as

$$\Phi(X) = \langle A, X \rangle B \tag{2.144}$$

for all $X \in \mathrm{L}(\mathcal{X})$. The natural representation of Φ is

$$K(\Phi) = \operatorname{vec}(B)\operatorname{vec}(A)^*, \tag{2.145}$$

and the Choi representation of Φ is

$$J(\Phi) = B \otimes \overline{A}. \tag{2.146}$$

Kraus and Stinespring representations of Φ may also be constructed, although they are not necessarily enlightening in this particular case. One way to obtain a Kraus representation of Φ is to first write

$$A = \sum_{a \in \Sigma} u_a x_a^* \quad \text{and} \quad B = \sum_{b \in \Gamma} v_b y_b^*, \tag{2.147}$$

for some choice of alphabets Σ and Γ and four sets of vectors:

$$\begin{aligned} \{u_a : a \in \Sigma\},\ \{x_a : a \in \Sigma\} \subset \mathcal{X}, \\ \{v_b : b \in \Gamma\},\ \{y_b : b \in \Gamma\} \subset \mathcal{Y}. \end{aligned} \tag{2.148}$$

It then follows that one Kraus representation of Φ is given by

$$\Phi(X) = \sum_{(a,b) \in \Sigma \times \Gamma} C_{a,b} X D_{a,b}^* \tag{2.149}$$

where $C_{a,b} = v_b u_a^*$ and $D_{a,b} = y_b x_a^*$ for each $a \in \Sigma$ and $b \in \Gamma$, and one Stinespring representation is given by

$$\Phi(X) = \operatorname{Tr}_{\mathcal{Z}}(CXD^*), \tag{2.150}$$

where

$$C = \sum_{(a,b) \in \Sigma \times \Gamma} C_{a,b} \otimes e_{(a,b)}, \quad D = \sum_{(a,b) \in \Sigma \times \Gamma} D_{a,b} \otimes e_{(a,b)}, \tag{2.151}$$

and $\mathcal{Z} = \mathbb{C}^{\Sigma \times \Gamma}$.

If A and B are positive semidefinite operators and the map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is defined by (2.144) for all $X \in \mathrm{L}(\mathcal{X})$, then $J(\Phi) = B \otimes \overline{A}$ is positive semidefinite, and therefore Φ is completely positive by Theorem 2.22. In the case that $A = \mathbb{1}_{\mathcal{X}}$ and $B = \sigma$ for some density operator $\sigma \in \mathrm{D}(\mathcal{Y})$, the map Φ is also trace preserving, and is therefore a channel. Such a channel is a *replacement channel*: it effectively discards its input, replacing it with the state σ.

The *completely depolarizing channel* $\Omega \in \mathrm{C}(\mathcal{X})$ is an important example of a replacement channel. This channel is defined as

$$\Omega(X) = \operatorname{Tr}(X)\omega \tag{2.152}$$

for all $X \in \mathrm{L}(\mathcal{X})$, where

$$\omega = \frac{\mathbb{1}_{\mathcal{X}}}{\dim(\mathcal{X})} \tag{2.153}$$

denotes the completely mixed state defined with respect to the space $\mathcal{X}$. Equivalently, Ω is the unique channel transforming every density operator into this completely mixed state: $\Omega(\rho) = \omega$ for all $\rho \in \mathrm{D}(\mathcal{X})$. From the equations (2.145) and (2.146), one has that the natural representation of the completely depolarizing channel $\Omega \in \mathrm{C}(\mathcal{X})$ is

$$K(\Omega) = \frac{\operatorname{vec}(\mathbb{1}_{\mathcal{X}})\operatorname{vec}(\mathbb{1}_{\mathcal{X}})^*}{\dim(\mathcal{X})}, \tag{2.154}$$

while the Choi representation of this channel is

$$J(\Omega) = \frac{\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{X}}}{\dim(\mathcal{X})}. \tag{2.155}$$

The transpose map

Let Σ be an alphabet, let $\mathcal{X} = \mathbb{C}^{\Sigma}$, and let $\mathrm{T} \in \mathrm{T}(\mathcal{X})$ denote the transpose map, defined as

$$\mathrm{T}(X) = X^{\mathsf{T}} \tag{2.156}$$

for all $X \in \mathrm{L}(\mathcal{X})$. This map will play an important role in Chapter 6, due to its connections to properties of entangled states.

The natural representation $K(\mathrm{T})$ of T must, by definition, satisfy

$$K(\mathrm{T})\,\mathrm{vec}(X) = \mathrm{vec}(X^{\mathsf{T}}) \tag{2.157}$$

for all $X \in \mathrm{L}(\mathcal{X})$. By considering those operators of the form $X = uv^{\mathsf{T}}$ for vectors $u, v \in \mathcal{X}$, one finds that

$$K(\mathrm{T})(u \otimes v) = v \otimes u. \tag{2.158}$$

It follows that $K(\mathrm{T}) = W$, for $W \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$ being the *swap operator*, which is defined by the action $W(u \otimes v) = v \otimes u$ for all vectors $u, v \in \mathcal{X}$.

The Choi representation of T is also equal to the swap operator, as

$$J(\mathrm{T}) = \sum_{a,b \in \Sigma} E_{b,a} \otimes E_{a,b} = W. \tag{2.159}$$

Under the assumption that $|\Sigma| \geq 2$, it therefore follows from Theorem 2.22 that T is not a completely positive map, as W is not a positive semidefinite operator in this case.

One example of a Kraus representation of T is

$$\mathrm{T}(X) = \sum_{a,b \in \Sigma} E_{a,b} X E_{b,a}^* \tag{2.160}$$

for all $X \in \mathrm{L}(\mathcal{X})$, from which it follows that $\mathrm{T}(X) = \mathrm{Tr}_{\mathcal{Z}}(AXB^*)$ is a Stinespring representation of T for $\mathcal{Z} = \mathbb{C}^{\Sigma \times \Sigma}$,

$$A = \sum_{a,b \in \Sigma} E_{a,b} \otimes e_{(a,b)}, \quad \text{and} \quad B = \sum_{a,b \in \Sigma} E_{b,a} \otimes e_{(a,b)}. \tag{2.161}$$

The completely dephasing channel

Let Σ be an alphabet and let $\mathcal{X} = \mathbb{C}^{\Sigma}$. The map $\Delta \in \mathrm{T}(\mathcal{X})$ defined as

$$\Delta(X) = \sum_{a \in \Sigma} X(a,a) E_{a,a} \tag{2.162}$$

for every $X \in \mathrm{L}(\mathcal{X})$ is an example of a channel known as the *completely dephasing channel*. This channel has the effect of replacing every off-diagonal entry of a given operator $X \in \mathrm{L}(\mathcal{X})$ by 0 and leaving the diagonal entries unchanged.

Through the association of diagonal density operators with probabilistic states, as discussed in Section 2.1.2, one may view the channel Δ as an ideal channel for *classical communication*: it acts as the identity mapping on every diagonal density operator, so that it effectively transmits classical probabilistic states without error, while all other states are mapped to the probabilistic state given by their diagonal entries.

The natural representation of Δ must satisfy the equation

$$K(\Delta)\,\mathrm{vec}(E_{a,b}) = \begin{cases} \mathrm{vec}(E_{a,b}) & \text{if } a = b \\ 0 & \text{if } a \neq b, \end{cases} \tag{2.163}$$

which is equivalent to

$$K(\Delta)(e_a \otimes e_b) = \begin{cases} e_a \otimes e_b & \text{if } a = b \\ 0 & \text{if } a \neq b, \end{cases} \tag{2.164}$$

for every $a, b \in \Sigma$. It follows that

$$K(\Delta) = \sum_{a \in \Sigma} E_{a,a} \otimes E_{a,a}. \tag{2.165}$$

Similar to the transpose mapping, the Choi representation of Δ happens to coincide with its natural representation, as the calculation

$$J(\Delta) = \sum_{a,b \in \Sigma} \Delta(E_{a,b}) \otimes E_{a,b} = \sum_{a \in \Sigma} E_{a,a} \otimes E_{a,a} \tag{2.166}$$

reveals. It is evident from this expression, together with Corollary 2.27, that Δ is indeed a channel.

One example of a Kraus representation of Δ is

$$\Delta(X) = \sum_{a \in \Sigma} E_{a,a} X E_{a,a}^*, \tag{2.167}$$

and an example of a Stinespring representation of Δ is

$$\Delta(X) = \mathrm{Tr}_{\mathcal{Z}}(AXA^*) \tag{2.168}$$

for $\mathcal{Z} = \mathbb{C}^{\Sigma}$ and

$$A = \sum_{a \in \Sigma} (e_a \otimes e_a) e_a^*. \tag{2.169}$$

A digression on classical registers

Classical probabilistic states of registers may be associated with diagonal density operators, as discussed in Section 2.1.2. The term *classical register* was mentioned in that discussion but not fully explained. It is appropriate to make this notion more precise, now that channels (and the completely dephasing channel in particular) have been introduced.

From a mathematical point of view, classical registers are not defined in a manner that is distinct from ordinary (quantum) registers. Rather, the term

classical register will be used to refer to any register that, by the nature of the processes under consideration, would be unaffected by an application of the completely dephasing channel Δ at any moment during its existence. Every state of a classical register is necessarily a diagonal density operator, corresponding to a probabilistic state, as these are the density operators that are invariant under the action of the channel Δ. Moreover, the correlations that may exist between a classical register and one or more other registers are limited. For example, for a classical register X and an arbitrary register Y, the only states of the compound register (X,Y) that are consistent with the term *classical register* being applied to X are those taking the form

$$\sum_{a\in\Sigma} p(a)E_{a,a}\otimes\rho_a, \tag{2.170}$$

for Σ being the classical state set of X, $\{\rho_a : a\in\Sigma\}\subseteq \mathrm{D}(\mathcal{Y})$ being an arbitrary collection of states of Y, and $p\in\mathcal{P}(\Sigma)$ being a probability vector. States of this form are commonly called *classical-quantum* states. It is both natural and convenient in some situations to associate the state (2.170) with the ensemble $\eta:\Sigma\to\mathrm{Pos}(\mathcal{Y})$ defined as $\eta(a)=p(a)\rho_a$ for each $a\in\Sigma$.

2.2.4 Extremal channels

For any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, the set of channels $\mathrm{C}(\mathcal{X},\mathcal{Y})$ is compact and convex (by Proposition 2.28). A characterization of the extreme points of this set is given by Theorem 2.31 below. The following lemma will be used in the proof of this theorem.

Lemma 2.30 *Let $A\in\mathrm{L}(\mathcal{Y},\mathcal{X})$ be an operator, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. It holds that*

$$\{P\in\mathrm{Pos}(\mathcal{X}) : \mathrm{im}(P)\subseteq\mathrm{im}(A)\} = \{AQA^* : Q\in\mathrm{Pos}(\mathcal{Y})\}. \tag{2.171}$$

Proof For every $Q\in\mathrm{Pos}(\mathcal{Y})$, it holds that AQA^* is positive semidefinite and satisfies $\mathrm{im}(AQA^*)\subseteq\mathrm{im}(A)$. The set on the right-hand side of (2.171) is therefore contained in the set on the left-hand side.

For the reverse containment, if $P\in\mathrm{Pos}(\mathcal{X})$ satisfies $\mathrm{im}(P)\subseteq\mathrm{im}(A)$, then by setting

$$Q = A^+P(A^+)^*, \tag{2.172}$$

for A^+ denoting the Moore–Penrose pseudo-inverse of A, one obtains

$$AQA^* = (AA^+)P(AA^+)^* = \Pi_{\mathrm{im}(A)}P\Pi_{\mathrm{im}(A)} = P, \tag{2.173}$$

which completes the proof. □

Theorem 2.31 (Choi) *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\Phi\in\mathrm{C}(\mathcal{X},\mathcal{Y})$ be a channel, and let $\{A_a : a\in\Sigma\}\subset\mathrm{L}(\mathcal{X},\mathcal{Y})$ be a linearly independent set of operators satisfying*

$$\Phi(X) = \sum_{a\in\Sigma} A_aXA_a^* \tag{2.174}$$

for all $X\in\mathrm{L}(\mathcal{X})$. The channel Φ is an extreme point of the set $\mathrm{C}(\mathcal{X},\mathcal{Y})$ if and only if the collection

$$\{A_b^*A_a : (a,b)\in\Sigma\times\Sigma\}\subset\mathrm{L}(\mathcal{X}) \tag{2.175}$$

of operators is linearly independent.

Proof Let $\mathcal{Z}=\mathbb{C}^\Sigma$, define an operator $M\in\mathrm{L}(\mathcal{Z},\mathcal{Y}\otimes\mathcal{X})$ as

$$M = \sum_{a\in\Sigma}\mathrm{vec}(A_a)e_a^*, \tag{2.176}$$

and observe that

$$MM^* = \sum_{a\in\Sigma}\mathrm{vec}(A_a)\,\mathrm{vec}(A_a)^* = J(\Phi). \tag{2.177}$$

As $\{A_a : a\in\Sigma\}$ is a linearly independent collection of operators, it must hold that $\ker(M)=\{0\}$.

Assume first that Φ is not an extreme point of $\mathrm{C}(\mathcal{X},\mathcal{Y})$. It follows that there exist channels $\Psi_0,\Psi_1\in\mathrm{C}(\mathcal{X},\mathcal{Y})$, with $\Psi_0\neq\Psi_1$, along with a scalar $\lambda\in(0,1)$, such that

$$\Phi = \lambda\Psi_0 + (1-\lambda)\Psi_1. \tag{2.178}$$

Let $P=J(\Phi)$, $Q_0=J(\Psi_0)$, and $Q_1=J(\Psi_1)$, so that

$$P = \lambda Q_0 + (1-\lambda)Q_1. \tag{2.179}$$

As Φ, Ψ_0, and Ψ_1 are channels, the operators $P,Q_0,Q_1\in\mathrm{Pos}(\mathcal{Y}\otimes\mathcal{X})$ are positive semidefinite and satisfy

$$\mathrm{Tr}_{\mathcal{Y}}(P) = \mathrm{Tr}_{\mathcal{Y}}(Q_0) = \mathrm{Tr}_{\mathcal{Y}}(Q_1) = \mathbb{1}_{\mathcal{X}}, \tag{2.180}$$

by Corollary 2.27.

As λ is positive and the operators Q_0 and Q_1 are positive semidefinite, the equation (2.179) implies

$$\mathrm{im}(Q_0)\subseteq\mathrm{im}(P) = \mathrm{im}(M). \tag{2.181}$$

It follows by Lemma 2.30 that there exists a positive semidefinite operator $R_0\in\mathrm{Pos}(\mathcal{Z})$ for which $Q_0=MR_0M^*$. By similar reasoning, there exists a positive semidefinite operator $R_1\in\mathrm{Pos}(\mathcal{Z})$ for which $Q_1=MR_1M^*$.

Letting $H = R_0 - R_1$, one finds that

$$0 = \mathrm{Tr}_{\mathcal{Y}}(Q_0) - \mathrm{Tr}_{\mathcal{Y}}(Q_1) = \mathrm{Tr}_{\mathcal{Y}}(MHM^*) = \sum_{a,b\in\Sigma} H(a,b)(A_b^*A_a)^{\mathsf{T}}, \quad (2.182)$$

and therefore

$$\sum_{a,b\in\Sigma} H(a,b)A_b^*A_a = 0. \quad (2.183)$$

Because $\Psi_0 \neq \Psi_1$, it holds that $Q_0 \neq Q_1$, so $R_0 \neq R_1$, and therefore $H \neq 0$. It has therefore been proved that $\{A_b^*A_a : (a,b) \in \Sigma\times\Sigma\}$ is a linearly dependent collection of operators.

Now assume the set (2.175) is linearly dependent:

$$\sum_{a,b\in\Sigma} Z(a,b)A_b^*A_a = 0 \quad (2.184)$$

for some choice of a nonzero operator $Z \in \mathrm{L}(\mathcal{Z})$. By taking the adjoint of both sides of this equation, one finds that

$$\sum_{a,b\in\Sigma} Z^*(a,b)A_b^*A_a = 0, \quad (2.185)$$

from which it follows that

$$\sum_{a,b\in\Sigma} H(a,b)A_b^*A_a = 0 \quad (2.186)$$

for both of the Hermitian operators

$$H = \frac{Z+Z^*}{2} \quad \text{and} \quad H = \frac{Z-Z^*}{2i}. \quad (2.187)$$

At least one of these operators must be nonzero, which implies that (2.186) must hold for some choice of a nonzero Hermitian operator H. Let such a choice of H be fixed, and assume moreover that $\|H\| = 1$ (which causes no loss of generality as (2.186) still holds if H is replaced by $H/\|H\|$).

Let $\Psi_0, \Psi_1 \in \mathrm{T}(\mathcal{X},\mathcal{Y})$ be the mappings defined by the equations

$$J(\Psi_0) = M(\mathbb{1}+H)M^* \quad \text{and} \quad J(\Psi_1) = M(\mathbb{1}-H)M^*. \quad (2.188)$$

Because H is Hermitian and satisfies $\|H\| = 1$, one has that the operators $\mathbb{1}+H$ and $\mathbb{1}-H$ are both positive semidefinite. The operators $M(\mathbb{1}+H)M^*$ and $M(\mathbb{1}-H)M^*$ are therefore positive semidefinite as well, implying that

Ψ_0 and Ψ_1 are completely positive, by Theorem 2.22. It holds that

$$\begin{aligned}\mathrm{Tr}_{\mathcal{Y}}(MHM^*) &= \sum_{a,b\in\Sigma} H(a,b)\,(A_b^*A_a)^{\mathsf{T}} \\ &= \left(\sum_{a,b\in\Sigma} H(a,b)A_b^*A_a\right)^{\mathsf{T}} = 0\end{aligned} \quad (2.189)$$

and therefore the following two equations hold:

$$\begin{aligned}\mathrm{Tr}_{\mathcal{Y}}(J(\Psi_0)) &= \mathrm{Tr}_{\mathcal{Y}}(MM^*) + \mathrm{Tr}_{\mathcal{Y}}(MHM^*) = \mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_{\mathcal{X}}, \\ \mathrm{Tr}_{\mathcal{Y}}(J(\Psi_1)) &= \mathrm{Tr}_{\mathcal{Y}}(MM^*) - \mathrm{Tr}_{\mathcal{Y}}(MHM^*) = \mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_{\mathcal{X}}.\end{aligned} \quad (2.190)$$

Thus, Ψ_0 and Ψ_1 are trace preserving by Theorem 2.26, and are therefore channels.

Finally, given that $H \neq 0$ and $\ker(M) = \{0\}$, it holds that $J(\Psi_0) \neq J(\Psi_1)$, so that $\Psi_0 \neq \Psi_1$. As

$$\frac{1}{2}J(\Psi_0) + \frac{1}{2}J(\Psi_1) = MM^* = J(\Phi), \quad (2.191)$$

one has that

$$\Phi = \frac{1}{2}\Psi_0 + \frac{1}{2}\Psi_1, \quad (2.192)$$

which demonstrates that Φ is not an extreme point of $\mathrm{C}(\mathcal{X},\mathcal{Y})$. □

Example 2.32 Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces such that $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$, let $A \in \mathrm{U}(\mathcal{X},\mathcal{Y})$ be an isometry, and let $\Phi \in \mathrm{C}(\mathcal{X},\mathcal{Y})$ be the isometric channel defined by

$$\Phi(X) = AXA^* \quad (2.193)$$

for all $X \in \mathrm{L}(\mathcal{X})$. The set $\{A^*A\}$ contains a single nonzero operator, and is therefore linearly independent. By Theorem 2.31, Φ is an extreme point of the set $\mathrm{C}(\mathcal{X},\mathcal{Y})$.

Example 2.33 Let $\Sigma = \{0,1\}$ denote the binary alphabet, and let $\mathcal{X} = \mathbb{C}^{\Sigma}$ and $\mathcal{Y} = \mathbb{C}^{\Sigma\times\Sigma}$. Also define operators $A_0, A_1 \in \mathrm{L}(\mathcal{X},\mathcal{Y})$ as

$$\begin{aligned}A_0 &= \frac{1}{\sqrt{6}}(2E_{00,0} + E_{01,1} + E_{10,1}), \\ A_1 &= \frac{1}{\sqrt{6}}(2E_{11,1} + E_{01,0} + E_{10,0}).\end{aligned} \quad (2.194)$$

(Elements of the form $(a,b) \in \Sigma\times\Sigma$ have been written as ab for the sake of

clarity.) Expressed as matrices (with respect to the natural orderings of Σ and $\Sigma \times \Sigma$), these operators are as follows:

$$A_0 = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad A_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix}. \tag{2.195}$$

Now, define a channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ as

$$\Phi(X) = A_0 X A_0^* + A_1 X A_1^* \tag{2.196}$$

for every $X \in \mathrm{L}(\mathcal{X})$. It holds that

$$\begin{aligned} A_0^* A_0 &= \frac{1}{3} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A_0^* A_1 = \frac{1}{3} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\ A_1^* A_0 &= \frac{1}{3} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad A_1^* A_1 = \frac{1}{3} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}. \end{aligned} \tag{2.197}$$

The set

$$\{A_0^* A_0, A_0^* A_1, A_1^* A_0, A_1^* A_1\} \tag{2.198}$$

is linearly independent, and therefore Theorem 2.31 implies that Φ is an extreme point of $\mathrm{C}(\mathcal{X}, \mathcal{Y})$.

2.3 Measurements

Measurements provide the mechanism through which classical information may be extracted from quantum states. This section defines measurements, and various notions connected with measurements, and provides a basic mathematical development of this concept.

2.3.1 Two equivalent definitions of measurements

When a hypothetical observer measures a register, the observer obtains a classical measurement outcome (as opposed to a description of the state of the register, for instance). In general, this measurement outcome is generated at random, according to a probability distribution that is determined by the measurement together with the state of the register immediately before the measurement was performed. In this way, measurements allow one to associate a meaning to the density operator description of quantum states,

at least insofar as the density operators determine the probabilities with which different classical outcomes occur for each possible measurement.

Measurements can be defined in mathematical terms in two different, but equivalent, ways. Both ways will be described in this section, and their equivalence will be explained.

Measurements defined by measurement operators

The following definition represents the first formulation of measurements to be described in this book. The precise mathematical meaning of the term *measurement* used throughout this book coincides with this definition.

Definition 2.34 A *measurement* is a function of the form

$$\mu : \Sigma \to \mathrm{Pos}(\mathcal{X}), \tag{2.199}$$

for some choice of an alphabet Σ and a complex Euclidean space $\mathcal{X}$, satisfying the constraint

$$\sum_{a \in \Sigma} \mu(a) = \mathbb{1}_{\mathcal{X}}. \tag{2.200}$$

The set Σ is the set of *measurement outcomes* of this measurement, and each operator $\mu(a)$ is the *measurement operator* associated with the outcome $a \in \Sigma$.

When a measurement μ is performed on a given register X, it must be assumed that μ takes the form (2.199), for some choice of an alphabet Σ and for $\mathcal{X}$ being the complex Euclidean space associated with X. Two things happen when such a measurement is performed, assuming the state of X immediately prior to the measurement is $\rho \in \mathrm{D}(\mathcal{X})$:

1. An element of Σ is selected at random. The probability distribution that describes this random selection is represented by the probability vector $p \in \mathcal{P}(\Sigma)$ defined as
$$p(a) = \langle \mu(a), \rho \rangle \tag{2.201}$$
for each $a \in \Sigma$.
2. The register X ceases to exist, in the sense that it no longer has a defined state and cannot be considered in further calculations.

It is evident from the first item that the probabilities associated with the outcomes of a given measurement depend linearly on the state that is measured. It is also evident that the probability vector $p \in \mathcal{P}(\Sigma)$ defined by (2.201) is indeed a probability vector: as ρ and $\mu(a)$ are both positive

semidefinite, their inner product $\langle \mu(a), \rho \rangle$ is nonnegative, and summing these values gives

$$\sum_{a \in \Sigma} p(a) = \sum_{a \in \Sigma} \langle \mu(a), \rho \rangle = \langle \mathbb{1}_{\mathcal{X}}, \rho \rangle = \mathrm{Tr}(\rho) = 1. \tag{2.202}$$

The assumption that registers cease to exist after being measured is not universal within quantum information theory—an alternative definition, in which the states of registers after they are measured is specified, does not make this requirement. Measurements of this alternative type, which are called *nondestructive measurements* in this book, are discussed in greater detail in Section 2.3.2. Nondestructive measurements can, however, be described as compositions of ordinary measurements (as described above) and channels. For this reason, no generality is lost in making the assumption that registers cease to exist upon being measured.

It is sometimes convenient to specify a measurement by describing its measurement operators as a collection indexed by its set of measurement outcomes. In particular, when one refers to a measurement as a collection

$$\{P_a : a \in \Sigma\} \subset \mathrm{Pos}(\mathcal{X}), \tag{2.203}$$

it is to be understood that the measurement is given by $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$, where $\mu(a) = P_a$ for each $a \in \Sigma$.

Measurements as channels

The second formulation of measurements, which is equivalent to the first, essentially describes measurements as channels whose outputs are stored in classical registers. The following definition of *quantum-to-classical channels* makes this notion precise.

Definition 2.35 Let $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be a channel, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. It is said that Φ is a *quantum-to-classical channel* if

$$\Phi = \Delta\Phi, \tag{2.204}$$

for $\Delta \in \mathrm{C}(\mathcal{Y})$ denoting the completely dephasing channel, defined with respect to the space $\mathcal{Y}$.

An equivalent condition for a channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ to be a quantum-to-classical channel is that $\Phi(\rho)$ is a diagonal density operator for every $\rho \in \mathrm{D}(\mathcal{X})$. The following simple proposition establishes that this is so.

Proposition 2.36 *Let $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be a channel, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. It holds that Φ is a quantum-to-classical channel if and only if $\Phi(\rho)$ is diagonal for every $\rho \in \mathrm{D}(\mathcal{X})$.*

Proof If Φ is a quantum-to-classical channel, then

$$\Phi(\rho) = \Delta(\Phi(\rho)), \tag{2.205}$$

and therefore $\Phi(\rho)$ is diagonal, for every density operator $\rho \in \mathrm{D}(\mathcal{X})$.

Conversely, if $\Phi(\rho)$ is diagonal, then $\Phi(\rho) = \Delta(\Phi(\rho))$, and therefore

$$(\Phi - \Delta\Phi)(\rho) = 0, \tag{2.206}$$

for every $\rho \in \mathrm{D}(\mathcal{X})$. As the density operators $\mathrm{D}(\mathcal{X})$ span all of $\mathrm{L}(\mathcal{X})$, it follows that $\Phi = \Delta\Phi$, and therefore Φ is a quantum-to-classical channel. □

The next theorem reveals the equivalence between quantum-to-classical channels and measurements. In essence, quantum-to-classical channels of the form $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ represent precisely those channels that can be realized as a measurement of a register X, according to a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$, followed by the measurement outcome being stored in a register Y having classical state set Σ.

Theorem 2.37 *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, and let $\mathcal{Y} = \mathbb{C}^{\Sigma}$. The following two complementary facts hold:*

1. *For every quantum-to-classical channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$, there exists a unique measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ for which the equation*
$$\Phi(X) = \sum_{a \in \Sigma} \langle \mu(a), X \rangle \, E_{a,a} \tag{2.207}$$
holds for all $X \in \mathrm{L}(\mathcal{X})$.
2. *For every measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$, the mapping $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ defined by (2.207) for all $X \in \mathrm{L}(\mathcal{X})$ is a quantum-to-classical channel.*

Proof Assume first that $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ is a quantum-to-classical channel. It therefore holds that

$$\Phi(X) = \Delta(\Phi(X)) = \sum_{a \in \Sigma} \langle E_{a,a}, \Phi(X) \rangle E_{a,a} = \sum_{a \in \Sigma} \langle \Phi^*(E_{a,a}), X \rangle E_{a,a} \tag{2.208}$$

for all $X \in \mathrm{L}(\mathcal{X})$. Define a function $\mu : \Sigma \to \mathrm{L}(\mathcal{X})$ as

$$\mu(a) = \Phi^*(E_{a,a}) \tag{2.209}$$

for each $a \in \Sigma$. As Φ is positive, so too is Φ^* (by Proposition 2.18), and therefore $\mu(a) \in \mathrm{Pos}(\mathcal{X})$ for each $a \in \Sigma$. Moreover, as Φ preserves trace, it holds (by Theorem 2.26) that Φ^* is unital, and therefore

$$\sum_{a \in \Sigma} \mu(a) = \sum_{a \in \Sigma} \Phi^*(E_{a,a}) = \Phi^*(\mathbb{1}_{\mathcal{Y}}) = \mathbb{1}_{\mathcal{X}}. \tag{2.210}$$

It follows that μ is a measurement for which (2.207) holds for all $X \in \mathrm{L}(\mathcal{X})$.

Toward proving the uniqueness of the measurement μ satisfying (2.207) for all $X \in \mathrm{L}(\mathcal{X})$, let $\nu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be an arbitrary measurement for which the equation

$$\Phi(X) = \sum_{a \in \Sigma} \langle \nu(a), X \rangle \, E_{a,a} \tag{2.211}$$

holds for all $X \in \mathrm{L}(\mathcal{X})$. One then has that

$$\sum_{a \in \Sigma} \langle \mu(a) - \nu(a), X \rangle \, E_{a,a} = 0 \tag{2.212}$$

for all $X \in \mathrm{L}(\mathcal{X})$, which implies that $\nu(a) = \mu(a)$ for every $a \in \Sigma$, and completes the proof of the first fact.

Now assume that $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ is a measurement, and let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be defined by (2.207). The Choi representation of this map is

$$J(\Phi) = \sum_{a \in \Sigma} E_{a,a} \otimes \overline{\mu(a)}. \tag{2.213}$$

This is a positive semidefinite operator, and it holds that

$$\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \sum_{a \in \Sigma} \overline{\mu(a)} = \overline{\mathbb{1}_{\mathcal{X}}} = \mathbb{1}_{\mathcal{X}}. \tag{2.214}$$

By Corollary 2.27, it holds that Φ is a channel. It is evident from inspection that $\Phi(\rho)$ is diagonal for every $\rho \in \mathrm{D}(\mathcal{X})$, and therefore Φ is a quantum-to-classical channel by Proposition 2.36, which completes the proof of the second statement. □

As the following proposition establishes, the set of quantum-to-classical channels of the form $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ is both compact and convex.

Proposition 2.38 *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces. The set of quantum-to-classical channels having the form $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ is compact and convex.*

Proof It will first be observed that the set of quantum-to-classical channels of the form $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ is given by

$$\{\Delta\Psi \,:\, \Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})\}, \tag{2.215}$$

for $\Delta \in \mathrm{C}(\mathcal{Y})$ being the completely dephasing channel defined with respect to the space $\mathcal{Y}$. Indeed, for every channel $\Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$, it holds that $\Delta\Psi$ is a quantum-to-classical channel by virtue of the fact that the channel Δ is idempotent (i.e., $\Delta\Delta = \Delta$). On the other hand, every quantum-to-classical channel Φ satisfies $\Phi = \Delta\Phi$ by definition, and is therefore represented in the set (2.215) by taking $\Psi = \Phi$.

By Proposition 2.28, the set $\mathrm{C}(\mathcal{X}, \mathcal{Y})$ is compact and convex. The mapping $\Psi \mapsto \Delta\Psi$ defined on $\mathrm{C}(\mathcal{X}, \mathcal{Y})$ is continuous, and therefore it maps $\mathrm{C}(\mathcal{X}, \mathcal{Y})$ to a compact and convex set. The image of $\mathrm{C}(\mathcal{X}, \mathcal{Y})$ under this mapping is precisely the set (2.215), which coincides with the set of quantum-to-classical channels of the form $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$, so the proof is complete. □

2.3.2 Basic notions concerning measurements

The subsections that follow introduce various notions and facts connected with measurements.

Product measurements

Suppose $\mathsf{X} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$ is a compound register. One may then consider a collection of measurements

$$\begin{aligned} \mu_1 &: \Sigma_1 \to \mathrm{Pos}(\mathcal{Y}_1) \\ &\vdots \\ \mu_n &: \Sigma_n \to \mathrm{Pos}(\mathcal{Y}_n) \end{aligned} \tag{2.216}$$

to be performed independently on the registers $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$. Such a process may be viewed as a single measurement

$$\mu : \Sigma_1 \times \cdots \times \Sigma_n \to \mathrm{Pos}(\mathcal{X}) \tag{2.217}$$

on X that is defined as

$$\mu(a_1, \ldots, a_n) = \mu_1(a_1) \otimes \cdots \otimes \mu_n(a_n) \tag{2.218}$$

for each tuple $(a_1, \ldots, a_n) \in \Sigma_1 \times \cdots \times \Sigma_n$. A measurement μ of this sort is said to be a *product measurement* on X.

It may be verified that when a product measurement is performed on a product state, the measurement outcomes resulting from the individual measurements are independently distributed.

Partial measurements

Suppose $\mathsf{X} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$ is a compound register, and a measurement

$$\mu : \Sigma \to \mathrm{Pos}(\mathcal{Y}_k) \tag{2.219}$$

is performed only on the register Y_k, for a single choice of $k \in \{1, \ldots, n\}$. Such a measurement must not only produce a measurement outcome $a \in \Sigma$,

but must also determine the resulting state of the register

$$(\mathsf{Y}_1, \ldots, \mathsf{Y}_{k-1}, \mathsf{Y}_{k+1}, \ldots, \mathsf{Y}_n), \tag{2.220}$$

conditioned on the measurement outcome that was obtained. For a given state $\rho \in \mathrm{D}(\mathcal{X})$ of the register X, the probability for each measurement outcome to appear, along with the corresponding post-measurement state of the register (2.220), may be calculated by considering the quantum-to-classical channel that corresponds to the measurement μ.

Let this quantum-to-classical channel be denoted by $\Phi \in \mathrm{C}(\mathcal{Y}_k, \mathcal{Z})$, for $\mathcal{Z} = \mathbb{C}^{\Sigma}$, so that

$$\Phi(Y) = \sum_{a \in \Sigma} \langle \mu(a), Y \rangle E_{a,a} \tag{2.221}$$

for every $Y \in \mathrm{L}(\mathcal{Y}_k)$. Consider the state of the compound register

$$(\mathsf{Z}, \mathsf{Y}_1, \ldots, \mathsf{Y}_{k-1}, \mathsf{Y}_{k+1}, \ldots, \mathsf{Y}_n) \tag{2.222}$$

obtained by applying the channel Φ to Y_k, followed by the application of a channel that performs the permutation of registers

$$(\mathsf{Y}_1, \ldots, \mathsf{Y}_{k-1}, \mathsf{Z}, \mathsf{Y}_{k+1}, \ldots, \mathsf{Y}_n) \to (\mathsf{Z}, \mathsf{Y}_1, \ldots, \mathsf{Y}_{k-1}, \mathsf{Y}_{k+1}, \ldots, \mathsf{Y}_n) \tag{2.223}$$

without changing the contents of these individual registers. The state of the register (2.222) that results may be written explicitly as

$$\sum_{a \in \Sigma} E_{a,a} \otimes \mathrm{Tr}_{\mathcal{Y}_k}\big((\mathbb{1}_{\mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_{k-1}} \otimes \mu(a) \otimes \mathbb{1}_{\mathcal{Y}_{k+1} \otimes \cdots \otimes \mathcal{Y}_n})\rho\big). \tag{2.224}$$

The state (2.224) is a classical-quantum state, and is naturally associated with the ensemble

$$\eta : \Sigma \to \mathrm{Pos}(\mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_{k-1} \otimes \mathcal{Y}_{k+1} \otimes \cdots \otimes \mathcal{Y}_n) \tag{2.225}$$

defined as

$$\eta(a) = \mathrm{Tr}_{\mathcal{Y}_k}\big((\mathbb{1}_{\mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_{k-1}} \otimes \mu(a) \otimes \mathbb{1}_{\mathcal{Y}_{k+1} \otimes \cdots \otimes \mathcal{Y}_n})\rho\big) \tag{2.226}$$

for each $a \in \Sigma$. This ensemble describes the distribution of measurement outcomes of the measurement μ and the states of the remaining registers that result. That is, each measurement outcome $a \in \Sigma$ appears with probability

$$\mathrm{Tr}(\eta(a)) = \langle \mu(a), \rho[\mathsf{Y}_k] \rangle, \tag{2.227}$$

and conditioned on an outcome $a \in \Sigma$ that appears with positive probability, the resulting state of $(\mathsf{Y}_1, \ldots, \mathsf{Y}_{k-1}, \mathsf{Y}_{k+1}, \ldots, \mathsf{Y}_n)$ becomes

$$\frac{\eta(a)}{\mathrm{Tr}(\eta(a))} = \frac{\mathrm{Tr}_{\mathcal{Y}_k}\big((\mathbb{1}_{\mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_{k-1}} \otimes \mu(a) \otimes \mathbb{1}_{\mathcal{Y}_{k+1} \otimes \cdots \otimes \mathcal{Y}_n})\rho\big)}{\langle \mu(a), \rho[\mathsf{Y}_k] \rangle}. \tag{2.228}$$

Example 2.39 Let Σ be an alphabet, and let Y and Z be registers whose classical state sets are given by Σ, so that $\mathcal{Y} = \mathbb{C}^{\Sigma}$ and $\mathcal{Z} = \mathbb{C}^{\Sigma}$. Define a state $\tau \in \mathrm{D}(\mathcal{Y} \otimes \mathcal{Z})$ as

$$\tau = \frac{1}{|\Sigma|} \sum_{b,c \in \Sigma} E_{b,c} \otimes E_{b,c}\,, \tag{2.229}$$

and consider an arbitrary measurement $\mu : \Gamma \to \mathrm{Pos}(\mathcal{Y})$. If this measurement is performed on Y when the pair (Y, Z) is in the state τ, then each outcome $a \in \Gamma$ appears with probability

$$p(a) = \langle \mu(a), \rho[\mathsf{Y}] \rangle = \frac{\mathrm{Tr}(\mu(a))}{|\Sigma|}. \tag{2.230}$$

Conditioned on the event that the measurement outcome a appears, the state of Z becomes

$$\begin{aligned} &\frac{1}{p(a)} \mathrm{Tr}_{\mathcal{Y}}\big((\mu(a) \otimes \mathbb{1}_{\mathcal{Z}})\tau\big) \\ &\quad = \frac{|\Sigma|}{\mathrm{Tr}(\mu(a))} \frac{1}{|\Sigma|} \sum_{b,c \in \Sigma} \mathrm{Tr}(\mu(a) E_{b,c}) E_{b,c} = \frac{\mu(a)^{\mathsf{T}}}{\mathrm{Tr}(\mu(a))}. \end{aligned} \tag{2.231}$$

Projective measurements and Naimark's theorem

A measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ is said to be a *projective measurement* if each of its measurement operators is a projection: $\mu(a) \in \mathrm{Proj}(\mathcal{X})$ for every $a \in \Sigma$.

The following proposition demonstrates that the measurement operators of a projective measurement must be pairwise orthogonal, and must therefore project onto orthogonal subspaces. For a projective measurement of the form $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$, there can therefore be no more than $\dim(\mathcal{X})$ distinct values of $a \in \Sigma$ for which $\mu(a)$ is nonzero.

Proposition 2.40 *Let Σ be an alphabet, let $\mathcal{X}$ be a complex Euclidean space, and let $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be a projective measurement. The set*

$$\{\mu(a) \,:\, a \in \Sigma\} \tag{2.232}$$

is an orthogonal set.

Proof As μ is a measurement, it holds that

$$\sum_{a \in \Sigma} \mu(a) = \mathbb{1}_{\mathcal{X}}, \tag{2.233}$$

and therefore this sum must square to itself:

$$\sum_{a,b\in\Sigma} \mu(a)\mu(b) = \left(\sum_{a\in\Sigma} \mu(a)\right)^2 = \sum_{a\in\Sigma} \mu(a). \tag{2.234}$$

Because each operator $\mu(a)$ is a projection operator, it follows that

$$\sum_{a,b\in\Sigma} \mu(a)\mu(b) = \sum_{a\in\Sigma} \mu(a) + \sum_{\substack{a,b\in\Sigma \\ a\neq b}} \mu(a)\mu(b), \tag{2.235}$$

and therefore

$$\sum_{\substack{a,b\in\Sigma \\ a\neq b}} \mu(a)\mu(b) = 0. \tag{2.236}$$

Taking the trace of both sides of this equation yields

$$\sum_{\substack{a,b\in\Sigma \\ a\neq b}} \langle \mu(a), \mu(b)\rangle = 0. \tag{2.237}$$

The inner product of any two positive semidefinite operators is nonnegative, and therefore $\langle \mu(a), \mu(b)\rangle = 0$ for all $a, b \in \Sigma$ with $a \neq b$, which completes the proof. □

For any orthonormal basis $\{x_a : a \in \Sigma\}$ of a complex Euclidean space $\mathcal{X} = \mathbb{C}^\Sigma$, the measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ defined as

$$\mu(a) = x_a x_a^* \tag{2.238}$$

for each $a \in \Sigma$ is an example of a projective measurement. A measurement of this sort is known more specifically as a *complete projective measurement*. This is the measurement that is commonly referred to as the *measurement with respect to the basis* $\{x_a : a \in \Sigma\}$.

Example 2.41 Let Σ be an alphabet and let $\mathcal{X} = \mathbb{C}^\Sigma$. The *measurement with respect to the standard basis* of $\mathcal{X}$ is the measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ defined as

$$\mu(a) = E_{a,a} \tag{2.239}$$

for each $a \in \Sigma$. For a given state $\rho \in \mathrm{D}(\mathcal{X})$, the probability associated with each measurement outcome $a \in \Sigma$, were this state to be measured according to μ, is equal to the corresponding diagonal entry $\rho(a, a)$. One may also observe that the quantum-to-classical channel associated with this measurement is the completely dephasing channel $\Delta \in \mathrm{C}(\mathcal{X})$.

The following theorem, known as *Naimark's theorem*, establishes a link between arbitrary measurements and projective measurements. It implies that any measurement can be viewed as a projective measurement on a compound register that includes the original register as a subregister.

Theorem 2.42 (Naimark's theorem) *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, let $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be a measurement, and let $\mathcal{Y} = \mathbb{C}^\Sigma$. There exists an isometry $A \in \mathrm{U}(\mathcal{X}, \mathcal{X}\otimes\mathcal{Y})$ such that*

$$\mu(a) = A^*(\mathbb{1}_\mathcal{X} \otimes E_{a,a})A \tag{2.240}$$

for every $a \in \Sigma$.

Proof Define an operator $A \in \mathrm{L}(\mathcal{X}, \mathcal{X}\otimes\mathcal{Y})$ as

$$A = \sum_{a\in\Sigma} \sqrt{\mu(a)} \otimes e_a. \tag{2.241}$$

It holds that

$$A^*A = \sum_{a\in\Sigma} \mu(a) = \mathbb{1}_\mathcal{X}, \tag{2.242}$$

and therefore A is an isometry. The required equation (2.240) holds for each $a \in \Sigma$, so the proof is complete. □

Corollary 2.43 *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, and let $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be a measurement. Also let $\mathcal{Y} = \mathbb{C}^\Sigma$ and let $u \in \mathcal{Y}$ be a unit vector. There exists a projective measurement $\nu : \Sigma \to \mathrm{Pos}(\mathcal{X}\otimes\mathcal{Y})$ such that*

$$\langle \nu(a), X \otimes uu^*\rangle = \langle \mu(a), X\rangle \tag{2.243}$$

for every $X \in \mathrm{L}(\mathcal{X})$.

Proof Let $A \in \mathrm{U}(\mathcal{X}, \mathcal{X}\otimes\mathcal{Y})$ be the isometry whose existence is implied by Theorem 2.42. Choose $U \in \mathrm{U}(\mathcal{X}\otimes\mathcal{Y})$ to be any unitary operator for which the equation

$$U(\mathbb{1}_\mathcal{X} \otimes u) = A \tag{2.244}$$

is satisfied, and define $\nu : \Sigma \to \mathrm{Pos}(\mathcal{X}\otimes\mathcal{Y})$ as

$$\nu(a) = U^*(\mathbb{1}_\mathcal{X} \otimes E_{a,a})U \tag{2.245}$$

for each $a \in \Sigma$. It holds that ν is a projective measurement, and moreover

$$\begin{aligned} \langle \nu(a), X \otimes uu^*\rangle &= \langle (\mathbb{1}_\mathcal{X} \otimes u^*)U^*(\mathbb{1}_\mathcal{X} \otimes E_{a,a})U(\mathbb{1}_\mathcal{X} \otimes u), X\rangle \\ &= \langle A^*(\mathbb{1}_\mathcal{X} \otimes E_{a,a})A, X\rangle = \langle \mu(a), X\rangle \end{aligned} \tag{2.246}$$

for each $a \in \Sigma$, as required. □

Information-complete measurements

States of registers are uniquely determined by the measurement statistics they generate. More precisely, the knowledge of the probability associated with every outcome of every measurement that could be performed on a given register is sufficient to obtain a description of that register's state. In fact, something stronger may be said, which is that there exist choices of measurements that uniquely determine every possible state of a register by the measurement statistics that they alone generate. Such measurements, which are known as *information-complete measurements*, are characterized by the property that their measurement operators span the entire space of operators from which they are drawn.

In more explicit terms, a measurement $\mu : \Sigma \to \text{Pos}(\mathcal{X})$ on a complex Euclidean space $\mathcal{X}$ is said to be an *information-complete measurement* if it holds that

$$\text{span}\{\mu(a) \,:\, a \in \Sigma\} = \text{L}(\mathcal{X}). \tag{2.247}$$

For any such measurement, and for any choice of $\rho \in \text{D}(\mathcal{X})$, it holds that the probability vector $p \in \mathcal{P}(\Sigma)$ defined by $p(a) = \langle \mu(a), \rho \rangle$ uniquely determines the state ρ. This fact is evident from the following proposition.

Proposition 2.44 *Let Σ be an alphabet, let $\mathcal{X}$ be a complex Euclidean space, and let $\{A_a \,:\, a \in \Sigma\} \subset \text{L}(\mathcal{X})$ be a collection of operators for which*

$$\text{span}\{A_a \,:\, a \in \Sigma\} = \text{L}(\mathcal{X}). \tag{2.248}$$

The mapping $\phi : \text{L}(\mathcal{X}) \to \mathbb{C}^{\Sigma}$ defined by

$$(\phi(X))(a) = \langle A_a, X \rangle, \tag{2.249}$$

for each $X \in \text{L}(\mathcal{X})$ and $a \in \Sigma$, is an injective mapping.

Proof Let $X, Y \in \text{L}(\mathcal{X})$ satisfy $\phi(X) = \phi(Y)$, so that

$$\langle A_a, X - Y \rangle = 0 \tag{2.250}$$

for every $a \in \Sigma$. As $\{A_a \,:\, a \in \Sigma\}$ spans $\text{L}(\mathcal{X})$, it follows by the conjugate linearity of the inner product that

$$\langle Z, X - Y \rangle = 0 \tag{2.251}$$

for every $Z \in \text{L}(\mathcal{X})$, and consequently $X - Y = 0$, which completes the proof. □

The following example provides one way of constructing information-complete measurements, for any choice of a complex Euclidean space.

Example 2.45 Let Σ be an alphabet, let $\mathcal{X} = \mathbb{C}^{\Sigma}$, and let

$$\{\rho_{a,b} \,:\, (a,b) \in \Sigma \times \Sigma\} \subseteq \text{D}(\mathcal{X}) \tag{2.252}$$

be a collection of density operators that spans all of $\text{L}(\mathcal{X})$. One such set was constructed in Example 2.7. Also define

$$Q = \sum_{(a,b) \in \Sigma \times \Sigma} \rho_{a,b} \tag{2.253}$$

and observe that Q is necessarily positive definite; if this were not so, there would exist a nonzero vector $u \in \mathcal{X}$ satisfying $\langle \rho_{a,b}, uu^* \rangle = 0$ for each pair $(a,b) \in \Sigma \times \Sigma$, in contradiction with Proposition 2.44. It may be verified that the function $\mu : \Sigma \times \Sigma \to \text{Pos}(\mathcal{X})$, defined by

$$\mu(a,b) = Q^{-\frac{1}{2}} \rho_{a,b} Q^{-\frac{1}{2}} \tag{2.254}$$

for each $(a,b) \in \Sigma \times \Sigma$, is an information-complete measurement.

Nondestructive measurements and instruments

It is convenient in some situations to consider an alternative definition of measurements that does not dictate that registers are destroyed upon being measured. Instead, a measured register is left in some particular state that depends both on its initial state and on the measurement outcome obtained. More generally, one may consider that the measured register is transformed into another register as a result of the measurement process.

One specific alternative definition, which is frequently taken as the definition of a measurement by other authors, describes such a process by a collection

$$\{M_a \,:\, a \in \Sigma\} \subset \text{L}(\mathcal{X}), \tag{2.255}$$

where Σ is the alphabet of measurement outcomes and $\mathcal{X}$ is the complex Euclidean space corresponding to the register being measured, such that the constraint

$$\sum_{a \in \Sigma} M_a^* M_a = \mathbb{1}_{\mathcal{X}} \tag{2.256}$$

is satisfied. When this form of measurement is applied to a register X in a given state $\rho \in \text{D}(\mathcal{X})$, two things happen:

1. An element of Σ is selected at random, with each outcome $a \in \Sigma$ being obtained with probability $\langle M_a^* M_a, \rho \rangle$.

2. Conditioned on the measurement outcome $a \in \Sigma$ having been obtained, the state of the register X becomes

$$\frac{M_a \rho M_a^*}{\langle M_a^* M_a, \rho \rangle}. \tag{2.257}$$

Measurements of this sort will be referred to as *nondestructive measurements* in this book.

A somewhat more general notion of a measurement is described by a collection

$$\{\Phi_a : a \in \Sigma\} \subset \mathrm{CP}(\mathcal{X}, \mathcal{Y}), \tag{2.258}$$

where Σ is the measurement outcome alphabet, $\mathcal{X}$ is the complex Euclidean space corresponding to the register that is measured, and $\mathcal{Y}$ is an arbitrary complex Euclidean space. In this case, these mappings must necessarily sum to a channel:

$$\sum_{a \in \Sigma} \Phi_a \in \mathrm{C}(\mathcal{X}, \mathcal{Y}). \tag{2.259}$$

When this form of measurement is applied to a register X in a given state $\rho \in \mathrm{D}(\mathcal{X})$, two things happen:

1. An element of Σ is selected at random, with each outcome $a \in \Sigma$ being obtained with probability $\mathrm{Tr}(\Phi_a(\rho))$.
2. Conditioned on the measurement outcome $a \in \Sigma$ having been obtained, X is transformed into a new register Y having state

$$\frac{\Phi_a(\rho)}{\mathrm{Tr}(\Phi_a(\rho))}. \tag{2.260}$$

The generalized notion of a measurement obtained in this way is called an *instrument* (or a *quantum instrument*). Nondestructive measurements of the form (2.255) may be represented by instruments of the form (2.258) by defining

$$\Phi_a(X) = M_a X M_a^* \tag{2.261}$$

for each $a \in \Sigma$.

Processes that are expressible as instruments, including nondestructive measurements, can alternatively be described as compositions of channels and (ordinary) measurements. Specifically, for a given instrument of the form (2.258), one may introduce a (classical) register Z having classical state set Σ, and define a channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Z} \otimes \mathcal{Y})$ as

$$\Phi(X) = \sum_{a \in \Sigma} E_{a,a} \otimes \Phi_a(X) \tag{2.262}$$

for every $X \in \mathrm{L}(\mathcal{X})$. The fact that Φ is indeed a channel follows directly from the constraints placed on a function of the form (2.258) that must be satisfied for it to be considered an instrument: the complete positivity of the collection of mappings $\{\Phi_a : a \in \Sigma\}$ implies that Φ is completely positive, while the condition (2.259) implies that Φ preserves trace.

Now, if such a channel Φ is applied to a register X, and then the register Z is measured with respect to the standard basis of $\mathcal{Z}$, the distribution of measurement outcomes, as well as the corresponding state of Y conditioned on each possible outcome, is identical to the process associated with the instrument (2.258), as described above.

2.3.3 Extremal measurements and ensembles

Measurements and ensembles may be regarded as elements of convex sets in a fairly straightforward way. A characterization of the extreme points of these sets is obtained below.

Convex combinations of measurements

For $\mathcal{X}$ being a complex Euclidean space and Σ being an alphabet, one may take convex combinations of measurements of the form $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ in the following way. For an alphabet Γ, a probability vector $p \in \mathcal{P}(\Gamma)$, and a collection $\{\mu_b : b \in \Gamma\}$ of measurements taking the form $\mu_b : \Sigma \to \mathrm{Pos}(\mathcal{X})$ for each $b \in \Gamma$, one defines the measurement

$$\mu = \sum_{b \in \Gamma} p(b) \mu_b \tag{2.263}$$

by the equation

$$\mu(a) = \sum_{b \in \Gamma} p(b) \mu_b(a) \tag{2.264}$$

holding for all $a \in \Sigma$. Equivalently, such a convex combination is taken with respect to the most straightforward way of regarding the set of all functions of the form $\theta : \Sigma \to \mathrm{Herm}(\mathcal{X})$ as a vector space over the real numbers.

An equivalent description of this notion may be obtained through the identification of each measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ with its corresponding quantum-to-classical channel

$$\Phi_\mu(X) = \sum_{a \in \Sigma} \langle \mu(a), X \rangle E_{a,a}. \tag{2.265}$$

Convex combinations of measurements then correspond to ordinary convex combinations of their associated channels.

The measurement described by the convex combination (2.263) may be viewed as being equivalent to a process whereby $b \in \Gamma$ is chosen according to the probability vector p, and the measurement μ_b is performed for the chosen symbol $b \in \Gamma$. The outcome of the measurement μ_b is taken as the output of the new measurement, while the symbol $b \in \Gamma$ is discarded.

Extremal measurements

As was established by Proposition 2.38, the set of all quantum-to-classical channels is compact and convex. A measurement is said to be an *extremal measurement* if its corresponding quantum-to-classical channel corresponds to an extreme point of this set. The definition below states this condition in concrete terms. A characterization of extremal measurements is provided by the theorem that follows.

Definition 2.46 Let Σ be an alphabet and let $\mathcal{X}$ be a complex Euclidean space. A measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ is an *extremal measurement* if, for all choices of measurements μ_0, $\mu_1 : \Sigma \to \mathrm{Pos}(\mathcal{X})$ satisfying $\mu = \lambda\mu_0 + (1-\lambda)\mu_1$ for some real number $\lambda \in (0,1)$, one has $\mu_0 = \mu_1$.

Theorem 2.47 *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, and let $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be a measurement. It holds that μ is an extremal measurement if and only if, for every function $\theta : \Sigma \to \mathrm{Herm}(\mathcal{X})$ satisfying*

$$\sum_{a\in\Sigma} \theta(a) = 0 \tag{2.266}$$

and $\mathrm{im}(\theta(a)) \subseteq \mathrm{im}(\mu(a))$ for every $a \in \Sigma$, one necessarily has that θ is identically zero: $\theta(a) = 0$ for each $a \in \Sigma$.

Proof The theorem will be proved in the contrapositive form. Assume first that μ is not an extremal measurement, so there exist distinct measurements $\mu_0, \mu_1 : \Sigma \to \mathrm{Pos}(\mathcal{X})$ and a scalar value $\lambda \in (0,1)$ for which

$$\mu = \lambda\mu_0 + (1-\lambda)\mu_1. \tag{2.267}$$

It follows that distinct measurements $\nu_0, \nu_1 : \Sigma \to \mathrm{Pos}(\mathcal{X})$ exist for which

$$\mu = \frac{\nu_0 + \nu_1}{2}. \tag{2.268}$$

In particular, one may set

$$\begin{aligned} &\nu_0 = 2\lambda\mu_0 + (1-2\lambda)\mu_1 \quad \text{and} \quad \nu_1 = \mu_1, && \text{if } \lambda \leq 1/2; \\ &\nu_0 = \mu_0 \quad \text{and} \quad \nu_1 = (2\lambda-1)\mu_0 + (2-2\lambda)\mu_1, && \text{if } \lambda \geq 1/2. \end{aligned} \tag{2.269}$$

Define $\theta : \Sigma \to \mathrm{Herm}(\mathcal{X})$ as $\theta(a) = \nu_0(a) - \nu_1(a)$ for each $a \in \Sigma$. It holds that

$$\sum_{a\in\Sigma} \theta(a) = \sum_{a\in\Sigma} \nu_0(a) - \sum_{a\in\Sigma} \nu_1(a) = \mathbb{1}_{\mathcal{X}} - \mathbb{1}_{\mathcal{X}} = 0. \tag{2.270}$$

Moreover,

$$\mathrm{im}(\theta(a)) \subseteq \mathrm{im}(\nu_0(a)) + \mathrm{im}(\nu_1(a)) = \mathrm{im}(\mu(a)) \tag{2.271}$$

for each $a \in \Sigma$, where the equality is a consequence of the facts that $\nu_0(a)$ and $\nu_1(a)$ are positive semidefinite and $\mu(a) = (\nu_0(a) + \nu_1(a))/2$. Finally, given that ν_0 and ν_1 are distinct, it is not the case that θ is identically zero.

Now assume that $\theta : \Sigma \to \mathrm{Herm}(\mathcal{X})$ is a function that is not identically zero, and that satisfies

$$\sum_{a\in\Sigma} \theta(a) = 0 \tag{2.272}$$

and $\mathrm{im}(\theta(a)) \subseteq \mathrm{im}(\mu(a))$ for every $a \in \Sigma$. For each $a \in \Sigma$, there must exist a positive real number $\varepsilon_a > 0$ for which

$$\mu(a) + \varepsilon_a\theta(a) \geq 0 \quad \text{and} \quad \mu(a) - \varepsilon_a\theta(a) \geq 0, \tag{2.273}$$

by virtue of the fact that $\mu(a)$ is positive semidefinite and $\theta(a)$ is a Hermitian operator with $\mathrm{im}(\theta(a)) \subseteq \mathrm{im}(\mu(a))$. Let

$$\varepsilon = \min\{\varepsilon_a \,:\, a \in \Sigma\} \tag{2.274}$$

and define

$$\mu_0 = \mu - \varepsilon\theta \quad \text{and} \quad \mu_1 = \mu + \varepsilon\theta. \tag{2.275}$$

It is evident that $\mu = (\mu_0 + \mu_1)/2$. As θ is not identically zero and ε is positive, it holds that μ_0 and μ_1 are distinct. Finally, it holds that μ_0 and μ_1 are measurements: the assumption (2.272) implies that

$$\sum_{a\in\Sigma} \mu_0(a) = \sum_{a\in\Sigma} \mu_1(a) = \sum_{a\in\Sigma} \mu(a) = \mathbb{1}_{\mathcal{X}} \tag{2.276}$$

while the inequalities (2.273) imply that the measurement operators $\mu_0(a)$ and $\mu_1(a)$ are positive semidefinite for each $a \in \Sigma$. It has therefore been established that μ is not an extremal measurement, which completes the proof. □

Theorem 2.47 has various implications, including the corollaries below. The first corollary makes the observation that extremal measurements can have at most $\dim(\mathcal{X})^2$ nonzero measurement operators.

Corollary 2.48 *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, and let $\mu : \Sigma \to \text{Pos}(\mathcal{X})$ be a measurement. If μ is an extremal measurement, then*

$$\big|\{a \in \Sigma : \mu(a) \neq 0\}\big| \leq \dim(\mathcal{X})^2. \tag{2.277}$$

Proof The corollary will be proved in the contrapositive form. Let

$$\Gamma = \{a \in \Sigma : \mu(a) \neq 0\}, \tag{2.278}$$

assume that $|\Gamma| > \dim(\mathcal{X})^2$, and consider the collection of measurement operators $\{\mu(a) : a \in \Gamma\}$ as a subset of the real vector space $\text{Herm}(\mathcal{X})$. By the assumption $|\Gamma| > \dim(\mathcal{X})^2$, it must hold that the set $\{\mu(a) : a \in \Gamma\}$ is linearly dependent, and therefore there exist real numbers $\{\alpha_a : a \in \Gamma\}$, not all of which are zero, so that

$$\sum_{a \in \Gamma} \alpha_a \mu(a) = 0. \tag{2.279}$$

Define a function $\theta : \Sigma \to \text{Herm}(\mathcal{X})$ as

$$\theta(a) = \begin{cases} \alpha_a \mu(a) & \text{if } a \in \Gamma \\ 0 & \text{if } a \notin \Gamma. \end{cases} \tag{2.280}$$

It holds that θ is not identically zero, and satisfies

$$\sum_{a \in \Sigma} \theta(a) = 0 \tag{2.281}$$

and $\text{im}(\theta(a)) \subseteq \text{im}(\mu(a))$ for every $a \in \Sigma$. By Theorem 2.47, measurement μ is therefore not an extremal measurement, which completes the proof. □

Corollary 2.48, together with Proposition 2.38 and Theorem 1.10, implies the following corollary.

Corollary 2.49 *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, and let $\mu : \Sigma \to \text{Pos}(\mathcal{X})$ be a measurement. There exists an alphabet Γ, a probability vector $p \in \mathcal{P}(\Gamma)$, and a collection of measurements $\{\mu_b : b \in \Gamma\}$, taking the form $\mu_b : \Sigma \to \text{Pos}(\mathcal{X})$ and satisfying*

$$\big|\{a \in \Sigma : \mu_b(a) \neq 0\}\big| \leq \dim(\mathcal{X})^2 \tag{2.282}$$

for each $b \in \Gamma$, such that

$$\mu = \sum_{b \in \Gamma} p(b) \mu_b. \tag{2.283}$$

For measurements whose measurement operators all have rank equal to one, Theorem 2.47 yields a simple criterion for extremality, as represented by the following corollary.

Corollary 2.50 *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, and let $\{x_a : a \in \Sigma\} \subset \mathcal{X}$ be a collection of nonzero vectors satisfying*

$$\sum_{a \in \Sigma} x_a x_a^* = \mathbb{1}_{\mathcal{X}}. \tag{2.284}$$

The measurement $\mu : \Sigma \to \text{Pos}(\mathcal{X})$ defined by $\mu(a) = x_a x_a^*$ for each $a \in \Sigma$ is an extremal measurement if and only if $\{x_a x_a^* : a \in \Sigma\} \subset \text{Herm}(\mathcal{X})$ is a linearly independent set.

Proof The corollary follows from Theorem 2.47, together with the fact that a Hermitian operator $H \in \text{Herm}(\mathcal{X})$ and a vector $u \in \mathcal{X}$ satisfy the condition $\text{im}(H) \subseteq \text{im}(uu^*)$ if and only if $H = \alpha uu^*$ for some $\alpha \in \mathbb{C}$. □

Another implication of Theorem 2.47 is that projective measurements are necessarily extremal.

Corollary 2.51 *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, and let $\mu : \Sigma \to \text{Pos}(\mathcal{X})$ be a projective measurement. It holds that μ is an extremal measurement.*

Proof Let $\theta : \Sigma \to \text{Herm}(\mathcal{X})$ be a function satisfying

$$\sum_{a \in \Sigma} \theta(a) = 0 \tag{2.285}$$

and $\text{im}(\theta(a)) \subseteq \text{im}(\mu(a))$ for every $a \in \Sigma$. For each $b \in \Sigma$, it therefore holds that

$$\sum_{a \in \Sigma} \mu(b)\theta(a) = 0. \tag{2.286}$$

By Proposition 2.40, the collection $\{\mu(b) : b \in \Sigma\}$ is orthogonal. Therefore, every vector in the image of $\theta(a)$ must be orthogonal to every vector in the image of $\mu(b)$ whenever $a \neq b$, so that

$$\mu(b)\theta(a) = \begin{cases} \theta(a) & \text{if } a = b \\ 0 & \text{if } a \neq b. \end{cases} \tag{2.287}$$

It follows that $\theta(b) = 0$ for every $b \in \Sigma$, and therefore the function θ is identically zero. As this is so for every choice of θ, as described above, it follows from Theorem 2.47 that μ is an extremal measurement. □

Convex combinations of ensembles of states

Convex combinations of ensembles of states may be defined in essentially the same way that convex combinations of measurements are defined. That is, if $\mathcal{X}$ is a complex Euclidean space, Σ and Γ are alphabets, $p \in \mathcal{P}(\Gamma)$ is a probability vector, and

$$\eta_b : \Sigma \to \mathrm{Pos}(\mathcal{X}) \tag{2.288}$$

is an ensemble of states for each $b \in \Gamma$, then the function $\eta : \Sigma \to \mathrm{Pos}(\mathcal{X})$ defined by

$$\eta(a) = \sum_{b \in \Gamma} p(b) \eta_b(a) \tag{2.289}$$

for every $a \in \Sigma$ is also an ensemble. One writes

$$\eta = \sum_{b \in \Gamma} p(b) \eta_b \tag{2.290}$$

in this situation. If a density operator $\rho_b \in \mathrm{D}(\mathcal{X})$, representing the average state of the ensemble η_b, is defined as

$$\rho_b = \sum_{a \in \Sigma} \eta_b(a) \tag{2.291}$$

for each $b \in \Gamma$, then it must hold that the average state of the ensemble η is given by

$$\sum_{a \in \Sigma} \eta(a) = \sum_{b \in \Gamma} p(b) \rho_b. \tag{2.292}$$

It is straightforward consequence of the spectral theorem (as represented by Corollary 1.4) that the extreme points of the set of all ensembles of the form $\eta : \Sigma \to \mathrm{Pos}(\mathcal{X})$ take a simple form; they are the ensembles η that are defined as

$$\eta(a) = \begin{cases} uu^* & \text{if } a = b \\ 0 & \text{if } a \neq b, \end{cases} \tag{2.293}$$

for some choice of a unit vector $u \in \mathcal{X}$ and a symbol $b \in \Sigma$.

In some situations, however, it is appropriate to consider just the subset of ensembles of the form $\eta : \Sigma \to \mathrm{Pos}(\mathcal{X})$ that have a particular average state ρ. This set possesses essentially the same convex structure as the set of measurements of the same form. The following proposition establishes one useful fact along these lines.

Proposition 2.52 *Let $\eta : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be an ensemble, for $\mathcal{X}$ a complex Euclidean space and Σ an alphabet, and let*

$$\rho = \sum_{a \in \Sigma} \eta(a). \tag{2.294}$$

There exists an alphabet Γ and a collection of ensembles $\{\eta_b : b \in \Gamma\}$ taking the form $\eta_b : \Sigma \to \mathrm{Pos}(\mathcal{X})$ so that the following properties are satisfied:

1. *For each $b \in \Gamma$, the average state of η_b is ρ:*
$$\sum_{a \in \Sigma} \eta_b(a) = \rho. \tag{2.295}$$
2. *For each $b \in \Gamma$, it holds that*
$$\big|\{a \in \Sigma \,:\, \eta_b(a) \neq 0\}\big| \leq \mathrm{rank}(\rho)^2. \tag{2.296}$$
3. *The ensemble η is a convex combination of the ensembles $\{\eta_b : b \in \Gamma\}$. Equivalently, it holds that*
$$\eta = \sum_{b \in \Gamma} p(b) \eta_b \tag{2.297}$$
for some choice of a probability vector $p \in \mathcal{P}(\Gamma)$.

Proof Let $\mathcal{Y}$ be a complex Euclidean space satisfying $\dim(\mathcal{Y}) = \mathrm{rank}(\rho)$, and let $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be an operator satisfying $AA^* = \rho$. Such an operator A must necessarily satisfy $\ker(A) = \{0\}$ and $\mathrm{im}(A) = \mathrm{im}(\rho)$. For each $a \in \Sigma$, it holds that

$$\mathrm{im}(\eta(a)) \subseteq \mathrm{im}(\rho) = \mathrm{im}(A). \tag{2.298}$$

By Lemma 2.30, one may therefore conclude that there exists a positive semidefinite operator $Q_a \in \mathrm{Pos}(\mathcal{Y})$ such that

$$\eta(a) = AQ_aA^*, \tag{2.299}$$

for each $a \in \Sigma$.

Now define $\mu : \Sigma \to \mathrm{Pos}(\mathcal{Y})$ as $\mu(a) = Q_a$ for each $a \in \Sigma$. As

$$AA^* = \rho = \sum_{a \in \Sigma} \eta(a) = A\left(\sum_{a \in \Sigma} \mu(a)\right)A^*, \tag{2.300}$$

the fact that $\ker(A) = \{0\}$ implies that

$$\sum_{a \in \Sigma} \mu(a) = \mathbb{1}_{\mathcal{Y}}, \tag{2.301}$$

and therefore μ is a measurement.

By Corollary 2.49, there exists an alphabet Γ, a collection of measurements $\{\mu_b : b \in \Gamma\}$ taking the form $\mu_b : \Sigma \to \mathrm{Pos}(\mathcal{Y})$ and satisfying

$$\big|\{a \in \Sigma : \mu_b(a) \neq 0\}\big| \leq \dim(\mathcal{Y})^2 \tag{2.302}$$

for each $b \in \Gamma$, and a probability vector $p \in \mathcal{P}(\Gamma)$, such that

$$\mu = \sum_{b \in \Gamma} p(b)\mu_b. \tag{2.303}$$

Define a function $\eta_b : \Sigma \to \mathrm{Pos}(\mathcal{X})$ for each $b \in \Gamma$ as

$$\eta_b(a) = A\mu_b(a)A^* \tag{2.304}$$

for each $a \in \Sigma$. It is evident that each η_b is an ensemble whose average state is ρ, by virtue of the fact that each μ_b is a measurement, and the requirement (2.296) follows directly from (2.302). Finally, one has

$$\sum_{b \in \Gamma} p(b)\eta_b(a) = A\left(\sum_{b \in \Gamma} p(b)\mu_b(a)\right)A^* = A\mu(a)A^* = \eta(a) \tag{2.305}$$

for each $a \in \Sigma$, and therefore (2.297) holds, which completes the proof. □

2.4 Exercises

Exercise 2.1 Let Σ be an alphabet, let $\mathcal{X}$ be a complex Euclidean space, and let $\phi : \mathrm{Herm}(\mathcal{X}) \to \mathbb{R}^\Sigma$ be a linear function. Prove that these two statements are equivalent:

1. It holds that $\phi(\rho) \in \mathcal{P}(\Sigma)$ for every density operator $\rho \in \mathrm{D}(\mathcal{X})$.
2. There exists a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ such that

$$(\phi(H))(a) = \langle \mu(a), H \rangle \tag{2.306}$$

for every $H \in \mathrm{Herm}(\mathcal{X})$ and $a \in \Sigma$.

Exercise 2.2 Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let Σ be an alphabet, and let $\eta : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be an ensemble of states. Suppose further that $u \in \mathcal{X} \otimes \mathcal{Y}$ is a vector such that

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = \sum_{a \in \Sigma} \eta(a). \tag{2.307}$$

Prove that there exists a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{Y})$ such that

$$\eta(a) = \mathrm{Tr}_{\mathcal{Y}}\big((\mathbb{1}_{\mathcal{X}} \otimes \mu(a))uu^*\big) \tag{2.308}$$

for all $a \in \Sigma$.

Exercise 2.3 Let $\Phi \in \mathrm{CP}(\mathcal{X}, \mathcal{Y})$ be a nonzero completely positive map, for $\mathcal{X}$ and $\mathcal{Y}$ being complex Euclidean spaces, and let $r = \mathrm{rank}(J(\Phi))$ be the Choi rank of Φ. Prove that there exists a complex Euclidean space $\mathcal{Z}$ having dimension r, along with an operator $A \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Z}, \mathcal{Y})$, such that

$$\Phi(X) = A(X \otimes \mathbb{1}_{\mathcal{Z}})A^* \tag{2.309}$$

for all $X \in \mathrm{L}(\mathcal{X})$. Give a simple equation involving the operator A that is equivalent to Φ preserving trace.

Exercise 2.4 Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a positive map, and let $\Delta \in \mathrm{C}(\mathcal{Y})$ denote the completely dephasing channel with respect to the space $\mathcal{Y}$. Prove that $\Delta\Phi$ is completely positive.

Exercise 2.5 Let $\Phi \in \mathrm{C}(\mathcal{X} \otimes \mathcal{Z}, \mathcal{Y} \otimes \mathcal{W})$ be a channel, for complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$. Prove that the following two statements are equivalent:

1. There exists a channel $\Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ such that

$$\mathrm{Tr}_{\mathcal{W}}(J(\Phi)) = J(\Psi) \otimes \mathbb{1}_{\mathcal{Z}}. \tag{2.310}$$

2. There exists a complex Euclidean space $\mathcal{V}$ with $\dim(\mathcal{V}) \leq \dim(\mathcal{X} \otimes \mathcal{Y})$, along with channels $\Phi_0 \in \mathrm{C}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{V})$ and $\Phi_1 \in \mathrm{C}(\mathcal{V} \otimes \mathcal{Z}, \mathcal{W})$, such that

$$\Phi = (\mathbb{1}_{\mathrm{L}(\mathcal{Y})} \otimes \Phi_1)(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}). \tag{2.311}$$

Exercise 2.6 Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$ be complex Euclidean spaces.

(a) Prove that every operator $P \in \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{X})$ satisfying the equation

$$\langle P, J(\Phi) \rangle = 1 \tag{2.312}$$

for every channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ must take the form

$$P = \mathbb{1}_{\mathcal{Y}} \otimes \rho \tag{2.313}$$

for some choice of $\rho \in \mathrm{D}(\mathcal{X})$.

(b) Let $\Xi \in \mathrm{CP}(\mathcal{Y} \otimes \mathcal{X}, \mathcal{W} \otimes \mathcal{Z})$ be a completely positive map for which the following statement holds: for every channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$, there exists a channel $\Psi \in \mathrm{C}(\mathcal{Z}, \mathcal{W})$ such that

$$\Xi(J(\Phi)) = J(\Psi). \tag{2.314}$$

Prove that there must exist a unital map $\Lambda \in \mathrm{CP}(\mathcal{X}, \mathcal{Z})$ such that

$$\mathrm{Tr}_{\mathcal{W}}(\Xi(X)) = \Lambda(\mathrm{Tr}_{\mathcal{Y}}(X)) \tag{2.315}$$

for all $X \in \mathrm{L}(\mathcal{Y} \otimes \mathcal{X})$.

(c) Let $\Xi \in \mathrm{CP}(\mathcal{Y} \otimes \mathcal{X}, \mathcal{W} \otimes \mathcal{Z})$ be a completely positive map satisfying the same requirement as described in part (b). Prove that there exist channels $\Xi_0 \in \mathrm{C}(\mathcal{Z}, \mathcal{X} \otimes \mathcal{V})$ and $\Xi_1 \in \mathrm{C}(\mathcal{Y} \otimes \mathcal{V}, \mathcal{W})$, for some choice of a complex Euclidean space $\mathcal{V}$, for which the following property holds: for every channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$, the channel $\Psi \in \mathrm{C}(\mathcal{Z}, \mathcal{W})$ that is uniquely determined by (2.314) is given by

$$\Psi = \Xi_1(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{V})})\Xi_0. \tag{2.316}$$

2.5 Bibliographic remarks

The theory of quantum information represents a mathematical formulation of certain aspects of quantum physics, particularly aspects relating to the storage and processing of information in abstract physical systems. While the history of quantum physics is not within the scope of this book, it is appropriate to mention that the mathematical theory discussed in this book is rooted in the work of the many physicists who first developed that field, including Planck, Einstein, Bohr, Heisenberg, Schrödinger, Born, Dirac, and Pauli. Much of this work was placed on a firm mathematical foundation by von Neumann's *Mathematical Foundations of Quantum Mechanics* (1955).

The description of quantum states as density operators was independently proposed by von Neumann (1927b) and Landau (1927), a notion equivalent to that of quantum channels was proposed by Haag and Kastler (1964), and the definition of measurements adopted in this book was proposed by Davies and Lewis (1970). The importance of this definition of measurements was articulated by Holevo (1972, 1973b,c,d); earlier formulations of the theory considered only projective measurements. The books of Helstrom (1976) and Kraus (1983) further refined these key foundational aspects of the theory of quantum information.

Further information on the history of quantum information can be found in the books of Peres (1993), Nielsen and Chuang (2000), and Wilde (2013), which are also indispensable references on the theory itself. Kitaev, Shen, and Vyalyi (2002) and Bengtsson and Życzkowski (2006) also describe the mathematical formalism that has been presented in this chapter, and include discussions of some specific topics connected with quantum information and computation.

The Choi representation is so-named for Choi (1975), who characterized completely positive maps (as represented by the equivalence of statements 1 and 3 in Theorem 2.22). Theorem 2.31 was proved in the same paper. A similar representation to the Choi representation was used earlier by de Pillis (1967) and Jamiołkowski (1972), and there are arguments to be made for the claim that the representation may be considered as folklore.

Theorem 2.22 is an amalgamation of results that are generally attributed to Stinespring (1955), Kraus (1971, 1983), and Choi (1975). Stinespring and Kraus also proved more general results holding for infinite-dimensional spaces; Theorem 2.22 presents only the finite-dimensional analogues of the results they proved. (Several theorems to be presented in this book have a similar character, often having originally been proved in the setting of C*-algebras, as compared with the simpler setting of complex Euclidean spaces.) Theorems 2.25 and 2.26 include equivalences that may be derived from the work of de Pillis (1967) and Jamiołkowski (1972), respectively.

Theorem 2.42 is a simplified variant of a theorem commonly known as Naimark's theorem (or Naimark's dilation theorem). A more general form of this theorem, holding for certain infinite-dimensional spaces and measure-theoretic formulations of measurements having infinitely many outcomes, was proved by Naimark (1943), whose name is sometimes alternatively transliterated as Neumark. This theorem is now commonly viewed as being a direct consequence of the later work of Stinespring mentioned above.

The characterization of extremal measurements given by Theorem 2.47 is equivalent to one obtained by Parthasarathy (1999). Results equivalent to Corollaries 2.48, 2.50, and 2.51 were observed in the same paper. The fact that projective measurements are extremal (Corollary 2.51) was also proved earlier by Holevo (1973d).

Exercise 2.2 is representative of a fact first proved by Hughston, Jozsa, and Wootters (1993). The fact represented by Exercise 2.5 is due to Eggeling, Schlingemann, and Werner (2002), answering a question raised by Beckman, Gottesman, Nielsen, and Preskill (2001) (who credit DiVincenzo for raising the question). Gutoski and Watrous (2007) and Chiribella, D'Ariano, and Perinotti (2009) generalized this result to quantum processes having inputs and outputs that alternate for multiple steps. Exercise 2.6 is representative of a related result of Chiribella, D'Ariano, and Perinotti (2008).