Chapter 4

Unital channels and majorization

This chapter studies the class of *unital channels*, together with the related notion of *majorization* for Hermitian operators. The following definition of unital channels will be used throughout the chapter.

Definition 4.1. Let $\mathcal{X}$ be a complex Euclidean space. A channel $\Phi \in \mathrm{C}(\mathcal{X})$ is a *unital channel* if $\Phi(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{X}}$.[1]

The first section of the chapter introduces various subclasses of unital channels, and the second section investigates properties of unital channels in general. The third section discusses majorization for Hermitian operators, together with an analogous notion for real vectors.

4.1 Subclasses of unital channels

Three subclasses of unital channels are introduced in the subsections that follow: *mixed-unitary channels*, *Weyl-covariant channels*, and *Schur channels*. Various properties of these classes, as well as relationships among them, and to general unital channels, are discussed.

[1] The requirement that unital channels take the form $\Phi \in \mathrm{C}(\mathcal{X})$, for some choice of a complex Euclidean space $\mathcal{X}$, is both natural and convenient with respect to the specific topics to be discussed in this chapter. One could, more generally, consider any channel of the form $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$, for some choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, to be a unital channel if the condition $\Phi(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{Y}}$ is met. In order for this requirement to be met by a channel, it must hold that $\dim(\mathcal{Y}) = \dim(\mathcal{X})$; and for this reason there is little generality lost in restricting the definition of unital channels to those of the form $\Phi \in \mathrm{C}(\mathcal{X})$.

4.1.1 Mixed-unitary channels

Every unitary channel is evidently unital, as is any convex combination of unitary channels. Channels of the later sort will be referred to as *mixed-unitary* channels, as the following definition makes precise.

Definition 4.2. Let $\mathcal{X}$ be a complex Euclidean space and let $\Phi \in \mathrm{C}(\mathcal{X})$ be a channel. It is said that Φ is a *mixed-unitary channel* if there exists an alphabet Σ, a probability vector $p \in \mathcal{P}(\Sigma)$, and a collection of unitary operators $\{U_a : a \in \Sigma\} \subset \mathrm{U}(\mathcal{X})$ such that

$$\Phi(X) = \sum_{a \in \Sigma} p(a) U_a X U_a^* \tag{4.1}$$

for every $X \in \mathrm{L}(\mathcal{X})$. Equivalently, a mapping $\Phi \in \mathrm{C}(\mathcal{X})$ is a mixed-unitary channel if it is a convex combination of unitary channels.

An example of a unital channel that is not mixed-unitary

While every mixed-unitary channel is necessarily unital, the converse of this statement does not hold, as the following example illustrates.

Example 4.3. Let $\mathcal{X} = \mathbb{C}^3$ and define $\Phi \in \mathrm{C}(\mathcal{X})$ as

$$\Phi(X) = \frac{1}{2}\operatorname{Tr}(X)\mathbb{1} - \frac{1}{2}X^{\mathsf{T}} \tag{4.2}$$

for all $X \in \mathrm{L}(\mathcal{X})$. Example 3.39 has established that Φ is a channel, and it is evident that Φ is unital. The channel Φ is, however, not a mixed-unitary channel. To see this, observe first that

$$\Phi(X) = A_1 X A_1^* + A_2 X A_2^* + A_3 X A_3^* \tag{4.3}$$

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

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{-1}{\sqrt{2}} & 0 \end{pmatrix}, \; A_2 = \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 \\ \frac{-1}{\sqrt{2}} & 0 & 0 \end{pmatrix}, \; A_3 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{-1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.4}$$

The fact that the expression (4.3) does indeed hold for all $X \in \mathrm{L}(\mathcal{X})$ follows from the observation that the Choi representation of the map defined by the

right-hand side of that equation is in agreement with $J(\Phi)$, as calculated in Example 3.39:

$$\frac{1}{2}\mathbb{1}\otimes\mathbb{1} - \frac{1}{2}W = \sum_{k=1}^{3} \operatorname{vec}(A_k)\operatorname{vec}(A_k)^*. \tag{4.5}$$

The collection $\{A_j^* A_k : 1 \leq j,k \leq 3\}$ includes the following operators:

$$A_1^*A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad A_1^*A_2 = \begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; A_1^*A_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{-1}{2} & 0 & 0 \end{pmatrix},$$

$$A_2^*A_1 = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2^*A_2 = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \; A_2^*A_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}, \tag{4.6}$$

$$A_3^*A_1 = \begin{pmatrix} 0 & 0 & \frac{-1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; A_3^*A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}, \; A_3^*A_3 = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This is a linearly independent collection, as an inspection reveals. It follows from Theorem 2.31 that Φ is an extreme point of the set of channels $\mathrm{C}(\mathcal{X})$. As Φ is not itself a unitary channel, it follows that it cannot be expressed as a convex combination of unitary channels.

Pinching channels

Many interesting examples of mixed-unitary channels are known. One type of channel, called a *pinching channel*, provides a collection of examples.

Definition 4.4. Let $\mathcal{X}$ be a complex Euclidean space. A channel $\Phi \in \mathrm{C}(\mathcal{X})$ is said to be a *pinching channel*, or simply a *pinching*, if there exists a collection $\{\Pi_a : a \in \Sigma\}$ of projection operators satisfying

$$\sum_{a\in\Sigma} \Pi_a = \mathbb{1}_{\mathcal{X}} \tag{4.7}$$

(i.e., such that the set $\{\Pi_a : a \in \Sigma\}$ represents a projective measurement) for which

$$\Phi(X) = \sum_{a\in\Sigma} \Pi_a X \Pi_a \tag{4.8}$$

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

The action of the channel defined by (4.8) on a register X is equivalent to X being measured with respect to a nondestructive measurement defined by $\{\Pi_a : a \in \Sigma\}$, followed by the measurement outcome being discarded.

Example 4.5. The channel $\Phi \in \mathrm{C}(\mathbb{C}^5)$ defined as

$$\Phi(X) = \Pi_0 X \Pi_0 + \Pi_1 X \Pi_1 \tag{4.9}$$

for

$$\Pi_0 = E_{1,1} + E_{2,2} \qquad \text{and} \qquad \Pi_1 = E_{3,3} + E_{4,4} + E_{5,5} \tag{4.10}$$

is an example of a pinching channel. This channel has the following action on a general operator in $\mathrm{L}(\mathcal{X})$, expressed in matrix form:

$$\Phi \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} & \alpha_{1,4} & \alpha_{1,5} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} & \alpha_{2,4} & \alpha_{2,5} \\ \alpha_{3,1} & \alpha_{3,2} & \alpha_{3,3} & \alpha_{3,4} & \alpha_{3,5} \\ \alpha_{4,1} & \alpha_{4,2} & \alpha_{4,3} & \alpha_{4,4} & \alpha_{4,5} \\ \alpha_{5,1} & \alpha_{5,2} & \alpha_{5,3} & \alpha_{5,4} & \alpha_{5,5} \end{pmatrix} = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & 0 & 0 & 0 \\ \alpha_{2,1} & \alpha_{2,2} & 0 & 0 & 0 \\ 0 & 0 & \alpha_{3,3} & \alpha_{3,4} & \alpha_{3,5} \\ 0 & 0 & \alpha_{4,3} & \alpha_{4,4} & \alpha_{4,5} \\ 0 & 0 & \alpha_{5,3} & \alpha_{5,4} & \alpha_{5,5} \end{pmatrix}. \tag{4.11}$$

The action of this channel is suggestive of the matrix representing the input operator being "pinched," causing a certain pattern of off-diagonal entries to become 0, which explains the terminology used to describe such maps. When a pinching channel is defined by a collection of projection operators that are not diagonal in the standard basis, the term is not descriptive in this way, but it is used nevertheless.

While it is not immediate from the definition that every pinching channel is a mixed-unitary channel, it is fairly straightforward to establish that this is so, as the proof of the following proposition reveals.

Proposition 4.6. *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, and let $\{\Pi_a : a \in \Sigma\}$ be a collection of projection operators on $\mathcal{X}$ satisfying*

$$\sum_{a \in \Sigma} \Pi_a = \mathbb{1}_{\mathcal{X}}. \tag{4.12}$$

The channel $\Phi \in \mathrm{C}(\mathcal{X})$ defined by

$$\Phi(X) = \sum_{a \in \Sigma} \Pi_a X \Pi_a \tag{4.13}$$

for all $X \in \mathrm{L}(\mathcal{X})$ is a mixed-unitary channel.

Proof. Consider the collection $\{-1,1\}^{\Sigma}$ containing those vectors in $\mathcal{X}$ having entries drawn from the set $\{-1,1\}$, and define a unitary operator

$$U_w = \sum_{a\in\Sigma} w(a)\Pi_a \tag{4.14}$$

for every such vector $w \in \{-1,1\}^{\Sigma}$. It holds that

$$\frac{1}{2^{|\Sigma|}} \sum_{w\in\{-1,1\}^{\Sigma}} U_w X U_w^* = \frac{1}{2^{|\Sigma|}} \sum_{a,b\in\Sigma} \sum_{w\in\{-1,1\}^{\Sigma}} w(a)w(b)\Pi_a X \Pi_b \tag{4.15}$$

for every $X \in \mathrm{L}(\mathcal{X})$. To simplify this expression, one may observe that

$$\frac{1}{2^{|\Sigma|}} \sum_{w\in\{-1,1\}^{\Sigma}} w(a)w(b) = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b \end{cases} \tag{4.16}$$

for every choice of $a, b \in \Sigma$, and therefore

$$\frac{1}{2^{|\Sigma|}} \sum_{w\in\{-1,1\}^{\Sigma}} U_w X U_w^* = \sum_{a\in\Sigma} \Pi_a X \Pi_a = \Phi(X) \tag{4.17}$$

for every $X \in \mathrm{L}(\mathcal{X})$. This demonstrates that Φ is a mixed-unitary channel, as required. □

Example 4.7. The completely dephasing channel $\Delta \in \mathrm{C}(\mathcal{X})$ defined on any complex Euclidean space $\mathcal{X} = \mathbb{C}^{\Sigma}$ is an example of a pinching channel, as it is defined according to Definition 4.4 by the collection of projection operators $\{E_{a,a} : a \in \Sigma\}$. By Proposition 4.6, it follows that Δ is a mixed-unitary channel.

Environment-assisted channel correction

Mixed-unitary channels have an alternative characterization based on the notion of *environment-assisted channel correction*. Let $\Phi \in \mathrm{C}(\mathcal{X})$ be a channel, represented in Stinespring form as

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}(AXA^*) \tag{4.18}$$

for all $X \in \mathrm{L}(\mathcal{X})$, for some choice of a complex Euclidean space $\mathcal{Z}$ and an isometry $A \in \mathrm{U}(\mathcal{X}, \mathcal{X}\otimes\mathcal{Z})$. Environment-assisted channel correction refers to the existence of an alphabet Σ, a collection of channels

$$\{\Psi_a : a \in \Sigma\} \subset \mathrm{C}(\mathcal{X}), \tag{4.19}$$

and a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{Z})$, for which the equation

$$X = \sum_{a \in \Sigma} \Psi_a\big(\mathrm{Tr}_{\mathcal{Z}}\big((\mathbb{1}_{\mathcal{X}} \otimes \mu(a))AXA^*\big)\big) \tag{4.20}$$

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

An interpretation of the equation (4.20) is as follows. One imagines that a register X contains a quantum state $\rho \in \mathrm{D}(\mathcal{X})$. The action of the mapping $X \mapsto AXA^*$ has the effect of encoding this state into the state of the pair (X, Z), for Z being a second register. By discarding the register Z, the register X is left in the state $\Phi(\rho)$, which may potentially be quite different from ρ. In essence, the register Z represents an "environment," to which some part of the encoding of ρ may have escaped or leaked. The measurement μ on Z, followed by the application of Ψ_a to X (for whichever outcome $a \in \Sigma$ resulted from the measurement), is viewed as an attempt to *correct* X, so that it is transformed back into ρ. The equation (4.20) represents the situation in which a perfect correction of this sort is accomplished.

The following theorem implies that a perfect correction of the sort just described is possible if and only if Φ is a mixed-unitary channel.

Theorem 4.8. *Let $A \in \mathrm{U}(\mathcal{X}, \mathcal{X} \otimes \mathcal{Z})$ be an isometry, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Z}$, and let $\Phi \in \mathrm{C}(\mathcal{X})$ be the channel defined by*

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(AXA^*\big) \tag{4.21}$$

for all $X \in \mathrm{L}(\mathcal{X})$. The following two are statements equivalent:

1. *Φ is a mixed-unitary channel.*
2. *There exists an alphabet Σ, a collection of channels $\{\Psi_a : a \in \Sigma\} \subset \mathrm{C}(\mathcal{X})$, and a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{Z})$ for which*

$$X = \sum_{a \in \Sigma} \Psi_a\big(\mathrm{Tr}_{\mathcal{Z}}\big((\mathbb{1}_{\mathcal{X}} \otimes \mu(a))AXA^*\big)\big) \tag{4.22}$$

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

Proof. Assume first that statement 1 holds, so that

$$\Phi(X) = \sum_{a \in \Sigma} p(a) U_a X U_a^* \tag{4.23}$$

for every $X \in \mathrm{L}(\mathcal{X})$, for some choice of an alphabet Σ, a collection of unitary operators $\{U_a : a \in \Sigma\} \subset \mathrm{U}(\mathcal{X})$, and a probability vector $p \in \mathcal{P}(\Sigma)$. There

is no loss of generality in assuming $|\Sigma| \geq \dim(\mathcal{Z})$; one may add any finite number of elements to Σ, take $p(a) = 0$, and choose $U_a \in \mathrm{U}(\mathcal{X})$ arbitrarily for the added elements, maintaining the validity of the expression (4.23). By this assumption, there must exist a collection $\{v_a : a \in \Sigma\} \subset \mathcal{Z}$ of vectors for which

$$\sum_{a \in \Sigma} v_a v_a^* = \mathbb{1}_{\mathcal{Z}}. \tag{4.24}$$

Fix such collection, and define operators $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X})$ as

$$A_a = (\mathbb{1}_{\mathcal{X}} \otimes v_a^*)A \tag{4.25}$$

for each $a \in \Sigma$. It holds that

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(AXA^*\big) = \sum_{a \in \Sigma} A_a X A_a^* \tag{4.26}$$

for every $X \in \mathrm{L}(\mathcal{X})$. Therefore, by Corollary 2.23, there must exist a unitary operator $W \in \mathrm{U}(\mathbb{C}^{\Sigma})$ such that

$$\sqrt{p(a)}U_a = \sum_{b \in \Sigma} W(a,b) A_b \tag{4.27}$$

for every $a \in \Sigma$.

For each symbol $a \in \Sigma$, define a vector $u_a \in \mathcal{Z}$ as

$$u_a = \sum_{b \in \Sigma} \overline{W(a,b)} v_b, \tag{4.28}$$

and define $\mu : \Sigma \to \mathrm{Pos}(\mathcal{Z})$ as $\mu(a) = u_a u_a^*$ for each $a \in \Sigma$. Because W is a unitary operator, it holds that

$$\sum_{a \in \Sigma} \mu(a) = \sum_{a,b,c \in \Sigma} \overline{W(a,b)} W(a,c) v_b v_c^* = \sum_{b \in \Sigma} v_b v_b^* = \mathbb{1}_{\mathcal{Z}}\,, \tag{4.29}$$

and therefore μ is a measurement. Also define a collection $\{\Psi_a : a \in \Sigma\}$ of channels as

$$\Psi_a(X) = U_a^* X U_a \tag{4.30}$$

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

Now, it holds that

$$(\mathbb{1}_{\mathcal{X}} \otimes u_a^*)A = \sum_{b \in \Sigma} W(a,b) A_b = \sqrt{p(a)}U_a\,, \tag{4.31}$$

and therefore

$$\mathrm{Tr}_{\mathcal{Z}}\big((\mathbb{1}_{\mathcal{X}} \otimes \mu(a))AXA^*\big) = p(a)U_a X U_a^*, \tag{4.32}$$

for each $a \in \Sigma$. It follows that

$$\sum_{a\in\Sigma} \Psi_a\big(\mathrm{Tr}_{\mathcal{Z}}\big((\mathbb{1}_{\mathcal{X}} \otimes \mu(a))AXA^*\big)\big) = \sum_{a\in\Sigma} p(a)U_a^* U_a X U_a^* U_a = X \tag{4.33}$$

for all $X \in \mathrm{L}(\mathcal{X})$. Statement 1 therefore implies statement 2.

Next, assume statement 2 holds. For each $a \in \Sigma$, define $\Phi_a \in \mathrm{CP}(\mathcal{X})$ as

$$\Phi_a(X) = \mathrm{Tr}_{\mathcal{Z}}\big((\mathbb{1}_{\mathcal{X}} \otimes \mu(a))AXA^*\big) \tag{4.34}$$

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

$$\{A_{a,b} : a \in \Sigma,\, b \in \Gamma\} \quad \text{and} \quad \{B_{a,b} : a \in \Sigma,\, b \in \Gamma\} \tag{4.35}$$

be collections of operators in $\mathrm{L}(\mathcal{X})$, for a suitable choice of an alphabet Γ, yielding Kraus representations

$$\Psi_a(X) = \sum_{b\in\Gamma} A_{a,b} X A_{a,b}^* \quad \text{and} \quad \Phi_a(X) = \sum_{c\in\Gamma} B_{a,c} X B_{a,c}^* \tag{4.36}$$

for all $a \in \Sigma$ and $X \in \mathrm{L}(\mathcal{X})$. (Taking a common alphabet Γ as an index set for these representations is only done to simplify notation and causes no loss of generality; one is free to include the zero operator among the Kraus operators of either map any number of times.) By the assumption that statement 2 holds, one has

$$\sum_{a\in\Sigma} \Psi_a \Phi_a = \mathbb{1}_{\mathrm{L}(\mathcal{X})}, \tag{4.37}$$

and therefore the Choi representations of the two sides equation (4.37) must agree:

$$\sum_{a\in\Sigma} \sum_{b,c\in\Gamma} \mathrm{vec}(A_{a,b}B_{a,c})\,\mathrm{vec}(A_{a,b}B_{a,c})^* = \mathrm{vec}(\mathbb{1}_{\mathcal{X}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*. \tag{4.38}$$

There must therefore exist a collection $\{\alpha_{a,b,c} : a \in \Sigma,\, b, c \in \Gamma\}$ of complex numbers for which the equation

$$A_{a,b}B_{a,c} = \alpha_{a,b,c}\mathbb{1}_{\mathcal{X}} \tag{4.39}$$

holds for all $a \in \Sigma$ and $b, c \in \Gamma$. This collection must also evidently satisfy the constraint

$$\sum_{a \in \Sigma} \sum_{b,c \in \Gamma} |\alpha_{a,b,c}|^2 = 1. \tag{4.40}$$

Consequently, one has

$$\sum_{b \in \Gamma} |\alpha_{a,b,c}|^2 \mathbb{1}_{\mathcal{X}} = \sum_{b \in \Gamma} B_{a,c}^* A_{a,b}^* A_{a,b} B_{a,c} = B_{a,c}^* B_{a,c} \tag{4.41}$$

for every $a \in \Sigma$ and $c \in \Gamma$, owing to the fact that each mapping Ψ_a is a channel. For every $a \in \Sigma$ and $c \in \Gamma$ it must therefore hold that

$$B_{a,c} = \beta_{a,c} U_{a,c} \tag{4.42}$$

for some choice of a unitary operator $U_{a,c} \in \mathrm{U}(\mathcal{X})$ and a complex number $\beta_{a,c} \in \mathbb{C}$ satisfying

$$|\beta_{a,c}|^2 = \sum_{b \in \Gamma} |\alpha_{a,b,c}|^2. \tag{4.43}$$

It follows that

$$\Phi(X) = \sum_{a \in \Sigma} \Phi_a(X) = \sum_{a \in \Sigma} \sum_{c \in \Gamma} p(a,c) U_{a,c} X U_{a,c}^*, \tag{4.44}$$

for $p \in \mathcal{P}(\Sigma \times \Gamma)$ being the probability vector defined as

$$p(a,c) = |\beta_{a,c}|^2 \tag{4.45}$$

for each $a \in \Sigma$ and $c \in \Gamma$. The channel Φ is therefore mixed-unitary, so it has been proved that statement 2 implies statement 1. □

Mixed-unitary channels and Carathéodory's theorem

Every mixed-unitary channel $\Phi \in \mathrm{C}(\mathcal{X})$ is, by definition, an element of the convex hull of the set of unitary channels. Using Carathéodory's theorem (Theorem 1.9), one may obtain upper-bounds on the number of unitary channels that must be averaged to obtain any mixed-unitary channel. The following proposition proves one bound along these lines.

Proposition 4.9. *Let $\mathcal{X}$ be a complex Euclidean space, let $n = \dim(\mathcal{X})$, and let $\Phi \in \mathrm{C}(\mathcal{X})$ be a mixed-unitary channel. There exists a positive integer m satisfying*

$$m \leq n^4 - 2n^2 + 2, \tag{4.46}$$

a collection of unitary operators $\{U_1, \ldots, U_m\} \subset \mathrm{U}(\mathcal{X})$*, and a probability vector* $(p_1, \ldots, p_m)$ *such that*

$$\Phi(X) = \sum_{k=1}^{m} p_k U_k X U_k^* \tag{4.47}$$

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

Proof. Consider the linear map $\Xi : \mathrm{Herm}(\mathcal{X} \otimes \mathcal{X}) \to \mathrm{Herm}(\mathcal{X} \oplus \mathcal{X})$ defined by the equation

$$\Xi(X \otimes Y) = \begin{pmatrix} \mathrm{Tr}(X)Y & 0 \\ 0 & \mathrm{Tr}(Y)X \end{pmatrix} \tag{4.48}$$

for all $X, Y \in \mathrm{Herm}(\mathcal{X})$, and fix any orthogonal basis

$$\bigl\{\mathbb{1}, H_1, \ldots, H_{n^2-1}\bigr\} \tag{4.49}$$

of $\mathrm{Herm}(\mathcal{X})$ containing the identity operator. It holds that

$$\Xi(H_j \otimes H_k) = 0 \tag{4.50}$$

for every choice of $j, k \in \{1, \ldots, n^2 - 1\}$, while the operators

$$\Xi(\mathbb{1} \otimes H_k), \quad \Xi(H_k \otimes \mathbb{1}), \quad \text{and} \quad \Xi(\mathbb{1} \otimes \mathbb{1}), \tag{4.51}$$

ranging over all choices of $k \in \{1, \ldots, n^2 - 1\}$, are all nonzero and pairwise orthogonal. The kernel of Ξ is therefore equal to the subspace spanned by the orthogonal collection

$$\bigl\{H_j \otimes H_k \,:\, 1 \leq j, k \leq n^2 - 1\bigr\}. \tag{4.52}$$

In particular, the dimension of the kernel of the mapping Ξ is

$$(n^2 - 1)^2 = n^4 - 2n^2 + 1. \tag{4.53}$$

Next, consider any unitary operator $U \in \mathrm{U}(\mathcal{X})$, and let $\Psi_U \in \mathrm{C}(\mathcal{X})$ be the unitary channel defined as

$$\Psi_U(X) = UXU^* \tag{4.54}$$

for every $X \in \mathrm{L}(\mathcal{X})$. Evaluating the mapping Ξ defined above on the Choi representation of Ψ_U yields

$$\Xi(J(\Psi_U)) = \Xi(\mathrm{vec}(U)\,\mathrm{vec}(U)^*) = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \tag{4.55}$$

The Choi representation of Ψ_U is therefore drawn from an affine subspace of $\mathrm{Herm}(\mathcal{X} \otimes \mathcal{X})$ having dimension $n^4 - 2n^2 + 1$.

Because Φ is a mixed-unitary channel, the Choi representation $J(\Phi)$ of Φ is equal to a convex combination of operators having the form $J(\Psi_U)$, for U ranging over a finite set of unitary operators. It therefore follows from Carathéodory's theorem that

$$J(\Phi) = \sum_{k=1}^{m} p_k J\bigl(\Psi_{U_k}\bigr) \tag{4.56}$$

for some choice of a positive integer

$$m \leq n^4 - 2n^2 + 2, \tag{4.57}$$

unitary operators $U_1, \ldots, U_m \in \mathrm{U}(\mathcal{X})$, and a probability vector $(p_1, \ldots, p_m)$. Equivalently,

$$\Phi(X) = \sum_{k=1}^{m} p_k U_k X U_k^* \tag{4.58}$$

for all $X \in \mathrm{L}(\mathcal{X})$, for the same choice of m, $U_1, \ldots, U_m$, and $(p_1, \ldots, p_m)$, which completes the proof. □

A similar technique to the one used in the proof above may be used to obtain an upper bound on the number of channels, drawn from an arbitrary collection, that must be averaged to obtain a given element in the convex hull of that collection. As a corollary, one obtains a different bound (which is almost always better than the one from the previous proposition) on the number of unitary channels that must be averaged to obtain a given mixed-unitary channel.

Theorem 4.10. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\mathcal{A} \subseteq \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be any nonempty collection of channels, and let $\Phi \in \mathrm{conv}(\mathcal{A})$ be a channel in the convex hull of $\mathcal{A}$. There exists a positive integer*

$$m \leq \mathrm{rank}(J(\Phi))^2, \tag{4.59}$$

a probability vector $(p_1, \ldots, p_m)$, and a selection of channels $\Psi_1, \ldots, \Psi_m \in \mathcal{A}$ such that

$$\Phi = p_1 \Psi_1 + \cdots + p_m \Psi_m. \tag{4.60}$$

Proof. Let $r = \operatorname{rank}(J(\Phi))$ and take Π to be the projection operator onto the image of $J(\Phi)$. Define a linear map

$$\Xi : \operatorname{Herm}(\mathcal{Y} \otimes \mathcal{X}) \to \operatorname{Herm}(\mathbb{C} \oplus (\mathcal{Y} \otimes \mathcal{X}) \oplus (\mathcal{Y} \otimes \mathcal{X})) \tag{4.61}$$

as

$$\Xi(H) = \begin{pmatrix} \operatorname{Tr}(H) & 0 & 0 \\ 0 & (\mathbb{1} - \Pi)H(\mathbb{1} - \Pi) & (\mathbb{1} - \Pi)H\Pi \\ 0 & \Pi H(\mathbb{1} - \Pi) & 0 \end{pmatrix} \tag{4.62}$$

for each $H \in \operatorname{Herm}(\mathcal{Y} \otimes \mathcal{X})$. It holds that $\Xi(H) = 0$ for precisely those Hermitian operators H satisfying

$$H = \Pi H \Pi \quad \text{and} \quad \operatorname{Tr}(H) = 0, \tag{4.63}$$

and therefore the kernel of Ξ has dimension $r^2 - 1$.

Let

$$\mathcal{B} = \{\Psi \in \mathcal{A} : \operatorname{im}(J(\Psi)) \subseteq \operatorname{im}(J(\Phi))\}, \tag{4.64}$$

and observe that $\Phi \in \operatorname{conv}(\mathcal{B})$, by virtue of the fact that $\Phi \in \operatorname{conv}(\mathcal{A})$. For each channel $\Psi \in \mathcal{B}$ it holds that

$$\Xi(J(\Psi)) = \begin{pmatrix} \dim(\mathcal{X}) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.65}$$

There is therefore an affine subspace of $\operatorname{Herm}(\mathcal{Y} \otimes \mathcal{X})$ of dimension $r^2 - 1$ that contains $J(\Psi)$, for every $\Psi \in \mathcal{B}$. As $J(\Phi)$ is a convex combination of operators in this affine subspace, it follows from Carathéodory's theorem that there exists an integer $m \leq (r^2 - 1) + 1 = r^2$, a selection of channels $\Psi_1, \ldots, \Psi_m \in \mathcal{B} \subseteq \mathcal{A}$, and a probability vector $(p_1, \ldots, p_m)$ such that

$$J(\Phi) = p_1 J(\Psi_1) + \cdots + p_m J(\Psi_m). \tag{4.66}$$

The equation (4.66) is equivalent to (4.60), which completes the proof. □

Corollary 4.11. *Let $\mathcal{X}$ be a complex Euclidean space and let $\Phi \in \mathrm{C}(\mathcal{X})$ be a mixed-unitary channel. There exists a positive integer $m \leq \operatorname{rank}(J(\Phi))^2$, a selection of unitary operators $U_1, \ldots, U_m \in \mathrm{U}(\mathcal{X})$, and a probability vector $(p_1, \ldots, p_m)$ such that*

$$\Phi(X) = \sum_{k=1}^{m} p_k U_k X U_k^* \tag{4.67}$$

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

4.1.2 Weyl-covariant channels

This section concerns the *Weyl-covariant channels*, which are a class of unital channels that relate (in multiple ways) to a collection of operators called the *discrete Weyl operators*.

Discrete Weyl operators

For every positive integer n, the set $\mathbb{Z}_n$ is defined as

$$\mathbb{Z}_n = \{0, \ldots, n-1\}. \tag{4.68}$$

This set forms a ring, with respect to addition and multiplication modulo n, and whenever elements of $\mathbb{Z}_n$ appear in arithmetic expressions in this book, the default assumption is that the operations are to be taken modulo n.

Now, assume that a positive integer n has been fixed, and let $\mathcal{X} = \mathbb{C}^{\mathbb{Z}_n}$. The *discrete Weyl operators* are a collection of unitary operators acting on $\mathcal{X}$, defined in the following way.[2] One first defines a scalar value

$$\zeta = \exp\left(\frac{2\pi i}{n}\right), \tag{4.69}$$

along with unitary operators

$$U = \sum_{c \in \mathbb{Z}_n} E_{c+1,c} \quad \text{and} \quad V = \sum_{c \in \mathbb{Z}_n} \zeta^c E_{c,c}. \tag{4.70}$$

For each pair $(a, b) \in \mathbb{Z}_n \times \mathbb{Z}_n$, the discrete Weyl operator $W_{a,b} \in \mathrm{U}(\mathcal{X})$ is then defined as

$$W_{a,b} = U^a V^b, \tag{4.71}$$

or equivalently as

$$W_{a,b} = \sum_{c \in \mathbb{Z}_n} \zeta^{bc} E_{a+c,c}. \tag{4.72}$$

[2] It is sometimes convenient to extend the definition of the discrete Weyl operators from complex Euclidean spaces of the form $\mathcal{X} = \mathbb{C}^{\mathbb{Z}_n}$ to arbitrary complex Euclidean spaces $\mathcal{X} = \mathbb{C}^{\Sigma}$, simply by placing Σ in correspondence with $\mathbb{Z}_n$, for $n = |\Sigma|$, in some fixed but otherwise arbitrary way.

Example 4.12. For $n = 2$, the discrete Weyl operators (in matrix form) are given by

$$\begin{aligned} W_{0,0} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad W_{0,1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \\ W_{1,0} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad W_{1,1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \end{aligned} \tag{4.73}$$

Equivalently,

$$W_{0,0} = \mathbb{1}, \quad W_{0,1} = \sigma_z, \quad W_{1,0} = \sigma_x, \quad W_{1,1} = -i\sigma_y, \tag{4.74}$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.75}$$

are the *Pauli operators*.

It holds that

$$UV = \sum_{c \in \mathbb{Z}_n} \zeta^c E_{c+1,c} \quad \text{and} \quad VU = \sum_{c \in \mathbb{Z}_n} \zeta^{c+1} E_{c+1,c}, \tag{4.76}$$

from which the commutation relation

$$VU = \zeta UV \tag{4.77}$$

follows. Identities that may be derived using this relation, together with straightforward calculations, include

$$\overline{W_{a,b}} = W_{a,-b}, \quad W_{a,b}^{\mathsf{T}} = \zeta^{-ab} W_{-a,b}, \quad \text{and} \quad W_{a,b}^* = \zeta^{ab} W_{-a,-b} \tag{4.78}$$

for all $a, b \in \mathbb{Z}_n$, and

$$W_{a,b} W_{c,d} = \zeta^{bc} W_{a+c,b+d} = \zeta^{bc-ad} W_{c,d} W_{a,b} \tag{4.79}$$

for all $a, b, c, d \in \mathbb{Z}_n$.

From the equation

$$\sum_{c \in \mathbb{Z}_n} \zeta^{ac} = \begin{cases} n & \text{if } a = 0 \\ 0 & \text{if } a \in \{1, \ldots, n-1\} \end{cases} \tag{4.80}$$

it follows that

$$\mathrm{Tr}(W_{a,b}) = \begin{cases} n & \text{if } (a,b) = (0,0) \\ 0 & \text{otherwise.} \end{cases} \tag{4.81}$$

Combining this observation with (4.79) yields

$$\langle W_{a,b}, W_{c,d} \rangle = \begin{cases} n & \text{if } (a,b) = (c,d) \\ 0 & \text{if } (a,b) \neq (c,d) \end{cases} \tag{4.82}$$

for all $a, b, c, d \in \mathbb{Z}_n$. The set

$$\left\{ \frac{1}{\sqrt{n}} W_{a,b} \,:\, (a,b) \in \mathbb{Z}_n \times \mathbb{Z}_n \right\} \tag{4.83}$$

therefore forms an orthonormal set. Because the cardinality of this set is equal to the dimension of $\mathrm{L}(\mathcal{X})$, it therefore forms an orthonormal basis for this space.

The *discrete Fourier transform operator* $F \in \mathrm{U}(\mathcal{X})$, defined as

$$F = \frac{1}{\sqrt{n}} \sum_{a,b \in \mathbb{Z}_n} \zeta^{ab} E_{a,b}\,, \tag{4.84}$$

has a special connection with the discrete Weyl operators. The fact that F is unitary may be verified by a direct calculation:

$$F^* F = \frac{1}{n} \sum_{a,b,c \in \mathbb{Z}_n} \zeta^{a(b-c)} E_{c,b} = \sum_{b \in \mathbb{Z}_n} E_{b,b} = \mathbb{1}. \tag{4.85}$$

It may also be verified that $FU = VF$ and $FV = U^* F$, from which it follows that

$$F W_{a,b} = \zeta^{-ab} W_{-b,a} F \tag{4.86}$$

for all $a, b \in \mathbb{Z}_n$.

Weyl-covariant maps and channels

A map $\Phi \in \mathrm{T}(\mathcal{X})$, for $\mathcal{X} = \mathbb{C}^{\mathbb{Z}_n}$ as above, is a *Weyl-covariant mapping* if it commutes with the action of conjugation by every discrete Weyl operator, as the following definition makes precise.

Definition 4.13. Let $\mathcal{X} = \mathbb{C}^{\mathbb{Z}_n}$ for n a positive integer. A map $\Phi \in \mathrm{T}(\mathcal{X})$ is a *Weyl-covariant map* if

$$\Phi\big(W_{a,b} X W_{a,b}^*\big) = W_{a,b} \Phi(X) W_{a,b}^* \tag{4.87}$$

for every $X \in \mathrm{L}(\mathcal{X})$ and $(a, b) \in \mathbb{Z}_n \times \mathbb{Z}_n$. If, in addition to being a Weyl-covariant map, Φ is a channel, then Φ is said to be a *Weyl-covariant channel*.

From this definition it follows that the set of Weyl-covariant maps of the form $\Phi \in \mathrm{T}(\mathcal{X})$ is a linear subspace of $\mathrm{T}(\mathcal{X})$; for any two Weyl-covariant maps $\Phi, \Psi \in \mathrm{T}(\mathcal{X})$ and scalars $\alpha, \beta \in \mathbb{C}$, the map $\alpha\Phi + \beta\Psi$ is also Weyl-covariant. It follows from this observation that the set of Weyl-covariant channels of the form $\Phi \in \mathrm{C}(\mathcal{X})$ is a convex subset of $\mathrm{C}(\mathcal{X})$.

The next theorem provides two alternative characterizations of Weyl-covariant maps. One characterization states that a map is Weyl-covariant if and only if each discrete Weyl operator is an eigenoperator[3] of that map. The other characterization states that a map is Weyl-covariant if and only if it is a linear combination of conjugations by discrete Weyl operators. The two characterizations are related by the discrete Fourier transform operator.

Theorem 4.14. *Let $\mathcal{X} = \mathbb{C}^{\mathbb{Z}_n}$ for a positive integer n, and let $\Phi \in \mathrm{T}(\mathcal{X})$ be a map. The following statements are equivalent:*

1. *Φ is a Weyl-covariant map.*
2. *There exists an operator $A \in \mathrm{L}(\mathcal{X})$ such that*

   $$\Phi(W_{a,b}) = A(a, b) W_{a,b} \tag{4.88}$$

   for all $(a, b) \in \mathbb{Z}_n \times \mathbb{Z}_n$.
3. *There exists an operator $B \in \mathrm{L}(\mathcal{X})$ such that*

   $$\Phi(X) = \sum_{a,b \in \mathbb{Z}_n} B(a, b) W_{a,b} X W_{a,b}^* \tag{4.89}$$

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

Under the assumption that these three statements hold, the operators A and B in statements 2 and 3 are related by the equation

$$A^{\mathsf{T}} = n F^* B F. \tag{4.90}$$

[3] The term *eigenoperator* should be interpreted in the natural way—as an operator analogue of an eigenvector for a linear map defined on a space of operators.

Proof. Assume Φ is a Weyl-covariant map and consider the operator

$$W_{a,b}^* \Phi(W_{a,b}), \tag{4.91}$$

for $(a,b) \in \mathbb{Z}_n \times \mathbb{Z}_n$ chosen arbitrarily. For every choice of $(c,d) \in \mathbb{Z}_n \times \mathbb{Z}_n$, it holds that

$$\begin{aligned} W_{a,b}^* \Phi(W_{a,b}) W_{c,d}^* &= W_{a,b}^* W_{c,d}^* W_{c,d} \Phi(W_{a,b}) W_{c,d}^* \\ &= W_{a,b}^* W_{c,d}^* \Phi(W_{c,d} W_{a,b} W_{c,d}^*) = W_{c,d}^* W_{a,b}^* \Phi(W_{a,b} W_{c,d} W_{c,d}^*) \\ &= W_{c,d}^* W_{a,b}^* \Phi(W_{a,b}), \end{aligned} \tag{4.92}$$

where the second equality has used the Weyl-covariance of Φ and the third equality has used the fact that

$$W_{c,d} W_{a,b} = \alpha\, W_{a,b} W_{c,d} \quad \text{and} \quad W_{a,b}^* W_{c,d}^* = \overline{\alpha}\, W_{c,d}^* W_{a,b}^* \tag{4.93}$$

for $\alpha = \zeta^{ad-bc}$. It follows that

$$[W_{a,b}^* \Phi(W_{a,b}), W_{c,d}^*] = 0 \tag{4.94}$$

for all $(c,d) \in \mathbb{Z}_n \times \mathbb{Z}_n$. As the set of all discrete Weyl operators forms a basis for $\mathrm{L}(\mathcal{X})$, it must therefore hold that $W_{a,b}^* \Phi(W_{a,b})$ commutes with all operators in $\mathrm{L}(\mathcal{X})$, and is therefore equal to a scalar multiple of the identity operator.

As this is true for every choice of $(a,b) \in \mathbb{Z}_n \times \mathbb{Z}_n$, it follows that one may choose an operator $A \in \mathrm{L}(\mathcal{X})$ so that

$$W_{a,b}^* \Phi(W_{a,b}) = A(a,b)\mathbb{1}, \tag{4.95}$$

and therefore

$$\Phi(W_{a,b}) = A(a,b) W_{a,b}, \tag{4.96}$$

for all $(a,b) \in \mathbb{Z}_n \times \mathbb{Z}_n$. Statement 1 therefore implies statement 2.

The reverse implication, that statement 2 implies statement 1, is implied by the commutation relation (4.79). In greater detail, suppose statement 2 holds, and let $(a,b) \in \mathbb{Z}_n \times \mathbb{Z}_n$. For each pair $(c,d) \in \mathbb{Z}_n \times \mathbb{Z}_n$, one has

$$\begin{aligned} \Phi(W_{a,b} W_{c,d} W_{a,b}^*) &= \zeta^{bc-ad} \Phi(W_{c,d}) = A(c,d) \zeta^{bc-ad} W_{c,d} \\ &= A(c,d) W_{a,b} W_{c,d} W_{a,b}^* = W_{a,b} \Phi(W_{c,d}) W_{a,b}^*, \end{aligned} \tag{4.97}$$

and therefore, again using the fact that the discrete Weyl operators form a basis for $\mathrm{L}(\mathcal{X})$, one has

$$\Phi(W_{a,b} X W_{a,b}^*) = W_{a,b} \Phi(X) W_{a,b}^* \tag{4.98}$$

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

Now assume statement 3 holds for some choice of $B \in \mathrm{L}(\mathcal{X})$. Using the commutation relation (4.79), it follows that

$$\Phi(W_{c,d}) = \sum_{a,b \in \mathbb{Z}_n} B(a,b) W_{a,b} W_{c,d} W_{a,b}^* = \sum_{a,b \in \mathbb{Z}_n} \zeta^{bc-ad} B(a,b) W_{c,d} \tag{4.99}$$

for every pair $(c,d) \in \mathbb{Z}_n \times \mathbb{Z}_n$. Choosing $A \in \mathrm{L}(\mathcal{X})$ so that

$$A(c,d) = \sum_{a,b \in \mathbb{Z}_n} \zeta^{bc-ad} B(a,b) \tag{4.100}$$

for all $(c,d) \in \mathbb{Z}_n \times \mathbb{Z}_n$, which is equivalent to $A = (nF^* B F)^{\mathsf{T}}$, one has that

$$\Phi(W_{c,d}) = A(c,d) W_{c,d} \tag{4.101}$$

for all $(c,d) \in \mathbb{Z}_n \times \mathbb{Z}_n$. Statement 3 therefore implies statement 2, with the operators A and B being related as claimed.

Finally, assume statement 2 holds for some choice of $A \in \mathrm{L}(\mathcal{X})$, and define $B = \frac{1}{n} F A^{\mathsf{T}} F^*$. By a similar calculation to the one used to establish the previous implication, one has

$$\begin{aligned} \Phi(W_{c,d}) &= A(c,d) W_{c,d} \\ &= \sum_{a,b \in \mathbb{Z}_n} \zeta^{bc-ad} B(a,b) W_{c,d} = \sum_{a,b \in \mathbb{Z}_n} B(a,b) W_{a,b} W_{c,d} W_{a,b}^* \end{aligned} \tag{4.102}$$

for every pair $(c,d) \in \mathbb{Z}_n \times \mathbb{Z}_n$, and therefore

$$\Phi(X) = \sum_{a,b \in \mathbb{Z}_n} B(a,b) W_{a,b} X W_{a,b}^* \tag{4.103}$$

for all $X \in \mathrm{L}(\mathcal{X})$ by linearity. Statement 2 therefore implies statement 3, where again A and B are related as claimed. □

Corollary 4.15. *Let $\mathcal{X} = \mathbb{C}^{\mathbb{Z}_n}$ for a positive integer n, and let $\Phi \in \mathrm{C}(\mathcal{X})$ be a Weyl-covariant channel. There exists a probability vector $p \in \mathcal{P}(\mathbb{Z}_n \times \mathbb{Z}_n)$ such that*

$$\Phi(X) = \sum_{a,b \in \mathbb{Z}_n} p(a,b) W_{a,b} X W_{a,b}^* \tag{4.104}$$

for all $X \in \mathrm{L}(\mathcal{X})$. In particular, it holds that Φ is a mixed-unitary channel.

Proof. By Theorem 4.14, there exists an operator $B \in \mathrm{L}(\mathcal{X})$ such that

$$\Phi(X) = \sum_{a,b \in \mathbb{Z}_n} B(a,b) W_{a,b} X W_{a,b}^* \tag{4.105}$$

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

$$J(\Phi) = \sum_{a,b \in \mathbb{Z}_n} B(a,b) \operatorname{vec}(W_{a,b}) \operatorname{vec}(W_{a,b})^*, \tag{4.106}$$

which is a positive semidefinite operator given the assumption that Φ is completely positive. This implies that $B(a,b)$ is nonnegative for every pair $(a,b) \in \mathbb{Z}_n \times \mathbb{Z}_n$, by virtue of the fact that the vectors

$$\{\operatorname{vec}(W_{a,b}) \,:\, a,b \in \mathbb{Z}_n\} \tag{4.107}$$

form an orthogonal set. It holds that

$$\operatorname{Tr}(\Phi(X)) = \sum_{a,b \in \mathbb{Z}_n} B(a,b) \operatorname{Tr}\bigl(W_{a,b} X W_{a,b}^*\bigr) = \sum_{a,b \in \mathbb{Z}_n} B(a,b) \operatorname{Tr}(X) \tag{4.108}$$

for every $X \in \mathrm{L}(\mathcal{X})$, and therefore

$$\sum_{a,b \in \mathbb{Z}_n} B(a,b) = 1 \tag{4.109}$$

by the assumption that Φ preserves trace. Defining $p(a,b) = B(a,b)$ for every pair $(a,b) \in \mathbb{Z}_n \times \mathbb{Z}_n$, one has that p is a probability vector, which completes the proof. □

Completely depolarizing and dephasing channels

The *completely depolarizing channel* $\Omega \in \mathrm{C}(\mathcal{X})$ and the *completely dephasing channel* $\Delta \in \mathrm{C}(\mathcal{X})$ are defined, for any choice of a complex Euclidean space $\mathcal{X} = \mathbb{C}^{\Sigma}$, as follows:

$$\Omega(X) = \frac{\operatorname{Tr}(X)}{\dim(\mathcal{X})} \mathbb{1}_{\mathcal{X}} \quad \text{and} \quad \Delta(X) = \sum_{a \in \Sigma} X(a,a) E_{a,a} \tag{4.110}$$

for all $X \in \mathrm{L}(\mathcal{X})$ (q.v. Section 2.2.3). In the case that the complex Euclidean space $\mathcal{X}$ takes the form $\mathcal{X} = \mathbb{C}^{\mathbb{Z}_n}$ for a positive integer n, these channels are both examples of Weyl-covariant channels.

That the completely depolarizing channel is a Weyl-covariant channel follows from the observation that

$$\Omega(W_{a,b}) = \begin{cases} W_{a,b} & \text{if } (a,b) = (0,0) \\ 0 & \text{if } (a,b) \neq (0,0), \end{cases} \tag{4.111}$$

or equivalently $\Omega(W_{a,b}) = E_{0,0}(a,b)W_{a,b}$, for every $(a,b) \in \mathbb{Z}_n \times \mathbb{Z}_n$. Thus, by Theorem 4.14, together with the observation that

$$\frac{1}{n} F E_{0,0} F^* = \frac{1}{n^2} \sum_{a,b \in \mathbb{Z}_n} E_{a,b}, \tag{4.112}$$

one has that

$$\Omega(X) = \frac{1}{n^2} \sum_{a,b \in \mathbb{Z}_n} W_{a,b} X W_{a,b}^* \tag{4.113}$$

for all $X \in \mathrm{L}(\mathcal{X})$. An alternative way to establish the validity of the equation (4.113) is to observe that the Choi operator of the map defined by the right-hand side of that equation is in agreement with the Choi operator of Ω:

$$\frac{1}{n^2} \sum_{a,b \in \mathbb{Z}_n} \operatorname{vec}(W_{a,b}) \operatorname{vec}(W_{a,b})^* = \frac{1}{n} \mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{X}} = J(\Omega). \tag{4.114}$$

As mentioned in the footnote on page 231, one may translate the notion of a discrete Weyl operator from a space of the form $\mathbb{C}^{\mathbb{Z}_n}$ to an arbitrary complex Euclidean space $\mathbb{C}^{\Sigma}$ through any fixed correspondence between the elements of Σ and $\mathbb{Z}_n$ (assuming $n = |\Sigma|$). It follows that the completely depolarizing channel $\Omega \in \mathrm{C}(\mathcal{X})$ is a mixed-unitary channel for any choice of a complex Euclidean space $\mathcal{X} = \mathbb{C}^{\Sigma}$, as it is equal to the Weyl-covariant channel defined above with respect to any chosen correspondence between Σ and $\mathbb{Z}_n$.

The completely dephasing channel is a Weyl-covariant channel, as is evident from the observation that

$$\Delta(W_{a,b}) = \begin{cases} W_{a,b} & \text{if } a = 0 \\ 0 & \text{if } a \neq 0, \end{cases} \tag{4.115}$$

or equivalently $\Omega(W_{a,b}) = A(a,b)W_{a,b}$ for

$$A = \sum_{c \in \mathbb{Z}_n} E_{0,c}, \tag{4.116}$$

for all $(a, b) \in \mathbb{Z}_n \times \mathbb{Z}_n$. By Theorem 4.14, together with the observation that $FA^{\mathsf{T}}F^* = A$, it follows that

$$\Delta(X) = \frac{1}{n} \sum_{c \in \mathbb{Z}_n} W_{0,c} X W_{0,c}^* \tag{4.117}$$

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

4.1.3 Schur channels

Schur channels, which are defined as follows, represent another interesting sub-class of unital channels.

Definition 4.16. Let $\mathcal{X} = \mathbb{C}^{\Sigma}$ be the complex Euclidean space indexed by a given alphabet Σ. A map $\Phi \in \mathrm{T}(\mathcal{X})$ is said to be a *Schur map* if there exists an operator $A \in \mathrm{L}(\mathcal{X})$ satisfying

$$\Phi(X) = A \odot X, \tag{4.118}$$

where $A \odot X$ denotes the entry-wise product of A and X:

$$(A \odot X)(a, b) = A(a, b) X(a, b) \tag{4.119}$$

for all $a, b \in \Sigma$. If, in addition, the map Φ is a channel, then it is said to be a *Schur channel*.

The following proposition provides a simple condition under which a given Schur map is completely positive (or, equivalently, positive).

Proposition 4.17. *Let Σ be an alphabet, let $\mathcal{X} = \mathbb{C}^{\Sigma}$, let $A \in \mathrm{L}(\mathcal{X})$ be an operator, and let $\Phi \in \mathrm{T}(\mathcal{X})$ be the Schur map defined as*

$$\Phi(X) = A \odot X \tag{4.120}$$

for all $X \in \mathrm{L}(\mathcal{X})$. The following statements are equivalent:

1. *A is positive semidefinite.*
2. *Φ is positive.*
3. *Φ is completely positive.*

Proof. Suppose A is positive semidefinite. It holds that

$$J(\Phi) = \sum_{a,b\in\Sigma} \Phi(E_{a,b}) \otimes E_{a,b} = \sum_{a,b\in\Sigma} A(a,b) E_{a,b} \otimes E_{a,b} = VAV^* \tag{4.121}$$

for $V \in \mathrm{U}(\mathcal{X}, \mathcal{X} \otimes \mathcal{X})$ being the isometry defined as

$$V = \sum_{a\in\Sigma} (e_a \otimes e_a) e_a^*. \tag{4.122}$$

This implies that $J(\Phi)$ is positive semidefinite, so Φ is completely positive by Theorem 2.22. It has been proved that statement 1 implies statement 3.

Statement 3 trivially implies statement 2 as every completely positive map is positive.

Finally, assume that Φ is positive. The operator $X \in \mathrm{L}(\mathcal{X})$ defined as

$$X(a,b) = 1 \tag{4.123}$$

for all $a, b \in \Sigma$ is positive semidefinite. By the positivity of Φ, it therefore holds that $\Phi(X) = A$ is positive semidefinite. Statement 2 therefore implies statement 1, which completes the proof. □

In a similar spirit to the previous proposition, the following proposition provides a simple condition under which a given Schur map preserves trace (or, equivalently, is unital).

Proposition 4.18. *Let Σ be an alphabet, let $\mathcal{X} = \mathbb{C}^\Sigma$, let $A \in \mathrm{L}(\mathcal{X})$ be an operator, and let $\Phi \in \mathrm{T}(\mathcal{X})$ be the Schur mapping defined as*

$$\Phi(X) = A \odot X \tag{4.124}$$

for all $X \in \mathrm{L}(\mathcal{X})$. The following statements are equivalent:

1. *$A(a,a) = 1$ for every $a \in \Sigma$*
2. *Φ is trace-preserving.*
3. *Φ is unital.*

Proof. Suppose $A(a,a) = 1$ for every $a \in \Sigma$. It follows that Φ is unital, as

$$\Phi(\mathbb{1}) = A \odot \mathbb{1} = \sum_{a\in\Sigma} A(a,a) E_{a,a} = \sum_{a\in\Sigma} E_{a,a} = \mathbb{1}. \tag{4.125}$$

It also follows that Φ is trace-preserving, as

$$\begin{aligned} \operatorname{Tr}(\Phi(X)) &= \sum_{a\in\Sigma} (A \odot X)(a,a) \\ &= \sum_{a\in\Sigma} A(a,a)X(a,a) = \sum_{a\in\Sigma} X(a,a) = \operatorname{Tr}(X) \end{aligned} \tag{4.126}$$

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

The assumption that Φ is trace-preserving implies that

$$A(a,a) = \operatorname{Tr}(A(a,a)E_{a,a}) = \operatorname{Tr}(\Phi(E_{a,a})) = \operatorname{Tr}(E_{a,a}) = 1 \tag{4.127}$$

for all $a \in \Sigma$. Statements 1 and 2 are therefore equivalent.

Finally, the assumption that Φ is unital implies

$$\sum_{a\in\Sigma} A(a,a)E_{a,a} = \Phi(\mathbb{1}) = \mathbb{1} = \sum_{a\in\Sigma} E_{a,a}, \tag{4.128}$$

and therefore $A(a,a) = 1$ for every $a \in \Sigma$. Statements 1 and 3 are therefore equivalent. □

Completely positive Schur maps may alternatively be characterized as the class of maps having Kraus representations consisting only of equal pairs of diagonal operators, as the following theorem states.

Theorem 4.19. *Let Σ be an alphabet, let $\mathcal{X} = \mathbb{C}^{\Sigma}$ be the complex Euclidean space indexed by Σ, and let $\Phi \in \mathrm{CP}(\mathcal{X})$ be a completely positive map. The following statements are equivalent:*

1. *Φ is a Schur map.*
2. *There exists a Kraus representation of Φ having the form*

$$\Phi(X) = \sum_{a\in\Gamma} A_a X A_a^*, \tag{4.129}$$

   for some alphabet Γ, such that $A_a \in \mathrm{L}(\mathcal{X})$ is a diagonal operator for each $a \in \Gamma$.
3. *For every Kraus representation of Φ having the form* (4.129)*, A_a is a diagonal operator for each $a \in \Gamma$.*

Proof. Suppose first that Φ is a Schur map, given by

$$\Phi(X) = P \odot X \tag{4.130}$$

for all $X \in \mathrm{L}(\mathcal{X})$, for some operator $P \in \mathrm{L}(\mathcal{X})$. By the assumption that Φ is completely positive, Proposition 4.17 implies that P is positive semidefinite. As was computed in the proof of that proposition, the Choi representation of Φ is given by

$$J(\Phi) = VPV^* \tag{4.131}$$

for

$$V = \sum_{b \in \Sigma} (e_b \otimes e_b) e_b^*. \tag{4.132}$$

Consider an arbitrary Kraus representation of Φ having the form (4.129), for some alphabet Γ and a collection $\{A_a : a \in \Gamma\} \subset \mathrm{L}(\mathcal{X})$ of operators. As the Choi representation of the map defined by the right-hand side of that equation must agree with (4.131), it holds that

$$\sum_{a \in \Gamma} \mathrm{vec}(A_a)\, \mathrm{vec}(A_a)^* = VPV^*, \tag{4.133}$$

and therefore

$$\mathrm{vec}(A_a) \in \mathrm{im}(V) = \mathrm{span}\{e_b \otimes e_b \,:\, b \in \Sigma\} \tag{4.134}$$

for every $a \in \Gamma$. This is equivalent to the condition that A_a is diagonal for every $a \in \Gamma$, and so it has been proved that statement 1 implies statement 3.

Statement 3 trivially implies statement 2, so it remains to prove that statement 2 implies statement 1. For a Kraus representation of Φ having the form (4.129), where Γ is an alphabet and $\{A_a : a \in \Gamma\}$ is a collection of diagonal operators, let $\{v_a : a \in \Gamma\} \subset \mathcal{X}$ be the collection of vectors satisfying $A_a = \mathrm{Diag}(v_a)$ for each $a \in \Gamma$, and define

$$P = \sum_{a \in \Gamma} v_a v_a^*. \tag{4.135}$$

A calculation reveals that

$$P \odot X = \sum_{a \in \Gamma} \sum_{b,c \in \Sigma} X(b,c)\, v_a(b) \overline{v_a(c)}\, E_{b,c} = \sum_{a \in \Gamma} A_a X A_a^* \tag{4.136}$$

for every $X \in \mathrm{L}(\mathcal{X})$. It has therefore been proved that Φ is a Schur mapping, so statement 2 implies statement 1 as required. □

4.2 General properties of unital channels

This section proves a few basic facts holding for unital channels in general. In particular, the extreme points of the set of all unital channels defined with respect to a given space are characterized, and properties relating to fixed-points and norms of unital channels are established.

4.2.1 Extreme points of the set of unital channels

Theorem 2.31 provides a criterion through which one may determine if a given channel $\Phi \in \mathrm{C}(\mathcal{X})$ is an extreme point of the set of all channels $\mathrm{C}(\mathcal{X})$, based on any linearly independent set of Kraus operators of Φ. It will be demonstrated by Theorem 4.21 below that a similar criterion holds when the set $\mathrm{C}(\mathcal{X})$ is replaced by the set of all unital channels

$$\big\{\Phi \in \mathrm{C}(\mathcal{X}) \,:\, \Phi(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{X}}\big\}. \tag{4.137}$$

Indeed, the analogous theorem for unital channels will follow directly from Theorem 2.31, together with an embedding of the set (4.137) within the set of all channels of the form $\mathrm{C}(\mathcal{X} \oplus \mathcal{X})$.

Assume that a complex Euclidean space $\mathcal{X}$ has been fixed, and define an operator

$$V \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X}, (\mathcal{X} \oplus \mathcal{X}) \otimes (\mathcal{X} \oplus \mathcal{X})) \tag{4.138}$$

by the equation

$$V \operatorname{vec}(X) = \operatorname{vec}\begin{pmatrix} X & 0 \\ 0 & X^{\mathsf{T}} \end{pmatrix} \tag{4.139}$$

holding for all operators $X \in \mathrm{L}(\mathcal{X})$. It may be verified that $V^*V = 2\mathbb{1}_{\mathcal{X}}$. For every map $\Phi \in \mathrm{T}(\mathcal{X})$, define $\phi(\Phi) \in \mathrm{T}(\mathcal{X} \oplus \mathcal{X})$ to be the unique map for which the equation

$$J(\phi(\Phi)) = VJ(\Phi)V^* \tag{4.140}$$

holds, and observe that the mapping $\phi : \mathrm{T}(\mathcal{X}) \to \mathrm{T}(\mathcal{X} \oplus \mathcal{X})$ defined in this way is injective and linear. If $\Phi \in \mathrm{T}(\mathcal{X})$ is defined by a Kraus representation

$$\Phi(X) = \sum_{a \in \Sigma} A_a X B_a^*, \tag{4.141}$$

then it holds that

$$\phi(\Phi)\begin{pmatrix} X_{0,0} & X_{0,1} \\ X_{1,0} & X_{1,1} \end{pmatrix} = \sum_{a \in \Sigma} \begin{pmatrix} A_a & 0 \\ 0 & A_a^{\mathsf{T}} \end{pmatrix} \begin{pmatrix} X_{0,0} & X_{0,1} \\ X_{1,0} & X_{1,1} \end{pmatrix} \begin{pmatrix} B_a & 0 \\ 0 & B_a^{\mathsf{T}} \end{pmatrix}^* \tag{4.142}$$

is a Kraus representation of $\phi(\Phi)$. The following observations concerning the mapping $\phi : \mathrm{T}(\mathcal{X}) \to \mathrm{T}(\mathcal{X} \oplus \mathcal{X})$ may be verified:

1. A map $\Phi \in \mathrm{T}(\mathcal{X})$ is completely positive if and only if $\phi(\Phi) \in \mathrm{T}(\mathcal{X} \oplus \mathcal{X})$ is completely positive.
2. A map $\Phi \in \mathrm{T}(\mathcal{X})$ is both trace-preserving and unital if and only if $\phi(\Phi) \in \mathrm{T}(\mathcal{X} \oplus \mathcal{X})$ is trace-preserving.

In particular, $\Phi \in \mathrm{C}(\mathcal{X})$ is a unital channel if and only if $\phi(\Phi) \in \mathrm{C}(\mathcal{X} \oplus \mathcal{X})$ is a channel. (In this case, $\phi(\Phi)$ will also happen to be unital.)

Lemma 4.20. *Let $\mathcal{X}$ be a complex Euclidean space, let $\Phi \in \mathrm{C}(\mathcal{X})$ be a unital channel, and let $\phi(\Phi) \in \mathrm{C}(\mathcal{X} \oplus \mathcal{X})$ be the channel defined from Φ by the equation* (4.140)*. It holds that Φ is an extreme point in the set of all unital channels in $\mathrm{C}(\mathcal{X})$ if and only if $\phi(\Phi)$ is an extreme point of the set of channels $\mathrm{C}(\mathcal{X} \oplus \mathcal{X})$.*

Proof. Suppose first that Φ is not an extreme point in the set of all unital channels in $\mathrm{C}(\mathcal{X})$, so that

$$\Phi = \lambda \Psi_0 + (1 - \lambda)\Psi_1 \tag{4.143}$$

for distinct unital channels $\Psi_0, \Psi_1 \in \mathrm{C}(\mathcal{X})$ and a scalar $\lambda \in (0, 1)$. As the mapping ϕ is linear and injective, it therefore holds that

$$\phi(\Phi) = \lambda \phi(\Psi_0) + (1 - \lambda)\phi(\Psi_1), \tag{4.144}$$

which is a proper convex combination of distinct channels. This implies that $\phi(\Phi)$ is not an extreme point of the set of channels $\mathrm{C}(\mathcal{X} \oplus \mathcal{X})$.

Suppose, on the other hand, that $\phi(\Phi)$ is not an extreme point of the set of channels $\mathrm{C}(\mathcal{X} \oplus \mathcal{X})$, so that

$$\phi(\Phi) = \lambda \Xi_0 + (1 - \lambda)\Xi_1 \tag{4.145}$$

for distinct channels $\Xi_0, \Xi_1 \in \mathrm{C}(\mathcal{X} \oplus \mathcal{X})$ and a scalar $\lambda \in (0, 1)$. Taking the Choi representations of both sides of this equation yields

$$VJ(\Phi)V^* = \lambda J(\Xi_0) + (1 - \lambda)J(\Xi_1). \tag{4.146}$$

It therefore follows from Lemma 2.30 that

$$J(\Xi_0) = VQ_0V^* \quad \text{and} \quad J(\Xi_1) = VQ_1V^* \tag{4.147}$$

for some choice of positive semidefinite operators $Q_0, Q_1 \in \mathrm{Pos}(\mathcal{X})$. Letting $\Psi_0, \Psi_1 \in \mathrm{T}(\mathcal{X})$ be the maps defined by $J(\Psi_0) = Q_0$ and $J(\Psi_1) = Q_1$, one has $\Xi_0 = \phi(\Psi_0)$ and $\Xi_1 = \phi(\Psi_1)$. As $\phi(\Psi_0) = \Xi_0$ and $\phi(\Psi_1) = \Xi_1$ are distinct channels, it follows that Ψ_0 and Ψ_1 are distinct unital channels. It holds that

$$\phi(\Phi) = \lambda\phi(\Psi_0) + (1-\lambda)\phi(\Psi_1) \tag{4.148}$$

and therefore

$$\Phi = \lambda\Psi_0 + (1-\lambda)\Psi_1, \tag{4.149}$$

which implies that Φ is not an extreme point in the set of all unital channels in $\mathrm{C}(\mathcal{X})$. □

Theorem 4.21. *Let $\mathcal{X}$ be a complex Euclidean space, let $\Phi \in \mathrm{C}(\mathcal{X})$ be a unital channel, and let $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X})$ be a linearly independent set of operators satisfying*

$$\Phi(X) = \sum_{a\in\Sigma} A_a X A_a^* \tag{4.150}$$

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

$$\left\{ \begin{pmatrix} A_b^* A_a & 0 \\ 0 & A_a A_b^* \end{pmatrix} : (a,b) \in \Sigma \times \Sigma \right\} \tag{4.151}$$

of operators is linearly independent.

Proof. By Lemma 4.20, the channel Φ is an extreme point of the set of unital channels in $\mathrm{C}(\mathcal{X})$ if and only if the channel $\phi(\Phi)$ is an extreme point in the set $\mathrm{C}(\mathcal{X} \oplus \mathcal{X})$, for $\phi : \mathrm{T}(\mathcal{X}) \to \mathrm{T}(\mathcal{X} \oplus \mathcal{X})$ being the mapping defined by the equation (4.140). By Theorem 2.31, it follows that $\phi(\Phi)$ is an extreme point of the set of channels $\mathrm{C}(\mathcal{X} \oplus \mathcal{X})$ if and only if

$$\left\{ \begin{pmatrix} A_b^* A_a & 0 \\ 0 & \overline{A_b} A_a^{\mathsf{T}} \end{pmatrix} : (a,b) \in \Sigma \times \Sigma \right\} \tag{4.152}$$

is a linearly independent collection of operators. Taking the transpose of the lower-right-hand block, which does not change whether or not the set is linearly independent, it follows that $\phi(\Phi)$ is an extreme point of the set $\mathrm{C}(\mathcal{X} \oplus \mathcal{X})$ if and only if the set (4.151) is linearly independent. □

Unital qubit channels are mixed-unitary

There exist non-mixed-unitary unital channels, as shown in Example 4.3. The existence of such channels, however, requires that the underlying space has dimension at least 3; when Theorem 4.21 is combined with the following lemma, one concludes that every unital qubit channel is mixed-unitary.

Lemma 4.22. *Let $\mathcal{X}$ be a complex Euclidean space and let $A_0, A_1 \in \mathrm{L}(\mathcal{X})$ be operators such that*

$$A_0^* A_0 + A_1^* A_1 = \mathbb{1}_{\mathcal{X}} = A_0 A_0^* + A_1 A_1^*. \tag{4.153}$$

There exist unitary operators $U, V \in \mathrm{U}(\mathcal{X})$ such that VA_0U^* and VA_1U^* are diagonal operators.

Proof. It suffices to prove that there exists a unitary operator $W \in \mathrm{U}(\mathcal{X})$ such that the operators WA_0 and WA_1 are both normal and satisfy

$$[WA_0, WA_1] = 0, \tag{4.154}$$

for then it follows by Theorem 1.5 that one may choose U so that UWA_0U^* and UWA_1U^* are diagonal, then take $V = UW$.

Let $U_0, U_1 \in \mathrm{U}(\mathcal{X})$ and $P_0, P_1 \in \mathrm{Pos}(\mathcal{X})$ be operators providing the polar decompositions

$$A_0 = U_0 P_0 \quad \text{and} \quad A_1 = U_1 P_1, \tag{4.155}$$

and let $W = U_0^*$. It holds that $WA_0 = P_0$, which is positive semidefinite and therefore normal. To verify that WA_1 is normal, observe that the assumption (4.153) implies

$$U_1 P_1^2 U_1^* = \mathbb{1} - U_0 P_0^2 U_0^* \quad \text{and} \quad P_1^2 = \mathbb{1} - P_0^2, \tag{4.156}$$

and therefore

$$\begin{aligned} (WA_1)(WA_1)^* &= U_0^* U_1 P_1^2 U_1^* U_0 = U_0^* (\mathbb{1} - U_0 P_0^2 U_0^*) U_0 \\ &= \mathbb{1} - P_0^2 = P_1^2 = P_1 U_1^* U_0 U_0^* U_1 P_1 = (WA_1)^* (WA_1). \end{aligned} \tag{4.157}$$

It remains to prove that the operators WA_0 and WA_1 commute. It follows from the equation $P_1^2 = \mathbb{1} - P_0^2$ that P_0^2 and P_1^2 commute. As P_0^2 and P_1^2 are commuting positive semidefinite operators, it therefore holds that P_0 and P_1 commute. Substituting $P_1^2 = \mathbb{1} - P_0^2$ into the equation

$$U_1 P_1^2 U_1^* = \mathbb{1} - U_0 P_0^2 U_0^*, \tag{4.158}$$

one finds that

$$U_0 P_0^2 U_0^* = U_1 P_0^2 U_1^*, \tag{4.159}$$

and therefore, by taking the square root of both sides of this equation,

$$U_0 P_0 U_0^* = U_1 P_0 U_1^*. \tag{4.160}$$

This implies that

$$P_0 U_0^* U_1 = U_0^* U_1 P_0, \tag{4.161}$$

and therefore P_0 and $U_0^* U_1$ commute. It follows that

$$\begin{aligned} (WA_0)(WA_1) &= P_0 U_0^* U_1 P_1 = U_0^* U_1 P_0 P_1 \\ &= U_0^* U_1 P_1 P_0 = (WA_1)(WA_0), \end{aligned} \tag{4.162}$$

and so WA_0 and WA_1 commute as required. □

Theorem 4.23. *Let $\mathcal{X}$ be a complex Euclidean space with $\dim(\mathcal{X}) = 2$. Every unital channel $\Phi \in \mathrm{C}(\mathcal{X})$ is a mixed-unitary channel.*

Proof. The set

$$\{\Phi \in \mathrm{C}(\mathcal{X}) : \Phi(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{X}}\} \tag{4.163}$$

of unital channels, defined with respect to the space $\mathcal{X}$, is both compact and convex—both of these properties are consequences of the fact that this set is equal to the intersection of the compact and convex set $\mathrm{C}(\mathcal{X})$ with the (closed) affine subspace of all maps $\Phi \in \mathrm{T}(\mathcal{X})$ satisfying $\Phi(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{X}}$. As this set is compact and convex, Theorem 1.10 implies that it is equal to the convex hull of its extreme points. To complete the proof, it therefore suffices to establish that every unital channel $\Phi \in \mathrm{C}(\mathcal{X})$ that is not a unitary channel is not an extreme point of the set (4.163).

Toward this goal, let $\Phi \in \mathrm{C}(\mathcal{X})$ be an arbitrary unital channel, and let $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X})$ be a linearly independent collection of operators satisfying

$$\Phi(X) = \sum_{a \in \Sigma} A_a X A_a^* \tag{4.164}$$

for all $X \in \mathrm{L}(\mathcal{X})$. One has that Φ is a unitary channel if and only if $|\Sigma| = 1$, so it suffices to prove that Φ is not an extreme point of the set (4.163) whenever $|\Sigma| \geq 2$.

By Theorem 4.21, the channel Φ is an extreme point of the set (4.163) if and only if

$$\left\{ \begin{pmatrix} A_b^* A_a & 0 \\ 0 & A_a A_b^* \end{pmatrix} \,:\, (a,b) \in \Sigma \times \Sigma \right\} \subset \mathrm{L}(\mathcal{X} \oplus \mathcal{X}) \tag{4.165}$$

is a linearly independent collection of operators. There are two cases that must be considered: the first case is that $|\Sigma| \geq 3$ and the second case is that $|\Sigma| = 2$.

For the first case, one has that the collection (4.165) includes at least 9 operators drawn from the 8-dimensional subspace

$$\left\{ \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \,:\, X, Y \in \mathrm{L}(\mathcal{X}) \right\}. \tag{4.166}$$

Thus, if $|\Sigma| \geq 3$, then the collection (4.165) cannot be linearly independent, and therefore Φ is not an extreme point of the set (4.163).

It remains to consider the case $|\Sigma| = 2$. There is no loss of generality in assuming $\Sigma = \{0,1\}$ and $\mathcal{X} = \mathbb{C}^{\Sigma}$. By Lemma 4.22, there must exist unitary operators $U, V \in \mathrm{U}(\mathcal{X})$ such that VA_0U^* and VA_1U^* are diagonal operators:

$$\begin{aligned} VA_0U^* &= \alpha_0 E_{0,0} + \beta_0 E_{1,1}\,, \\ VA_1U^* &= \alpha_1 E_{0,0} + \beta_1 E_{1,1}. \end{aligned} \tag{4.167}$$

The following equations therefore hold for every choice of $a, b \in \Sigma$:

$$\begin{aligned} A_b^* A_a &= \alpha_a \overline{\alpha_b} U^* E_{0,0} U + \beta_a \overline{\beta_b} U^* E_{1,1} U, \\ A_a A_b^* &= \alpha_a \overline{\alpha_b} V^* E_{0,0} V + \beta_a \overline{\beta_b} V^* E_{1,1} V. \end{aligned} \tag{4.168}$$

The set (4.165) is therefore contained in the subspace spanned by the set of operators

$$\left\{ \begin{pmatrix} U^* E_{0,0} U & 0 \\ 0 & V^* E_{0,0} V \end{pmatrix}, \begin{pmatrix} U^* E_{1,1} U & 0 \\ 0 & V^* E_{1,1} V \end{pmatrix} \right\}. \tag{4.169}$$

The collection (4.165) contains 4 operators drawn from a two-dimensional space, and therefore cannot be linearly independent. This implies that the channel Φ is not an extreme point of the set (4.163), which completes the proof. □

4.2.2 Fixed-points, spectra, and norms of unital channels

Every channel of the form $\Phi \in \mathrm{C}(\mathcal{X})$ must have at least one density operator fixed point, meaning a density operator $\rho \in \mathrm{D}(\mathcal{X})$ satisfying

$$\Phi(\rho) = \rho. \tag{4.170}$$

One may see this fact as a consequence of the Brouwer fixed-point theorem, which states that every continuous function mapping a compact, convex set in a Euclidean space to itself must have a fixed point. The full power of the Brouwer fixed-point theorem is, however, really not needed in this case; the fact that channels are linear maps allows for a simpler proof. The following theorem establishes this fact in slightly greater generality, for any positive and trace-preserving map $\Phi \in \mathrm{T}(\mathcal{X})$.

Theorem 4.24. *Let $\mathcal{X}$ be a complex Euclidean space and let $\Phi \in \mathrm{T}(\mathcal{X})$ be a positive and trace-preserving map. There exists a density operator $\rho \in \mathrm{D}(\mathcal{X})$ such that $\Phi(\rho) = \rho$.*

Proof. For each nonnegative integer $n \in \mathbb{N}$, define a map $\Psi_n \in \mathrm{T}(\mathcal{X})$ as

$$\Psi_n(X) = \frac{1}{2^n} \sum_{k=0}^{2^n - 1} \Phi^k(X) \tag{4.171}$$

for each $X \in \mathrm{L}(\mathcal{X})$, and define a set

$$\mathcal{C}_n = \{\Psi_n(\rho) \,:\, \rho \in \mathrm{D}(\mathcal{X})\}. \tag{4.172}$$

As Φ is linear, positive, and trace-preserving, the same is true of Ψ_n, and so it follows that $\mathcal{C}_n$ is a compact and convex subset of $\mathrm{D}(\mathcal{X})$ for each $n \in \mathbb{N}$. By the convexity of the set $\mathcal{C}_n$, it holds that

$$\Psi_{n+1}(\rho) = \frac{1}{2}\Psi_n(\rho) + \frac{1}{2}\Psi_n\Big(\Phi^{2^n}(\rho)\Big) \in \mathcal{C}_n \tag{4.173}$$

for every $\rho \in \mathrm{D}(\mathcal{X})$, and therefore $\mathcal{C}_{n+1} \subseteq \mathcal{C}_n$, for every $n \in \mathbb{N}$. As each $\mathcal{C}_n$ is compact and $\mathcal{C}_{n+1} \subseteq \mathcal{C}_n$ for all $n \in \mathbb{N}$, it follows that there must exist an element

$$\rho \in \bigcap_{n \in \mathbb{N}} \mathcal{C}_n \tag{4.174}$$

contained in the intersection of all of these sets.

Now, fix any choice of ρ satisfying (4.174). For an arbitrary choice of $n \in \mathbb{N}$, it holds that $\rho = \Psi_n(\sigma)$ for some choice of $\sigma \in \mathrm{D}(\mathcal{X})$, and therefore

$$\Phi(\rho) - \rho = \Phi(\Psi_n(\sigma)) - \Psi_n(\sigma) = \frac{\Phi^{2^n}(\sigma) - \sigma}{2^n}. \tag{4.175}$$

As the trace distance between two density operators cannot exceed 2, it follows that

$$\big\| \Phi(\rho) - \rho \big\|_1 \leq \frac{1}{2^{n-1}}. \tag{4.176}$$

This bound holds for every $n \in \mathbb{N}$, which implies $\|\Phi(\rho) - \rho\|_1 = 0$, and therefore $\Phi(\rho) = \rho$ as required. □

There is, of course, no difficulty in proving the existence of a density operator fixed point of a unital channel: if $\Phi \in \mathrm{C}(\mathcal{X})$ is a unital channel, then $\omega = \mathbb{1}_{\mathcal{X}} / \dim(\mathcal{X})$ is a density operator fixed point of Φ. What is more interesting is the fact that the collection of all operators $X \in \mathrm{L}(\mathcal{X})$ satisfying $\Phi(X) = X$ forms a unital sub-algebra of $\mathrm{L}(\mathcal{X})$, as the following theorem implies.

Theorem 4.25. *Let $\mathcal{X}$ be a complex Euclidean space and let $\Phi \in \mathrm{C}(\mathcal{X})$ be a unital channel. Suppose further that Σ is an alphabet and $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X})$ is a collection of operators yielding a Kraus representation of Φ:*

$$\Phi(X) = \sum_{a \in \Sigma} A_a X A_a^* \tag{4.177}$$

for all $X \in \mathrm{L}(\mathcal{X})$. For every $X \in \mathrm{L}(\mathcal{X})$ it holds that $\Phi(X) = X$ if and only if $[X, A_a] = 0$ for every $a \in \Sigma$.

Proof. If $X \in \mathrm{L}(\mathcal{X})$ is an operator for which $[X, A_a] = 0$ for every $a \in \Sigma$, then

$$\Phi(X) = \sum_{a \in \Sigma} A_a X A_a^* = \sum_{a \in \Sigma} X A_a A_a^* = X \Phi(\mathbb{1}) = X, \tag{4.178}$$

where the last equality follows from the assumption that Φ is unital.

Now suppose that $X \in \mathrm{L}(\mathcal{X})$ is an operator for which $\Phi(X) = X$, and consider the positive semidefinite operator

$$\sum_{a \in \Sigma} [X, A_a] \, [X, A_a]^*. \tag{4.179}$$

Expanding this operator and using the assumptions that Φ is unital and $\Phi(X) = X$ (and therefore $\Phi(X^*) = X^*$), one has

$$\begin{aligned}
&\sum_{a\in\Sigma} [X, A_a]\,[X, A_a]^* \\
&\quad = \sum_{a\in\Sigma} \Big(\big(XA_a - A_aX\big)\big(A_a^*X^* - X^*A_a^*\big) \Big) \\
&\quad = \sum_{a\in\Sigma} \big(XA_aA_a^*X^* - A_aXA_a^*X^* - XA_aX^*A_a^* + A_aXX^*A_a^*\big) \\
&\quad = XX^* - \Phi(X)X^* - X\Phi(X^*) + \Phi(XX^*) \\
&\quad = \Phi(XX^*) - XX^*.
\end{aligned} \tag{4.180}$$

As Φ is a channel, and is therefore trace-preserving, it holds that the trace of the operator represented by the previous equation is zero. The only traceless positive semidefinite operator is the zero operator, and therefore

$$\sum_{a\in\Sigma} [X, A_a]\,[X, A_a]^* = 0. \tag{4.181}$$

This implies that each of the terms $[X, A_a]\,[X, A_a]^*$ is zero, and therefore each operator $[X, A_a]$ is zero. □

For any channel of the form $\Phi \in \mathrm{C}(\mathcal{X})$, for $\mathcal{X}$ being a complex Euclidean space, one has that the natural representation of Φ is a square operator of the form $K(\Phi) \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$. The following proposition establishes that the spectral radius of $K(\Phi)$ is necessarily equal to 1.

Proposition 4.26. *Let $\mathcal{X}$ be a complex Euclidean space and let $\Phi \in \mathrm{C}(\mathcal{X})$ be a channel. The spectral radius of $K(\Phi)$ is equal to 1.*

Proof. By Theorem 4.24, there must exist a density operator $\rho \in \mathrm{D}(\mathcal{X})$ such that $\Phi(\rho) = \rho$, which implies that $K(\Phi)$ has an eigenvalue equal to 1.

It remains to prove that every eigenvalue of $K(\Phi)$ is at most 1 in absolute value, which is equivalent to the statement that $|\lambda| \leq 1$ for every choice of a nonzero operator $X \in \mathrm{L}(\mathcal{X})$ and a complex number $\lambda \in \mathbb{C}$ satisfying

$$\Phi(X) = \lambda X. \tag{4.182}$$

Suppose that $X \in \mathrm{L}(\mathcal{X})$ and $\lambda \in \mathbb{C}$ satisfy (4.182). By Corollary 3.43, it holds that $\|\Phi\|_1 = 1$, and therefore

$$1 \geq \frac{\|\Phi(X)\|_1}{\|X\|_1} = \frac{\|\lambda X\|_1}{\|X\|_1} = |\lambda|. \tag{4.183}$$

The required bound on λ holds, which completes the proof. □

While the spectral radius of the natural representation $K(\Phi)$ of every channel $\Phi \in \mathrm{C}(\mathcal{X})$ must equal 1, it will not generally be the case that the spectral norm of $K(\Phi)$ will be 1. As the following theorem establishes, this happens if and only if Φ is a unital channel. Like Theorem 4.24, the property of complete positivity is not needed in the proof of this fact, and so it holds not only for channels, but for all positive and trace-preserving maps.

Theorem 4.27. *Let $\mathcal{X}$ be a complex Euclidean space and let $\Phi \in \mathrm{T}(\mathcal{X})$ be a positive and trace-preserving map. It holds that Φ is unital if and only if $\|K(\Phi)\| = 1$.*

Proof. Suppose that Φ is a unital channel. It is evident that $\|K(\Phi)\| \geq 1$, as Theorem 4.24 implies that

$$K(\Phi)\operatorname{vec}(\rho) = \operatorname{vec}(\rho) \tag{4.184}$$

for some choice of a density operator $\rho \in \mathrm{D}(\mathcal{X})$. It therefore suffices to prove that $\|K(\Phi)\| \leq 1$, which is equivalent to the condition that

$$\big\|\Phi(X)\big\|_2 \leq \big\|X\big\|_2 \tag{4.185}$$

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

Consider first an arbitrary Hermitian operator $H \in \mathrm{Herm}(\mathcal{X})$. Let

$$H = \sum_{k=1}^{n} \lambda_k x_k x_k^* \tag{4.186}$$

be a spectral decomposition of H, for $n = \dim(\mathcal{X})$, and let

$$\rho_k = \Phi(x_k x_k^*) \tag{4.187}$$

for each $k \in \{1, \ldots, n\}$. One has that $\rho_1, \ldots, \rho_n$ are density operators, as a consequence of the fact that Φ is positive and trace-preserving. Moreover, as Φ is unital, it follows that $\rho_1 + \cdots + \rho_n = \mathbb{1}$. It holds that

$$\big\|\Phi(H)\big\|_2^2 = \big\|\lambda_1\rho_1 + \cdots + \lambda_n\rho_n\big\|_2^2 = \sum_{1 \leq j,k \leq n} \lambda_j\lambda_k\langle\rho_j, \rho_k\rangle. \tag{4.188}$$

The Cauchy–Schwarz inequality implies that

$$\begin{aligned} &\sum_{1 \leq j,k \leq n} \lambda_j\lambda_k\langle\rho_j, \rho_k\rangle \\ &\leq \sqrt{\sum_{1 \leq j,k \leq n} \lambda_j^2\langle\rho_j, \rho_k\rangle}\sqrt{\sum_{1 \leq j,k \leq n} \lambda_k^2\langle\rho_j, \rho_k\rangle} = \sum_{k=1}^{n} \lambda_k^2 = \big\|H\big\|_2^2, \end{aligned} \tag{4.189}$$

where the first equality has followed from the fact that $\rho_1 + \cdots + \rho_n = \mathbb{1}$. It has therefore been established that $\|\Phi(H)\|_2 \leq \|H\|_2$ for all Hermitian operators $H \in \mathrm{Herm}(\mathcal{X})$.

Now consider any operator $X \in \mathrm{L}(\mathcal{X})$, written as $X = H + iK$ for

$$H = \frac{X + X^*}{2} \quad \text{and} \quad K = \frac{X - X^*}{2i} \tag{4.190}$$

being Hermitian operators, and observe that

$$\|X\|_2^2 = \|H\|_2^2 + \|K\|_2^2. \tag{4.191}$$

As Φ is necessarily Hermiticity-preserving, one finds that

$$\|\Phi(X)\|_2^2 = \|\Phi(H) + i\Phi(K)\|_2^2 = \|\Phi(H)\|_2^2 + \|\Phi(K)\|_2^2. \tag{4.192}$$

Therefore

$$\|\Phi(X)\|_2^2 = \|\Phi(H)\|_2^2 + \|\Phi(K)\|_2^2 \leq \|H\|_2^2 + \|K\|_2^2 = \|X\|_2^2, \tag{4.193}$$

so $\|\Phi(X)\|_2 \leq \|X\|_2$, as required.

Now suppose that Φ is a positive and trace-preserving map for which $\|K(\Phi)\| = 1$, which is equivalent to the condition that $\|\Phi(X)\|_2 \leq \|X\|_2$ for every $X \in \mathrm{L}(\mathcal{X})$. In particular, it must hold that

$$\|\Phi(\mathbb{1})\|_2 \leq \|\mathbb{1}\|_2 = \sqrt{n}, \tag{4.194}$$

for $n = \dim(\mathcal{X})$. As Φ is positive and trace-preserving, one has that $\Phi(\mathbb{1})$ is positive semidefinite and has trace equal to n. When these observations are combined with the Cauchy–Schwarz inequality, one finds that

$$n = \mathrm{Tr}\big(\Phi(\mathbb{1})\big) = \langle \mathbb{1}, \Phi(\mathbb{1}) \rangle \leq \|\mathbb{1}\|_2 \, \|\Phi(\mathbb{1})\|_2 \leq n. \tag{4.195}$$

Equality is therefore obtained in the Cauchy–Schwarz inequality, implying that $\Phi(\mathbb{1})$ and $\mathbb{1}$ are linearly dependent. As $\mathrm{Tr}(\mathbb{1}) = \mathrm{Tr}(\Phi(\mathbb{1}))$, it follows that $\Phi(\mathbb{1})$ and $\mathbb{1}$ must in fact be equal, and therefore Φ is unital. □

4.3 Majorization

This section introduces the *majorization* relation for Hermitian operators, which is a generalization of a similar concept for real vectors. Intuitively speaking, the majorization relation formalizes the notion of one object being obtained from another through a "random mixing process."

One may formalize the majorization relation, both for real vectors and for Hermitian operators, in multiple, equivalent ways. Once formalized, it is a very useful mathematical concept. In the theory of quantum information, majorization has a particularly striking application in the form of Nielsen's theorem (Theorem 6.37 in Chapter 6), which gives a precise characterization of the possible transformations between bipartite pure states that may be performed by two individuals whose communications with one another are restricted to classical information transmissions.

4.3.1 Majorization for real vectors

The definition of the majorization relation for real vectors to be presented in this book is based on the class of *doubly stochastic* operators. A discussion of such operators follows, after which the majorization relation for real vectors is defined.

Doubly stochastic operators

Let Σ be an alphabet, and consider the real Euclidean space $\mathbb{R}^{\Sigma}$. An operator $A \in \mathrm{L}(\mathbb{R}^{\Sigma})$ acting on this vector space is said to be *stochastic* if

1. $A(a,b) \geq 0$ for each $(a,b) \in \Sigma \times \Sigma$, and
2. $\sum_{a \in \Sigma} A(a,b) = 1$ for each $b \in \Sigma$.

This condition is equivalent to Ae_b being a probability vector for each $b \in \Sigma$, or equivalently, that A maps probability vectors to probability vectors. An operator $A \in \mathrm{L}(\mathbb{R}^{\Sigma})$ is said to be *doubly stochastic* if

1. $A(a,b) \geq 0$ for each $(a,b) \in \Sigma \times \Sigma$,
2. $\sum_{a \in \Sigma} A(a,b) = 1$ for each $b \in \Sigma$, and
3. $\sum_{b \in \Sigma} A(a,b) = 1$ for each $a \in \Sigma$.

That is, an operator A is doubly stochastic if and only if both A and A^{T} (or, equivalently, both A and A^*) are stochastic, which is equivalent to the condition that every row and every column of the matrix representation of A forms a probability vector.

Doubly stochastic operators have a close relationship to *permutation operators*. For each permutation $\pi \in \mathrm{Sym}(\Sigma)$, one defines the permutation operator $V_\pi \in \mathrm{L}(\mathbb{R}^\Sigma)$ as

$$V_\pi(a,b) = \begin{cases} 1 & \text{if } a = \pi(b) \\ 0 & \text{otherwise} \end{cases} \tag{4.196}$$

for every $(a,b) \in \Sigma \times \Sigma$. Equivalently, V_π is the unique operator satisfying the equation $V_\pi e_b = e_{\pi(b)}$ for each $b \in \Sigma$. It is evident that permutation operators are doubly stochastic. The next theorem establishes that the set of all doubly stochastic operators is, in fact, equal to the convex hull of the permutation operators.

Theorem 4.28 (Birkhoff–von Neumann theorem)**.** *Let Σ be an alphabet and let $A \in \mathrm{L}(\mathbb{R}^\Sigma)$ be an operator. It holds that A is doubly stochastic if and only if there exists a probability vector $p \in \mathcal{P}(\mathrm{Sym}(\Sigma))$ such that*

$$A = \sum_{\pi \in \mathrm{Sym}(\Sigma)} p(\pi) V_\pi. \tag{4.197}$$

Proof. The set of all doubly stochastic operators acting on $\mathbb{R}^\Sigma$ is convex and compact, and is therefore equal to the convex hull of its extreme points by Theorem 1.10. The theorem will therefore follow from the demonstration that every extreme point in this set is a permutation operator. With this fact in mind, let A be a doubly stochastic operator that is not a permutation operator. It will be proved that A is not an extreme point of the set of doubly stochastic operators, which is sufficient to complete the proof.

Given that A is doubly stochastic but not a permutation operator, there must exist at least one pair $(a_1, b_1) \in \Sigma \times \Sigma$ such that $A(a_1, b_1) \in (0,1)$. As $\sum_b A(a_1, b) = 1$ and $A(a_1, b_1) \in (0,1)$, one may conclude that there exists an index $b_2 \neq b_1$ such that $A(a_1, b_2) \in (0,1)$. Applying similar reasoning, but to the first index rather than the second, it follows that there must exist an index $a_2 \neq a_1$ such that $A(a_2, b_2) \in (0,1)$. Repeating this argument, one may eventually find a closed loop of even length among the entries of A that are contained in the interval $(0,1)$, alternating between the first and

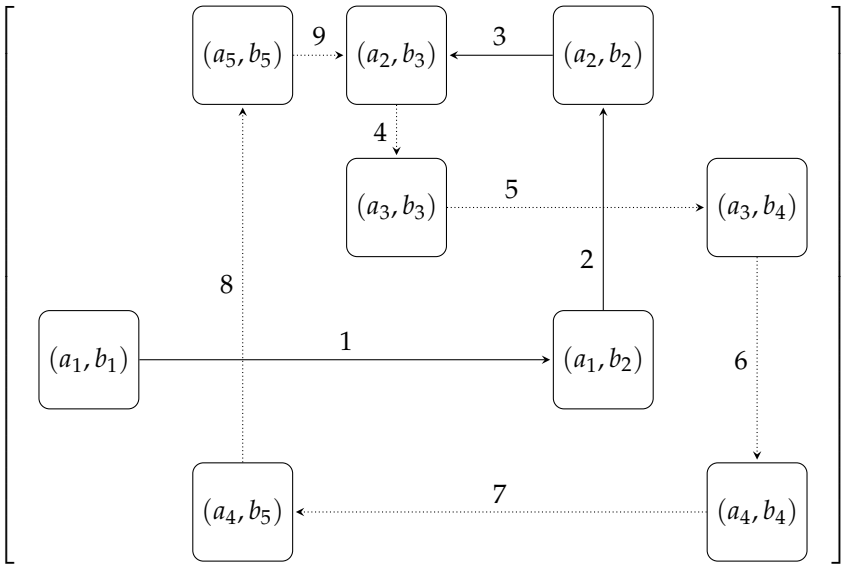

Figure 4.1: An example of a closed loop consisting of entries of A that are contained in the interval $(0, 1)$. The loop is indicated by the dotted lines.

second indices (i.e., between rows and columns). A loop must eventually be formed, given that there are only finitely many entries in the matrix A; and an odd-length loop can be avoided by an appropriate choice for the entry that closes the loop. This process is illustrated in Figure 4.1.

Let $\varepsilon \in (0, 1)$ be equal to the minimum value over the entries in a closed loop of the form just described, and define B to be the operator obtained by setting each entry in the closed loop to be $\pm\varepsilon$, alternating sign among the entries as suggested in Figure 4.2. All of the other entries in B are set to 0. Finally, consider the operators $A + B$ and $A - B$. As A is doubly stochastic and the row and column sums of B are all 0, it holds that both $A + B$ and $A - B$ also have row and column sums equal to 1. As ε was chosen to be no larger than the smallest entry within the chosen closed loop, none of the entries of $A + B$ or $A - B$ are negative, and therefore $A - B$ and $A + B$ are doubly stochastic. As B is nonzero, it holds that $A + B$ and $A - B$ are distinct. Thus,

$$A = \frac{1}{2}(A + B) + \frac{1}{2}(A - B) \tag{4.198}$$

is a proper convex combination of doubly stochastic operators, and is therefore not an extreme point in the set of doubly stochastic operators. □

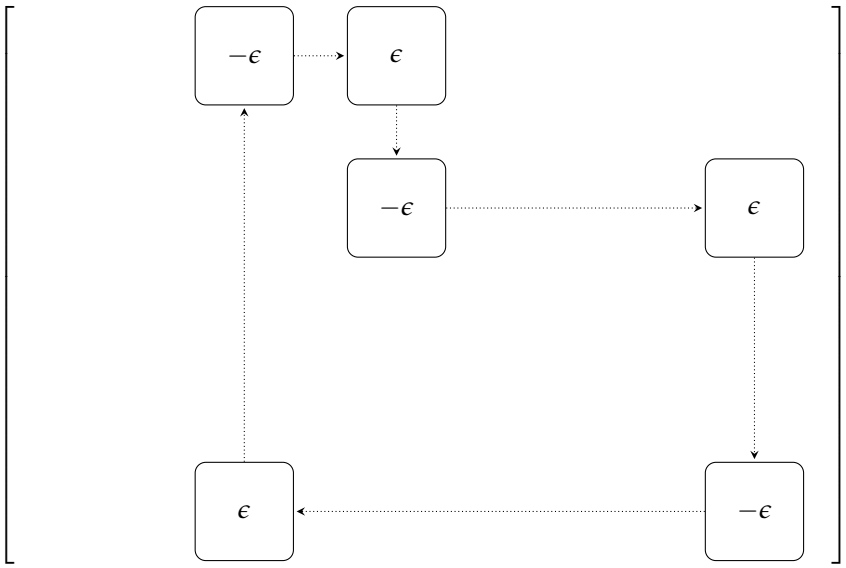

Figure 4.2: The operator B. All entries besides those indicated are 0.

Definition and characterizations of majorization for real vectors

A definition of the majorization relation for vectors of real numbers, based on the actions of doubly stochastic operators, is as follows.

Definition 4.29. Let Σ be an alphabet and let $u, v \in \mathbb{R}^{\Sigma}$ be vectors. It is said that u *majorizes* v, written $v \prec u$, if there exists a doubly stochastic operator $A \in \mathrm{L}\big(\mathbb{R}^{\Sigma}\big)$ for which $v = Au$.

By the Birkhoff–von Neumann theorem (Theorem 4.28), one may view this definition as formalizing the sort of "random mixing process" suggested at the beginning of the current section. An operator A is doubly stochastic if and only if it is equal to a convex combination of permutation operators, so the relation $v \prec u$ holds precisely when v can be obtained by randomly choosing a permutation $\pi \in \mathrm{Sym}(\Sigma)$, with respect to a chosen distribution $p \in \mathcal{P}(\mathrm{Sym}(\Sigma))$, shuffling the entries of u in accordance with the chosen permutation π, and then averaging the resulting vectors with respect to p.

The following theorem provides two alternative characterizations of the majorization relation for real vectors. The statement of the theorem makes use of the following notation: for every vector $u \in \mathbb{R}^{\Sigma}$ and for $n = |\Sigma|$, one

writes

$$r(u) = (r_1(u), \ldots, r_n(u)) \tag{4.199}$$

to denote the vector obtained by *sorting* the entries of u in decreasing order. In other words, one has

$$\{u(a) \,:\, a \in \Sigma\} = \{r_1(u), \ldots, r_n(u)\}, \tag{4.200}$$

where the equality considers the two sides of the equation to be multisets, and moreover

$$r_1(u) \geq \cdots \geq r_n(u). \tag{4.201}$$

Theorem 4.30. *Let Σ be an alphabet and let $u, v \in \mathbb{R}^{\Sigma}$. The following statements are equivalent:*

1. $v \prec u$.
2. *For $n = |\Sigma|$, one has*

$$r_1(u) + \cdots + r_m(u) \geq r_1(v) + \cdots + r_m(v) \tag{4.202}$$

   for every choice of $m \in \{1, \ldots, n-1\}$, as well as

$$r_1(u) + \cdots + r_n(u) = r_1(v) + \cdots + r_n(v). \tag{4.203}$$

3. *There exists a unitary operator $U \in \mathrm{U}(\mathbb{C}^{\Sigma})$ such that, for the doubly stochastic operator $A \in \mathrm{L}(\mathbb{R}^{\Sigma})$ defined as*

$$A(a,b) = |U(a,b)|^2 \tag{4.204}$$

   for each $(a,b) \in \Sigma \times \Sigma$, one has $v = Au$.

Proof. Assume first that statement 1 holds, so that there exists a doubly stochastic operator $A \in \mathrm{L}(\mathbb{R}^{\Sigma})$ such that $Au = v$. It will be proved that

$$\sum_{a \in \Sigma} u(a) = \sum_{a \in \Sigma} v(a), \tag{4.205}$$

and that, for every subset $S \subseteq \Sigma$, there exists a subset $T \subseteq \Sigma$ such that $|S| = |T|$ and

$$\sum_{a \in T} u(a) \geq \sum_{a \in S} v(a). \tag{4.206}$$

This will imply statement 2; the condition (4.205) is equivalent to (4.203), while (4.206) implies (4.202) when one considers the case that S comprises the indices of the m largest entries of v, for each $m \in \{1, \ldots, n-1\}$. The first condition (4.205) is immediate from the assumption that A is stochastic:

$$\sum_{a \in \Sigma} v(a) = \sum_{a \in \Sigma} (Au)(a) = \sum_{a,b \in \Sigma} A(a,b)u(b) = \sum_{b \in \Sigma} u(b). \tag{4.207}$$

To prove the second condition, observe first that the Birkhoff–von Neumann theorem (Theorem 4.28) implies that

$$A = \sum_{\pi \in \mathrm{Sym}(\Sigma)} p(\pi) V_\pi \tag{4.208}$$

for some choice of a probability vector $p \in \mathcal{P}(\mathrm{Sym}(\Sigma))$. For an arbitrary choice of a subset $S \subseteq \Sigma$, the expression (4.208) implies that

$$\sum_{a \in S} v(a) = \sum_{a \in S} (Au)(a) = \sum_{\pi \in \mathrm{Sym}(\Sigma)} p(\pi) \sum_{a \in S} u\big(\pi^{-1}(a)\big) \tag{4.209}$$

A convex combination of a collection of real numbers cannot exceed the maximal element in that set, and therefore there must exist a permutation $\pi \in \mathrm{Sym}(\Sigma)$ such that

$$\sum_{a \in T_\pi} u(a) = \sum_{a \in S} u\big(\pi^{-1}(a)\big) \geq \sum_{a \in S} v(a) \tag{4.210}$$

for $T_\pi = \big\{\pi^{-1}(a) \,:\, a \in S\big\}$. As $|T_\pi| = |S|$, the inequality (4.206) has been proved for a suitable choice of an index set T. It has therefore been proved that statement 1 implies statement 2.

Next it will be proved that statement 2 implies statement 3, which is the most difficult implication of the proof. The implication will be proved by induction on $n = |\Sigma|$, for which the base case $n = 1$ is trivial. It will therefore be assumed that $n \geq 2$ for the remainder of the proof. As the majorization relationship is invariant under renaming and independently reordering the indices of the vectors under consideration, there is no loss of generality in assuming that $\Sigma = \{1, \ldots, n\}$, that $u = (u_1, \ldots, u_n)$ satisfies $u_1 \geq \cdots \geq u_n$, and that $v = (v_1, \ldots, v_n)$ satisfies $v_1 \geq \cdots \geq v_n$.

Under the assumption that statement 2 holds, it must be the case that $u_1 \geq v_1 \geq u_k$ for some choice of $k \in \{1, \ldots, n\}$. Fix k to be minimal among all such indices. There are two cases: $k = 1$ and $k > 1$.

If it is the case that $k = 1$, then $u_1 = v_1$, from which it follows that

$$u_2 + \cdots + u_m \geq v_2 + \cdots + v_m \tag{4.211}$$

for every $m \in \{2, \ldots, n-1\}$, as well as

$$u_2 + \cdots + u_n = v_2 + \cdots + v_n. \tag{4.212}$$

Define vectors $x = (u_2, \ldots, u_n)$ and $y = (v_2, \ldots, v_n)$. By the hypothesis of induction, there must therefore exist a unitary operator V, whose entries are indexed by the set $\{2, \ldots, n\}$, having the property that the doubly stochastic operator B defined by

$$B(a, b) = |V(a, b)|^2 \tag{4.213}$$

for all $a, b \in \{2, \ldots, n\}$ satisfies $y = Bx$. Taking U to be the unitary operator

$$U = \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix} \tag{4.214}$$

and letting A be defined by

$$A(a, b) = |U(a, b)|^2 \tag{4.215}$$

for all $a, b \in \{1, \ldots, n\}$, one has that $v = Au$, as required.

If it is the case that $k > 1$, then $u_1 > v_1 \geq u_k$, and so there must exist a real number $\lambda \in [0, 1)$ such that $v_1 = \lambda u_1 + (1 - \lambda) u_k$. Define vectors $x = (x_2, \ldots, x_n)$ and $y = (y_2, \ldots, y_n)$ as

$$\begin{aligned} x &= (u_2, \ldots, u_{k-1}, (1 - \lambda) u_1 + \lambda u_k, u_{k+1}, \ldots, u_n), \\ y &= (v_2, \ldots, v_n). \end{aligned} \tag{4.216}$$

For $m \in \{2, \ldots, k-1\}$ it holds that

$$x_2 + \cdots + x_m = u_2 + \cdots + u_m > (m - 1) v_1 \geq v_2 + \cdots + v_m, \tag{4.217}$$

by virtue of the fact that k is the minimal index for which $v_1 \geq u_k$. For $m \in \{k, \ldots, n\}$ it holds that

$$\begin{aligned} &x_2 + \cdots + x_m \\ &\quad = (1 - \lambda) u_1 + u_2 + \cdots + u_{k-1} + \lambda u_k + u_{k+1} + \cdots + u_m \\ &\quad = u_1 + \cdots + u_m - v_1 \geq v_1 + \cdots + v_m - v_1 = v_2 + \cdots + v_m. \end{aligned} \tag{4.218}$$

By the hypothesis of induction, there must therefore exist a unitary operator V, whose entries are indexed by the set $\{2,\ldots,n\}$, having the property that the doubly stochastic operator B defined by

$$B(a,b) = |V(a,b)|^2 \tag{4.219}$$

for every $a,b \in \{2,\ldots,n\}$ satisfies $y = Bx$. Let W be the unitary operator defined by

$$\begin{aligned} We_1 &= \sqrt{\lambda}e_1 - \sqrt{1-\lambda}e_k, \\ We_k &= \sqrt{1-\lambda}e_1 + \sqrt{\lambda}e_k, \end{aligned} \tag{4.220}$$

and $We_a = e_a$ for $a \in \{2,\ldots,n\}\backslash\{k\}$, and let

$$U = \begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix} W. \tag{4.221}$$

The entries of U may be calculated explicitly—they are given by

$$\begin{aligned} U(1,1) &= \sqrt{\lambda} & U(a,1) &= -\sqrt{1-\lambda}V(a,k) \\ U(1,k) &= \sqrt{1-\lambda} & U(a,k) &= \sqrt{\lambda}V(a,k) \\ U(1,b) &= 0 & U(a,b) &= V(a,b) \end{aligned} \tag{4.222}$$

for $a \in \{2,\ldots,n\}$ and $b \in \{2,\ldots,n\}\backslash\{k\}$. Letting A be the doubly stochastic operator defined by

$$A(a,b) = |U(a,b)|^2 \tag{4.223}$$

for every $a,b \in \{1,\ldots,n\}$, one obtains an operator whose entries are given by

$$\begin{aligned} A(1,1) &= \lambda & A(a,1) &= (1-\lambda)B(a,k) \\ A(1,k) &= 1-\lambda & A(a,k) &= \lambda B(a,k) \\ A(1,b) &= 0 & A(a,b) &= B(a,b) \end{aligned} \tag{4.224}$$

for $a \in \{2,\ldots,n\}$ and $b \in \{2,\ldots,n\}\backslash\{k\}$. Equivalently,

$$A = \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix} D, \tag{4.225}$$

for D being the doubly stochastic operator defined by

$$\begin{aligned} De_1 &= \lambda e_1 + (1-\lambda)e_k, \\ De_k &= (1-\lambda)e_1 + \lambda e_k, \end{aligned} \tag{4.226}$$

and $De_a = e_a$ for $a \in \{2, \ldots, m\} \backslash \{k\}$. It holds that

$$Du = \begin{pmatrix} v_1 \\ x \end{pmatrix} \tag{4.227}$$

and therefore

$$Au = \begin{pmatrix} v_1 \\ Bx \end{pmatrix} = v. \tag{4.228}$$

It has therefore been proved that statement 2 implies statement 3.

The final step is to observe that statement 3 implies statement 1, which is trivial, as the operator A determined by statement 3 must be doubly stochastic. □

Remark 4.31. In light of the equivalence between the first and third statements in Theorem 4.30, it is natural to ask if every doubly stochastic operator $A \in \mathrm{L}(\mathbb{R}^\Sigma)$ is given by $A(a,b) = |U(a,b)|^2$ for some choice of a unitary operator $U \in \mathrm{U}(\mathbb{C}^\Sigma)$. This is not the case: the operator

$$A = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \tag{4.229}$$

in $\mathrm{L}(\mathbb{R}^3)$ is an example of a doubly stochastic operator that cannot be derived from a unitary operator in this fashion. Indeed, if A is to be derived from a unitary operator $U \in \mathrm{U}(\mathbb{C}^3)$, then U must take the form

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & \alpha_2 & \alpha_1 \\ \alpha_3 & 0 & \beta_1 \\ \beta_3 & \beta_2 & 0 \end{pmatrix} \tag{4.230}$$

for $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2$, and β_3 being complex numbers on the unit circle. However, if U is unitary, then it must hold that

$$\mathbb{1} = UU^* = \frac{1}{2} \begin{pmatrix} |\alpha_1|^2 + |\alpha_2|^2 & \alpha_1\overline{\beta_1} & \alpha_2\overline{\beta_2} \\ \overline{\alpha_1}\beta_1 & |\alpha_3|^2 + |\beta_1|^2 & \alpha_3\overline{\beta_3} \\ \overline{\alpha_2}\beta_2 & \overline{\alpha_3}\beta_3 & |\beta_2|^2 + |\beta_3|^2 \end{pmatrix}. \tag{4.231}$$

This, however, is impossible, as none of the off-diagonal entries of the operator on the right-hand-side of (4.231) can equal zero for $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2$, and β_3 being complex numbers on the unit circle.

4.3.2 Majorization for Hermitian operators

The majorization relation for Hermitian operators will now be defined. This relation inherits the essential characteristics of its real vector analogue; and similar to its real vector analogue it may be characterized in multiple ways. After a discussion of its alternative characterizations, two applications of majorization for Hermitian operators will be presented.

Definition and characterizations of majorization for Hermitian operators

In analogy to the intuitive description of the majorization relation for real vectors suggested previously, one may view that one Hermitian operator X majorizes another Hermitian operator Y if and only if it is the case that Y can be obtained from X through a "random mixing" process. One natural way to formalize the notion of "random mixing" in the quantum setting is to consider mixed-unitary channels to be representative of such processes. The definition of the majorization relation for Hermitian operators that follows adopts this viewpoint.

Definition 4.32. Let $X, Y \in \mathrm{Herm}(\mathcal{X})$ be Hermitian operators, for a complex Euclidean space $\mathcal{X}$. It is said that X *majorizes* Y, written $Y \prec X$, if there exists a mixed-unitary channel $\Phi \in \mathrm{C}(\mathcal{X})$ for which $\Phi(X) = Y$.

There is, *a priori*, no reason to prefer Definition 4.32 over one possible alternative, in which the condition that Φ is mixed-unitary is replaced by the condition that Φ is a unital channel. This is indeed a natural alternative because unital channels are, in some sense, analogous to doubly stochastic operators acting on real Euclidean spaces, while mixed-unitary channels are analogous to convex combinations of permutation operators. The failure of a direct quantum analogue to the Birkhoff–von Neumann theorem to hold is responsible for this apparent difference between two possible definitions of majorization for Hermitian operators.

The following theorem demonstrates that these two alternatives are, in fact, equivalent. The theorem also provides two additional characterizations of the majorization relation for Hermitian operators, the second of which establishes a direct link between majorization for Hermitian operators and majorization for real vectors (applied to the vectors of eigenvalues of the Hermitian operators under consideration).

Theorem 4.33 (Uhlmann). *Let $X, Y \in \mathrm{Herm}(\mathcal{X})$ be Hermitian operators, for a complex Euclidean space $\mathcal{X}$. The following statements are equivalent:*

1. $Y \prec X$.
2. *There exists a unital channel $\Phi \in \mathrm{C}(\mathcal{X})$ such that $Y = \Phi(X)$.*
3. *There exists a positive, trace-preserving, and unital map $\Phi \in \mathrm{T}(\mathcal{X})$ such that $Y = \Phi(X)$.*
4. $\lambda(Y) \prec \lambda(X)$.

Proof. Under the assumption that statement 1 holds, there exists a mixed-unitary channel $\Phi \in \mathrm{C}(\mathcal{X})$ such that $Y = \Phi(X)$. Any such channel is necessarily unital, and therefore statement 1 trivially implies statement 2. As every unital channel is positive, trace-preserving, and unital, statement 2 trivially implies statement 3.

Now assume that statement 3 holds. Let $n = \dim(\mathcal{X})$, and let

$$X = \sum_{j=1}^{n} \lambda_j(X)\, x_j x_j^* \quad \text{and} \quad Y = \sum_{k=1}^{n} \lambda_k(Y)\, y_k y_k^* \tag{4.232}$$

be spectral decompositions of X and Y, respectively. As $\Phi(X) = Y$, one concludes that

$$\lambda_k(Y) = \sum_{j=1}^{n} \lambda_j(X)\, y_k^* \Phi(x_j x_j^*) y_k \tag{4.233}$$

for each $k \in \{1, \ldots, n\}$. Equivalently, $\lambda(Y) = A\lambda(X)$ for $A \in \mathrm{L}(\mathbb{R}^n)$ being the operator defined as

$$A(k, j) = y_k^* \Phi(x_j x_j^*) y_k \tag{4.234}$$

for every $j, k \in \{1, \ldots, n\}$. Each entry of A is nonnegative by the positivity of Φ; by the fact that Φ is trace-preserving, it holds that

$$\sum_{k=1}^{n} A(k, j) = 1 \tag{4.235}$$

for each $j \in \{1, \ldots, n\}$; and by the fact that Φ is unital, it holds that

$$\sum_{j=1}^{n} A(k, j) = 1 \tag{4.236}$$

for each $k \in \{1, \ldots, n\}$. The operator A is therefore doubly stochastic, so that $\lambda(Y) \prec \lambda(X)$. It has therefore been proved that statement 3 implies statement 4.

Finally, assume $\lambda(Y) \prec \lambda(X)$, and again consider spectral decompositions of X and Y as in (4.232). As $\lambda(Y) \prec \lambda(X)$, one may conclude from Theorem 4.28 that there exists a probability vector $p \in \mathcal{P}(S_n)$ such that

$$\lambda_k(Y) = \sum_{\pi \in S_n} p(\pi) \lambda_{\pi(k)}(X) \tag{4.237}$$

for all $k \in \{1, \ldots, n\}$. By defining a unitary operator

$$U_\pi = \sum_{k=1}^{n} y_k x_{\pi(k)}^* \tag{4.238}$$

for each permutation $\pi \in S_n$, one has that

$$\begin{aligned} &\sum_{\pi \in S_n} p(\pi) U_\pi X U_\pi^* \\ &\quad = \sum_{k=1}^{n} \sum_{\pi \in S_n} p(\pi) \lambda_{\pi(k)}(X) y_k y_k^* = \sum_{k=1}^{n} \lambda_k(Y) y_k y_k^* = Y. \end{aligned} \tag{4.239}$$

It therefore holds that $Y \prec X$, and so statement 4 implies statement 1, which completes the proof. □

Two applications of Hermitian operator majorization

The theorems that follow offer a sample of the applications of majorization for Hermitian operators. The first theorem, whose proof makes essential use of Theorem 4.33, provides a precise characterization of those real vectors that may be obtained as the diagonal entries of a given Hermitian operator with respect to an arbitrary choice of an orthonormal basis.

Theorem 4.34 (Schur–Horn theorem)**.** *Let $\mathcal{X}$ be a complex Euclidean space, let $n = \dim(\mathcal{X})$, and let $X \in \mathrm{Herm}(\mathcal{X})$ be a Hermitian operator. The following two implications, which are converse to one another, hold:*

1. *For every orthonormal basis $\{x_1, \ldots, x_n\}$ of $\mathcal{X}$, the vector $v \in \mathbb{R}^n$ defined by $v(k) = x_k^* X x_k$ for each $k \in \{1, \ldots, n\}$ satisfies $v \prec \lambda(X)$.*
2. *For every vector $v \in \mathbb{R}^n$ satisfying $v \prec \lambda(X)$, there exists an orthonormal basis $\{x_1, \ldots, x_n\}$ of $\mathcal{X}$ for which $v(k) = x_k^* X x_k$ for each $k \in \{1, \ldots, n\}$.*

Proof. Suppose $\{x_1, \ldots, x_n\}$ is an orthonormal basis of $\mathcal{X}$ and $v \in \mathbb{R}^n$ is defined as $v(k) = x_k^* X x_k$ for each $k \in \{1, \ldots, n\}$. Define a map $\Phi \in \mathrm{T}(\mathcal{X})$ as

$$\Phi(Y) = \sum_{k=1}^{n} x_k x_k^* Y x_k x_k^* \tag{4.240}$$

for every operator $Y \in \mathrm{L}(\mathcal{X})$, and observe that Φ is a pinching channel. By Proposition 4.6, it follows that Φ is a mixed-unitary channel. One therefore has $\Phi(X) \prec X$, which implies $\lambda(\Phi(X)) \prec \lambda(X)$ by Theorem 4.33. As

$$\Phi(X) = \sum_{k=1}^{n} v(k) x_k x_k^*, \tag{4.241}$$

it is evident that

$$\operatorname{spec}(\Phi(X)) = \{v(1), \ldots, v(n)\}, \tag{4.242}$$

or equivalently that

$$\lambda(\Phi(X)) = W_\pi v \tag{4.243}$$

for a permutation operator W_π that has the effect of ordering the entries of v from largest to smallest:

$$(W_\pi v)(1) \geq \cdots \geq (W_\pi v)(n). \tag{4.244}$$

It follows that $v \prec \lambda(X)$, as is required to establish the first implication.

Now suppose $v \in \mathbb{R}^n$ is a vector satisfying $v \prec \lambda(X)$, and let

$$X = \sum_{k=1}^{n} \lambda_k(X) u_k u_k^* \tag{4.245}$$

be a spectral decomposition of X. By Theorem 4.30, the assumption that $v \prec \lambda(X)$ implies that there exists a unitary operator $U \in \mathrm{U}(\mathbb{C}^n)$ such that, for $A \in \mathrm{L}(\mathbb{R}^n)$ defined by

$$A(j,k) = |U(j,k)|^2 \tag{4.246}$$

for $j, k \in \{1, \ldots, n\}$, one has $v = A\lambda(X)$. Define $V \in \mathrm{U}(\mathcal{X}, \mathbb{C}^n)$ as

$$V = \sum_{k=1}^{n} e_k u_k^* \tag{4.247}$$

and let

$$x_k = V^* U^* V u_k \tag{4.248}$$

for each $k \in \{1, \ldots, n\}$. The operator $V^* U^* V \in \mathrm{U}(\mathcal{X})$ is a unitary operator, implying that $\{x_1, \ldots, x_n\}$ is an orthonormal basis of $\mathcal{X}$. It holds that

$$x_k^* X x_k = \sum_{j=1}^{n} |U(k,j)|^2 \lambda_j(X) = (A\lambda(X))(k) = v(k), \tag{4.249}$$

which establishes the second implication. □

The next theorem, representing a second application of majorization for Hermitian operators, characterizes the collection of probability vectors that are consistent with the representation of a given density operator as a mixture of pure states.

Theorem 4.35. *Let $\mathcal{X}$ be a complex Euclidean space, let $\rho \in \mathrm{D}(\mathcal{X})$ be a density operator, let $n = \dim(\mathcal{X})$, and let $p = (p_1, \ldots, p_n)$ be a probability vector. There exist a collection of (not necessarily orthogonal) unit vectors $\{u_1, \ldots, u_n\} \subset \mathcal{X}$ such that*

$$\rho = \sum_{k=1}^{n} p_k u_k u_k^* \tag{4.250}$$

if and only if $p \prec \lambda(\rho)$.

Proof. Assume first that

$$\rho = \sum_{k=1}^{n} p_k u_k u_k^* \tag{4.251}$$

for a collection $\{u_1, \ldots, u_n\} \subset \mathcal{X}$ of unit vectors. Define $A \in \mathrm{L}(\mathbb{C}^n, \mathcal{X})$ as

$$A = \sum_{k=1}^{n} \sqrt{p_k}\, u_k e_k^*, \tag{4.252}$$

and observe that $AA^* = \rho$. It holds that

$$A^* A = \sum_{j=1}^{n} \sum_{k=1}^{n} \sqrt{p_j p_k}\, \langle u_k, u_j \rangle\, E_{k,j}\,, \tag{4.253}$$

and therefore

$$e_k^* A^* A e_k = p_k \tag{4.254}$$

for every $k \in \{1, \ldots, n\}$. By Theorem 4.34, this implies $p \prec \lambda(A^* A)$. As

$$\lambda(A^* A) = \lambda(AA^*) = \lambda(\rho), \tag{4.255}$$

it follows that $p \prec \lambda(\rho)$. One of the required implications of the theorem has therefore been proved.

Now assume that $p \prec \lambda(\rho)$. By Theorem 4.34, there exists an orthonormal basis $\{x_1, \ldots, x_n\}$ of $\mathcal{X}$ with the property that

$$p_k = x_k^* \rho x_k \tag{4.256}$$

for each $k \in \{1, \ldots, n\}$. Let

$$y_k = \sqrt{\rho} x_k \tag{4.257}$$

and define

$$u_k = \begin{cases} \frac{y_k}{\|y_k\|} & \text{if } y_k \neq 0 \\ z & \text{if } y_k = 0 \end{cases} \tag{4.258}$$

for each $k \in \{1, \ldots, n\}$, where $z \in \mathcal{X}$ is an arbitrarily chosen unit vector. One has that

$$\|y_k\|^2 = \left\langle \sqrt{\rho} x_k, \sqrt{\rho} x_k \right\rangle = x_k^* \rho x_k = p_k, \tag{4.259}$$

for each $k \in \{1, \ldots, n\}$, and therefore

$$\sum_{k=1}^{n} p_k u_k u_k^* = \sum_{k=1}^{n} y_k y_k^* = \sum_{k=1}^{n} \sqrt{\rho} x_k x_k^* \sqrt{\rho} = \rho. \tag{4.260}$$

This proves the other required implication of the theorem. □

4.4 Exercises

4.1. Let $\mathcal{X}$ be a complex Euclidean space with $\dim(\mathcal{X}) = 3$ and let $\Phi \in \mathrm{C}(\mathcal{X})$ be a Schur channel. Prove that Φ is a mixed-unitary channel.

4.2. For every positive integer $n \geq 2$, define a unital channel $\Phi_n \in \mathrm{C}(\mathbb{C}^n)$ as

$$\Phi_n(X) = \frac{\mathrm{Tr}(X)\mathbb{1}_n - X^{\mathsf{T}}}{n-1} \tag{4.261}$$

for every $X \in \mathrm{L}(\mathbb{C}^n)$, where $\mathbb{1}_n$ denotes the identity operator on $\mathbb{C}^n$. Prove that Φ_n is not mixed-unitary when n is odd.

A correct solution to this exercise generalizes Example 4.3, but a different argument will be needed than the one in that example when $n \geq 5$.

4.3. Let n be a positive integer, let $\mathcal{X} = \mathbb{C}^{\mathbb{Z}_n}$, let

$$\{W_{a,b} \,:\, a, b \in \mathbb{Z}_n\} \subset \mathrm{U}(\mathcal{X}) \tag{4.262}$$

be the set of discrete Weyl operators acting on $\mathcal{X}$, and let $\Phi \in \mathrm{C}(\mathcal{X})$ be a channel. Prove that the following two statements are equivalent:

1. Φ is both a Schur channel and a Weyl-covariant channel.
2. There exists a probability vector $p \in \mathcal{P}(\mathbb{Z}_n)$ such that

$$\Phi(X) = \sum_{a \in \mathbb{Z}_n} p(a) W_{0,a} X W_{0,a}^* \tag{4.263}$$

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

4.4. Let $\mathcal{X}$ be a complex Euclidean space and let $\Phi \in \mathrm{T}(\mathcal{X})$ be a Hermiticity-preserving map. Prove that the following two statements are equivalent:

1. Φ is positive, trace-preserving, and unital.
2. $\Phi(H) \prec H$ for every Hermitian operator $H \in \mathrm{Herm}(\mathcal{X})$.

4.5. Let $\mathcal{X}$ be a complex Euclidean space, let $\rho \in \mathrm{D}(\mathcal{X})$ be a density operator, let $p = (p_1, \ldots, p_m)$ be a probability vector, and assume $p_1 \geq p_2 \geq \cdots \geq p_m$. Prove that there exist unit vectors $u_1, \ldots, u_m \in \mathcal{X}$ satisfying

$$\rho = \sum_{k=1}^{m} p_k u_k u_k^* \tag{4.264}$$

if and only if

$$p_1 + \cdots + p_k \leq \lambda_1(\rho) + \cdots + \lambda_k(\rho) \tag{4.265}$$

for all k satisfying $1 \leq k \leq \mathrm{rank}(\rho)$.

A correct solution to this problem generalizes Theorem 4.35, as m need not coincide with the dimension of $\mathcal{X}$.

4.6. Let $\mathcal{X}$ be a complex Euclidean space, let $n = \dim(\mathcal{X})$, and let $\Phi \in \mathrm{C}(\mathcal{X})$ be a unital channel. Following the conventions discussed in Section 1.1.3 of Chapter 1, let $s_1(Y) \geq \cdots \geq s_n(Y)$ denote the singular values of a given operator $Y \in \mathrm{L}(\mathcal{X})$, ordered from largest to smallest, and taking $s_k(Y) = 0$ when $k > \mathrm{rank}(Y)$. Prove that, for every operator $X \in \mathrm{L}(\mathcal{X})$, it holds that

$$s_1(X) + \cdots + s_m(X) \geq s_1(\Phi(X)) + \cdots + s_m(\Phi(X)) \tag{4.266}$$

for every $m \in \{1, \ldots, n\}$.

4.5 Bibliographic remarks

Unital channels are sometimes referred to as *doubly stochastic* maps in the mathematics literature, although that term has also been used in reference to positive (but not necessarily completely positive), trace-preserving, and unital maps. The extreme points of sets of unital channels were studied by Landau and Streater [142]; the facts represented by Theorem 4.21, Example 4.3, and Theorem 4.23 appear in their paper. Related results for positive, trace-preserving, and unital maps had previously been discovered by Tregub [202], who also gave a different example of a unital channel (that also happens to be a Schur channel) that is not mixed-unitary. Another class of examples of this type appear in the work of Kümmerer and Maassen [140].

Mixed-unitary channels have often been called *random unitary* channels, as in the case of Audenaert and Scheel [18]. One disadvantage of this terminology is that it clashes with the terminology associated with a different concept, which is the consideration of unitary operators chosen at random from a given distribution (often the distribution associated with the Haar measure, as will be discussed later in Chapter 7).

The notion of environment-assisted channel correction was suggested by Alber, Beth, Charnes, Delgado, Grassl, and Mussinger [4]. The characterization of mixed-unitary channels based on this notion, as established by Theorem 4.8, follows from a slightly more general result due to Gregoratti and Werner [80]. Corollary 4.11 was proved by Buscemi [43] through the use of this characterization together with Corollary 2.48.

The discrete Weyl operators appear in Weyl's work on group-theoretic aspects of quantum mechanics. (See, for instance, Sections 14 and 15 in Chapter IV of [228].) The notion of covariance applies not only to the discrete Weyl operators and quantum channels, but to other collections of unitary operators and algebraic objects. There is, for example, some discussion of this notion in [228], and it was considered more explicitly for quantum instruments by Davies [54]. Channel covariance with respect to the discrete Weyl operators was considered by Holevo [109, 110], and the facts represented by Theorem 4.14 may be derived from that work.

Schur [190] proved that the positive semidefinite cone is closed under entry-wise products—which is a fact now referred to as the *Schur product theorem*. The entry-wise product of operators is called the *Schur product*, and Schur maps are so named for this reason. The term *Hadamard product* is

also sometimes used to refer to the entry-wise product, and correspondingly Schur maps are sometimes referred to as *Hadamard maps*. Schur maps are also referred to as *diagonal maps* by some authors, as they correspond to maps with diagonal Kraus operators (as is stated in Theorem 4.19).

Theorem 4.25 is due to Kribs [138], whose proof made use of arguments that can be found in the paper of Lindblad [149]. Fixed points of quantum channels, unital channels, and other classes of completely positive maps have also been studied by other researchers, including Bratteli, Jorgensen, Kishimoto, and Werner [42], Arias, Gheondea, and Gutter [13], and others. Theorem 4.27 is a special case of a theorem due to Perez-García, Wolf, Petz, and Ruskai [172]. (The theorem holds for a more general class of norms, not just the spectral norm.)

The notion of majorization for real vectors was developed in the first half of the twentieth century by several mathematicians, including Hardy, Littlewood, Pólya, Schur, Rado, and Horn. Details on this history may be found in Marshall, Olkin, and Arnold [154]. The extension of this notion to Hermitian operators is due to Uhlmann [206, 207, 208], as is Theorem 4.33. (See also the book of Alberti and Uhlmann [6].) The two implications of Theorem 4.34 were proved by Schur [189] and Horn [113], respectively, and Theorem 4.35 is due to Nielsen [164].