7

Permutation invariance and unitarily invariant measures

This chapter introduces two notions—*permutation invariance* and *unitarily invariant measures*—having interesting applications in quantum information theory. A state of a collection of identical registers is said to be permutation invariant if it is unchanged under arbitrary permutations of the contents of the registers. Unitarily invariant measures are Borel measures, defined for sets of vectors or operators, that are unchanged by the action of all unitary operators acting on the underlying space. The two notions are distinct but nevertheless linked, with the interplay between them offering a useful tool for performing calculations in both settings.

7.1 Permutation-invariant vectors and operators

This section of the chapter discusses properties of permutation-invariant states of collections of identical registers. Somewhat more generally, one may consider permutation-invariant positive semidefinite operators, as well as permutation-invariant vectors.

It is to be assumed for the entirety of the section that an alphabet Σ and a positive integer $n \geq 2$ have been fixed, and that $\mathsf{X}_1, \ldots, \mathsf{X}_n$ is a sequence of registers, all sharing the same classical state set Σ. The assumption that the registers $\mathsf{X}_1, \ldots, \mathsf{X}_n$ share the same classical state set Σ allows one to identify the complex Euclidean spaces $\mathcal{X}_1, \ldots, \mathcal{X}_n$ associated with these registers with a single space $\mathcal{X} = \mathbb{C}^{\Sigma}$, and to write

$$\mathcal{X}^{\otimes n} = \mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n \tag{7.1}$$

for the sake of brevity.

Algebraic properties of states of the compound register $(\mathsf{X}_1, \ldots, \mathsf{X}_n)$ that relate to permutations and symmetries among the individual registers will be a primary focus of the section.

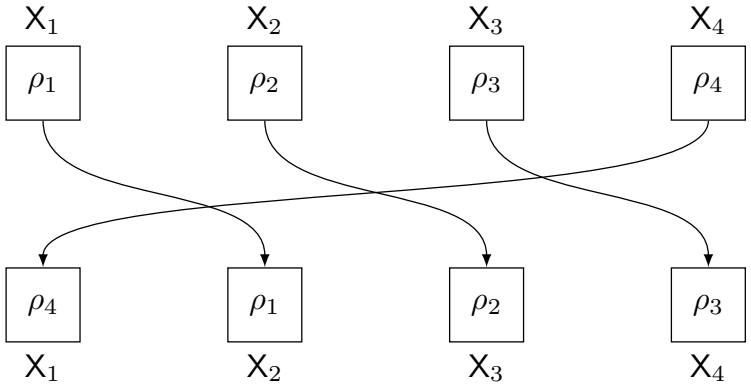

Figure 7.1 The action of the operator W_π on a register $(\mathsf{X}_1, \mathsf{X}_2, \mathsf{X}_3, \mathsf{X}_4)$ when $\pi = (1\ 2\ 3\ 4)$. If the register $(\mathsf{X}_1, \mathsf{X}_2, \mathsf{X}_3, \mathsf{X}_4)$ was initially in the product state $\rho = \rho_1 \otimes \rho_2 \otimes \rho_3 \otimes \rho_4$, and the contents of these registers were permuted according to π as illustrated, the resulting state would then be given by $W_\pi \rho W_\pi^* = \rho_4 \otimes \rho_1 \otimes \rho_2 \otimes \rho_3$. For non-product states, the action of W_π is determined by linearity.

7.1.1 The subspace of permutation-invariant vectors

Within the tensor product space

$$\mathcal{X}^{\otimes n} = \mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n, \tag{7.2}$$

some vectors are unchanged under all permutations of the tensor factors $\mathcal{X}_1, \ldots, \mathcal{X}_n$. The set of all such vectors forms a subspace that is known as the *symmetric subspace*. A more formal description of this subspace will be given shortly, following a short discussion of those operators that represent permutations among the tensor factors of the space (7.2).

Permutations of tensor factors

Define a unitary operator $W_\pi \in \mathrm{U}(\mathcal{X}^{\otimes n})$, for each permutation $\pi \in S_n$, by the action

$$W_\pi(x_1 \otimes \cdots \otimes x_n) = x_{\pi^{-1}(1)} \otimes \cdots \otimes x_{\pi^{-1}(n)} \tag{7.3}$$

for every choice of vectors $x_1, \ldots, x_n \in \mathcal{X}$. The action of the operator W_π, when considered as a channel acting on a state ρ as

$$\rho \mapsto W_\pi \rho W_\pi^*, \tag{7.4}$$

corresponds to permuting the contents of the registers $\mathsf{X}_1, \ldots, \mathsf{X}_n$ in the manner described by π. Figure 7.1 depicts an example of this action.

One may observe that

$$W_\pi W_\sigma = W_{\pi\sigma} \quad \text{and} \quad W_\pi^{-1} = W_\pi^* = W_{\pi^{-1}} \tag{7.5}$$

for all permutations $\pi, \sigma \in S_n$. Each operator W_π is a permutation operator, in the sense that it is a unitary operator with entries drawn from the set $\{0, 1\}$, and therefore one has

$$\overline{W_\pi} = W_\pi \quad \text{and} \quad W_\pi^{\mathsf{T}} = W_\pi^* \tag{7.6}$$

for every $\pi \in S_n$.

The symmetric subspace

As suggested above, some vectors in $\mathcal{X}^{\otimes n}$ are invariant under the action of W_π for every choice of $\pi \in S_n$, and it holds that the set of all such vectors forms a subspace known as the *symmetric subspace*. This subspace will be denoted $\mathcal{X}^{\circledcirc n}$, which is defined in more precise terms as

$$\mathcal{X}^{\circledcirc n} = \{x \in \mathcal{X}^{\otimes n} \,:\, x = W_\pi x \text{ for every } \pi \in S_n\}. \tag{7.7}$$

This space may alternatively be denoted $\mathcal{X}_1 \circledcirc \cdots \circledcirc \mathcal{X}_n$ when it is useful to do so. (The use of this notation naturally assumes that $\mathcal{X}_1, \ldots, \mathcal{X}_n$ have been identified with a single complex Euclidean space $\mathcal{X}$.)

The following proposition serves as a convenient starting point from which other facts regarding the symmetric subspace may be derived.

Proposition 7.1 *Let $\mathcal{X}$ be a complex Euclidean space and n a positive integer. The projection onto the symmetric subspace $\mathcal{X}^{\circledcirc n}$ is given by*

$$\Pi_{\mathcal{X}^{\circledcirc n}} = \frac{1}{n!} \sum_{\pi \in S_n} W_\pi. \tag{7.8}$$

Proof Using the equations (7.5), one may verify directly that the operator

$$\Pi = \frac{1}{n!} \sum_{\pi \in S_n} W_\pi \tag{7.9}$$

is Hermitian and squares to itself, implying that it is a projection operator. It holds that $W_\pi \Pi = \Pi$ for every $\pi \in S_n$, implying that

$$\operatorname{im}(\Pi) \subseteq \mathcal{X}^{\circledcirc n}. \tag{7.10}$$

On the other hand, for every $x \in \mathcal{X}^{\circledcirc n}$, it is evident that $\Pi x = x$, implying

$$\mathcal{X}^{\circledcirc n} \subseteq \operatorname{im}(\Pi). \tag{7.11}$$

As Π is a projection operator that satisfies $\operatorname{im}(\Pi) = \mathcal{X}^{\circledcirc n}$, the proposition is proved. □

An orthonormal basis for the symmetric subspace $\mathcal{X}^{\circledS n}$ will be identified next, and in the process the dimension of this space will be determined. It is helpful to make use of basic combinatorial concepts for this purpose.

First, for every alphabet Σ and every positive integer n, one defines the set $\mathrm{Bag}(n,\Sigma)$ to be the collection of all functions of the form $\phi : \Sigma \to \mathbb{N}$ (where $\mathbb{N} = \{0,1,2,\ldots\}$) possessing the property

$$\sum_{a\in\Sigma} \phi(a) = n. \tag{7.12}$$

Each function $\phi \in \mathrm{Bag}(n,\Sigma)$ may be viewed as describing a *bag* containing a total of n objects, each labeled by a symbol from the alphabet Σ. For each $a \in \Sigma$, the value $\phi(a)$ specifies the number of objects in the bag that are labeled by a. The objects are not considered to be ordered within the bag—it is only the number of objects having each possible label that is indicated by the function ϕ. Equivalently, a function $\phi \in \mathrm{Bag}(n,\Sigma)$ may be interpreted as a description of a multiset of size exactly n with elements drawn from Σ.

An n-tuple $(a_1,\ldots,a_n) \in \Sigma^n$ is *consistent* with a function $\phi \in \mathrm{Bag}(n,\Sigma)$ if and only if

$$\phi(a) = \big|\{k \in \{1,\ldots,n\} \,:\, a = a_k\}\big| \tag{7.13}$$

for every $a \in \Sigma$. In words, $(a_1,\ldots,a_n)$ is consistent with ϕ if and only if $(a_1,\ldots,a_n)$ represents one possible ordering of the elements in the multiset specified by ϕ. For each $\phi \in \mathrm{Bag}(n,\Sigma)$, the set Σ^n_ϕ is defined as the subset of Σ^n containing those elements $(a_1,\ldots,a_n) \in \Sigma^n$ that are consistent with ϕ. This yields a partition of Σ^n, as each n-tuple $(a_1,\ldots,a_n) \in \Sigma^n$ is consistent with precisely one function $\phi \in \mathrm{Bag}(n,\Sigma)$. For any two n-tuples

$$(a_1,\ldots,a_n),\ (b_1,\ldots,b_n) \in \Sigma^n_\phi \tag{7.14}$$

that are consistent with the same function $\phi \in \mathrm{Bag}(n,\Sigma)$, there must exist at least one permutation $\pi \in S_n$ for which

$$(a_1,\ldots,a_n) = \big(b_{\pi(1)},\ldots,b_{\pi(n)}\big). \tag{7.15}$$

The number of distinct functions $\phi \in \mathrm{Bag}(n,\Sigma)$ is given by the formula

$$|\mathrm{Bag}(n,\Sigma)| = \binom{|\Sigma|+n-1}{|\Sigma|-1}, \tag{7.16}$$

and for each $\phi \in \mathrm{Bag}(n,\Sigma)$ the number of distinct n-tuples within the subset Σ^n_ϕ is given by

$$|\Sigma^n_\phi| = \frac{n!}{\prod_{a\in\Sigma}(\phi(a)!)}. \tag{7.17}$$

As the following proposition establishes, an orthonormal basis for the symmetric subspace $\mathcal{X}^{\circledcirc n}$ may be obtained through the notions that were just introduced.

Proposition 7.2 *Let Σ be an alphabet, let n be a positive integer, and let $\mathcal{X} = \mathbb{C}^{\Sigma}$. Define a vector $u_\phi \in \mathcal{X}^{\otimes n}$ for each $\phi \in \mathrm{Bag}(n, \Sigma)$ as*

$$u_\phi = \left|\Sigma_\phi^n\right|^{-\frac{1}{2}} \sum_{(a_1,\ldots,a_n)\in\Sigma_\phi^n} e_{a_1} \otimes \cdots \otimes e_{a_n}. \tag{7.18}$$

The collection

$$\{u_\phi \,:\, \phi \in \mathrm{Bag}(n,\Sigma)\} \tag{7.19}$$

is an orthonormal basis for $\mathcal{X}^{\circledcirc n}$.

Proof It is evident that each vector u_ϕ is a unit vector. Moreover, for each choice of $\phi, \psi \in \mathrm{Bag}(n, \Sigma)$ with $\phi \neq \psi$, it holds that

$$\Sigma_\phi^n \cap \Sigma_\psi^n = \varnothing, \tag{7.20}$$

and therefore $\langle u_\phi, u_\psi \rangle = 0$, as each element $(a_1, \ldots, a_n) \in \Sigma^n$ is consistent with precisely one element of $\mathrm{Bag}(n, \Sigma)$. It therefore holds that (7.19) is an orthonormal set. As each vector u_ϕ is invariant under the action of W_π for every $\pi \in S_n$, it holds that

$$u_\phi \in \mathcal{X}^{\circledcirc n} \tag{7.21}$$

for every $\phi \in \mathrm{Bag}(n, \Sigma)$.

To complete the proof, it remains to prove that the set

$$\{u_\phi \,:\, \phi \in \mathrm{Bag}(n,\Sigma)\} \tag{7.22}$$

spans all of $\mathcal{X}^{\circledcirc n}$. This fact follows from the observation that, for every n-tuple $(a_1, \ldots, a_n) \in \Sigma^n$, it holds that

$$\begin{aligned} &\Pi_{\mathcal{X}^{\circledcirc n}}(e_{a_1} \otimes \cdots \otimes e_{a_n}) \\ &\quad = \frac{1}{n!} \sum_{\pi\in\mathcal{S}_n} W_\pi(e_{a_1} \otimes \cdots \otimes e_{a_n}) = \left|\Sigma_\phi^n\right|^{-\frac{1}{2}} u_\phi, \end{aligned} \tag{7.23}$$

for the unique element $\phi \in \mathrm{Bag}(n, \Sigma)$ with which the n-tuple $(a_1, \ldots, a_n)$ is consistent. □

Corollary 7.3 *Let $\mathcal{X}$ be a complex Euclidean space and let n be a positive integer. It holds that*

$$\dim(\mathcal{X}^{\circledvee n}) = \binom{\dim(\mathcal{X}) + n - 1}{\dim(\mathcal{X}) - 1} = \binom{\dim(\mathcal{X}) + n - 1}{n}. \tag{7.24}$$

Example 7.4 Suppose $\Sigma = \{0, 1\}$, $\mathcal{X} = \mathbb{C}^{\Sigma}$, and $n = 3$. The following four vectors form an orthonormal basis of $\mathcal{X}^{\circledvee 3}$:

$$\begin{aligned}
u_0 &= e_0 \otimes e_0 \otimes e_0 \\
u_1 &= \frac{1}{\sqrt{3}}(e_0 \otimes e_0 \otimes e_1 + e_0 \otimes e_1 \otimes e_0 + e_1 \otimes e_0 \otimes e_0) \\
u_2 &= \frac{1}{\sqrt{3}}(e_0 \otimes e_1 \otimes e_1 + e_1 \otimes e_0 \otimes e_1 + e_1 \otimes e_1 \otimes e_0) \\
u_3 &= e_1 \otimes e_1 \otimes e_1.
\end{aligned} \tag{7.25}$$

Tensor power spanning sets for the symmetric subspace

It is evident that the inclusion

$$v^{\otimes n} \in \mathcal{X}^{\circledvee n} \tag{7.26}$$

holds for every vector $v \in \mathcal{X}$. The following theorem demonstrates that the symmetric subspace $\mathcal{X}^{\circledvee n}$ is, in fact, spanned by the set of all vectors having this form. This fact remains true when the entries of v are restricted to finite subsets of $\mathbb{C}$, provided that those sets are sufficiently large.

Theorem 7.5 *Let Σ be an alphabet, let n be a positive integer, and let $\mathcal{X} = \mathbb{C}^{\Sigma}$. For any set $\mathcal{A} \subseteq \mathbb{C}$ satisfying $|\mathcal{A}| \geq n + 1$ it holds that*

$$\operatorname{span}\Big\{v^{\otimes n} \,:\, v \in \mathcal{A}^{\Sigma}\Big\} = \mathcal{X}^{\circledvee n}. \tag{7.27}$$

Theorem 7.5 can be proved in multiple ways. One proof makes use of the following elementary fact concerning multivariate polynomials.

Lemma 7.6 (Schwartz–Zippel) *Let P be a multivariate polynomial, with variables $Z_1, \ldots, Z_m$ and complex number coefficients, that is not identically zero and has total degree at most n, and let $\mathcal{A} \subset \mathbb{C}$ be a nonempty, finite set of complex numbers. It holds that*

$$\big|\{(\alpha_1, \ldots, \alpha_m) \in \mathcal{A}^m \,:\, P(\alpha_1, \ldots, \alpha_m) = 0\}\big| \leq n|\mathcal{A}|^{m-1}. \tag{7.28}$$

Proof The lemma is trivial in the case that $|\mathcal{A}| \leq n$, so it will be assumed that $|\mathcal{A}| \geq n + 1$ for the remainder of the proof, which is by induction on m. When $m = 1$, the lemma follows from the fact that a nonzero, univariate polynomial with degree at most n can have at most n roots.

Under the assumption that $m \geq 2$, one may write

$$P(Z_1, \ldots, Z_m) = \sum_{k=0}^{n} Q_k(Z_1, \ldots, Z_{m-1}) Z_m^k, \tag{7.29}$$

for $Q_0, \ldots, Q_n$ being complex polynomials in variables $Z_1, \ldots, Z_{m-1}$, and with the total degree of Q_k being at most $n - k$ for each $k \in \{0, \ldots, n\}$. Fix k to be the largest value in the set $\{0, \ldots, n\}$ for which Q_k is nonzero. Given that P is nonzero, there must exist such a choice of k.

As Q_k has total degree at most $n - k$, it follows from the hypothesis of induction that

$$\begin{aligned} \big|\{(\alpha_1, \ldots, \alpha_{m-1}) \in \mathcal{A}^{m-1} \,:\, Q_k(\alpha_1, \ldots, \alpha_{m-1}) \neq 0\}\big| \\ \geq |\mathcal{A}|^{m-1} - (n-k)|\mathcal{A}|^{m-2}. \end{aligned} \tag{7.30}$$

For each choice of $(\alpha_1, \ldots, \alpha_{m-1}) \in \mathcal{A}^{m-1}$ for which $Q_k(\alpha_1, \ldots, \alpha_{m-1}) \neq 0$, it holds that

$$P(\alpha_1, \ldots, \alpha_{m-1}, Z_m) = \sum_{j=0}^{k} Q_j(\alpha_1, \ldots, \alpha_{m-1}) Z_m^j \tag{7.31}$$

is a univariate polynomial of degree k in the variable Z_m, implying that there must exist at least $|\mathcal{A}| - k$ choices of $\alpha_m \in \mathcal{A}$ for which

$$P(\alpha_1, \ldots, \alpha_m) \neq 0. \tag{7.32}$$

It follows that there are at least

$$\big(|\mathcal{A}|^{m-1} - (n-k)|\mathcal{A}|^{m-2}\big)\big(|\mathcal{A}| - k\big) \geq |\mathcal{A}|^m - n|\mathcal{A}|^{m-1} \tag{7.33}$$

distinct m-tuples $(\alpha_1, \ldots, \alpha_m) \in \mathcal{A}^m$ for which $P(\alpha_1, \ldots, \alpha_m) \neq 0$, which completes the proof of the lemma. □

Remark Although it is irrelevant to its use in proving Theorem 7.5, one may observe that Lemma 7.6 holds for P being a multivariate polynomial over any field, not just the field of complex numbers. This fact is established by the proof above, which has not used properties of the complex numbers that do not hold for arbitrary fields.

Proof of Theorem 7.5 For every choice of a permutation $\pi \in S_n$ and a vector $v \in \mathbb{C}^\Sigma$, it holds that

$$W_\pi v^{\otimes n} = v^{\otimes n}. \tag{7.34}$$

It follows that $v^{\otimes n} \in \mathcal{X}^{\circledcirc n}$, and therefore

$$\operatorname{span}\Big\{v^{\otimes n} \,:\, v \in \mathcal{A}^\Sigma\Big\} \subseteq \mathcal{X}^{\circledcirc n}. \tag{7.35}$$

To prove the reverse inclusion, let $w \in \mathcal{X}^{\circledcirc n}$ be any nonzero vector, and write

$$w = \sum_{\phi \in \mathrm{Bag}(n,\Sigma)} \alpha_\phi u_\phi, \tag{7.36}$$

for some collection of complex number coefficients $\{\alpha_\phi : \phi \in \mathrm{Bag}(n,\Sigma)\}$, with each vector u_ϕ being defined as in (7.18). It will be proved that

$$\langle w, v^{\otimes n} \rangle \neq 0 \tag{7.37}$$

for at least one choice of a vector $v \in \mathcal{A}^\Sigma$. The required inclusion follows from this fact, for if the containment (7.35) were proper, it would be possible to choose $w \in \mathcal{X}^{\circledcirc n}$ that is orthogonal to $v^{\otimes n}$ for every $v \in \mathcal{A}^\Sigma$.

For the remainder of the proof it will be assumed that $\mathcal{A}$ is a finite set, which causes no loss of generality, for if $\mathcal{A}$ were infinite, one could restrict their attention to an arbitrary finite subset of $\mathcal{A}$ having size at least $n+1$, yielding the desired inclusion.

Define a multivariate polynomial

$$Q = \sum_{\phi \in \mathrm{Bag}(n,\Sigma)} \overline{\alpha_\phi} \sqrt{|\Sigma_\phi^n|} \prod_{a \in \Sigma} Z_a^{\phi(a)} \tag{7.38}$$

in a collection of variables $\{Z_a : a \in \Sigma\}$. As the monomials

$$\prod_{a \in \Sigma} Z_a^{\phi(a)} \tag{7.39}$$

are distinct as ϕ ranges over the elements of $\mathrm{Bag}(n,\Sigma)$, with each monomial having total degree n, it follows that Q is a nonzero polynomial with total degree n. A calculation reveals that

$$Q(v) = \langle w, v^{\otimes n} \rangle \tag{7.40}$$

for every vector $v \in \mathbb{C}^\Sigma$, where $Q(v)$ refers to the complex number obtained by the substitution of the value $v(a)$ for the variable Z_a in Q for each $a \in \Sigma$. As Q is a nonzero multivariate polynomial with total degree n, it follows from the Schwartz–Zippel lemma (Lemma 7.6) that $Q(v) = 0$ for at most

$$n|\mathcal{A}|^{|\Sigma|-1} < |\mathcal{A}|^{|\Sigma|} \tag{7.41}$$

choices of vectors $v \in \mathcal{A}^\Sigma$, implying that there exists at least one vector $v \in \mathcal{A}^\Sigma$ for which $\langle w, v^{\otimes n} \rangle \neq 0$, completing the proof. □

The anti-symmetric subspace

Along similar lines to the symmetric subspace $\mathcal{X}^{\circledcirc n}$ of the tensor product space $\mathcal{X}^{\otimes n}$, one may define the *anti-symmetric subspace* of the same tensor product space as

$$\mathcal{X}^{\circledcirc n} = \{x \in \mathcal{X}^{\otimes n} \,:\, W_\pi x = \operatorname{sign}(\pi) x \text{ for every } \pi \in S_n\}. \tag{7.42}$$

The short discussion on the anti-symmetric subspace that follows may, for the most part, be considered as an aside; with the exception of the case in which $n = 2$, the anti-symmetric subspace does not play a significant role elsewhere in this book. It is, nevertheless, natural to consider this subspace along side of the symmetric subspace. The following propositions establish a few basic facts about the anti-symmetric subspace.

Proposition 7.7 *Let $\mathcal{X}$ be a complex Euclidean space and n a positive integer. The projection onto the anti-symmetric subspace $\mathcal{X}^{\circledcirc n}$ is given by*

$$\Pi_{\mathcal{X}^{\circledcirc n}} = \frac{1}{n!} \sum_{\pi \in S_n} \operatorname{sign}(\pi) W_\pi. \tag{7.43}$$

Proof The proof is similar to the proof of Proposition 7.1. Using (7.5), along with the fact that $\operatorname{sign}(\pi)\operatorname{sign}(\sigma) = \operatorname{sign}(\pi\sigma)$ for every choice of $\pi, \sigma \in S_n$, it may be verified that the operator

$$\Pi = \frac{1}{n!} \sum_{\pi \in S_n} \operatorname{sign}(\pi) W_\pi \tag{7.44}$$

is Hermitian and squares to itself, implying that it is a projection operator. For every $\pi \in S_n$ it holds that

$$W_\pi \Pi = \operatorname{sign}(\pi) \Pi, \tag{7.45}$$

from which it follows that

$$\operatorname{im}(\Pi) \subseteq \mathcal{X}^{\circledcirc n}. \tag{7.46}$$

For every vector $x \in \mathcal{X}^{\circledcirc n}$, it holds that $\Pi x = x$, implying that

$$\mathcal{X}^{\circledcirc n} \subseteq \operatorname{im}(\Pi). \tag{7.47}$$

As Π is a projection operator satisfying $\operatorname{im}(\Pi) = \mathcal{X}^{\circledcirc n}$, the proposition is proved. □

When constructing an orthonormal basis of the anti-symmetric subspace $\mathcal{X}^{\circledcirc n}$, for $\mathcal{X} = \mathbb{C}^{\Sigma}$, it is convenient to assume that a total ordering of Σ has been fixed. For every n-tuple $(a_1, \ldots, a_n) \in \Sigma^n$ for which $a_1 < \cdots < a_n$, define a vector

$$u_{a_1,\ldots,a_n} = \frac{1}{\sqrt{n!}} \sum_{\pi \in S_n} \operatorname{sign}(\pi) W_\pi(e_{a_1} \otimes \cdots \otimes e_{a_n}). \tag{7.48}$$

Proposition 7.8 *Let Σ be an alphabet, let $n \geq 2$ be a positive integer, let $\mathcal{X} = \mathbb{C}^{\Sigma}$, and define $u_{a_1,\ldots,a_n} \in \mathcal{X}^{\otimes n}$ for each n-tuple $(a_1, \ldots, a_n) \in \Sigma^n$ satisfying $a_1 < \cdots < a_n$ as in* (7.48)*. The collection*

$$\{u_{a_1,\ldots,a_n} \,:\, (a_1, \ldots, a_n) \in \Sigma^n,\ a_1 < \cdots < a_n\} \tag{7.49}$$

is an orthonormal basis for $\mathcal{X}^{\circledcirc n}$.

Proof Each vector $u_{a_1,\ldots,a_n}$ is evidently a unit vector, and is contained in the space $\mathcal{X}^{\circledcirc n}$. For distinct n-tuples $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ with $a_1 < \cdots < a_n$ and $b_1 < \cdots < b_n$ it holds that

$$\langle u_{a_1,\ldots,a_n}, u_{b_1,\ldots,b_n} \rangle = 0, \tag{7.50}$$

as these vectors are linear combinations of disjoint sets of standard basis vectors. It therefore remains to prove that the collection (7.49) spans $\mathcal{X}^{\circledcirc n}$.

For any choice of distinct indices $j, k \in \{1, \ldots, n\}$, and for $(j\ k) \in S_n$ being the permutation that swaps j and k, leaving all other elements of $\{1, \ldots, n\}$ fixed, one has

$$W_{(j\,k)} \Pi_{\mathcal{X}^{\circledcirc n}} = -\Pi_{\mathcal{X}^{\circledcirc n}} = \Pi_{\mathcal{X}^{\circledcirc n}} W_{(j\,k)}. \tag{7.51}$$

Consequently, for any choice of an n-tuple $(a_1, \ldots, a_n) \in \Sigma^n$ for which there exist distinct indices $j, k \in \{1, \ldots, n\}$ for which $a_j = a_k$, it holds that

$$\begin{aligned} \Pi_{\mathcal{X}^{\circledcirc n}}(e_{a_1} \otimes \cdots \otimes e_{a_n}) &= \Pi_{\mathcal{X}^{\circledcirc n}} W_{(j\,k)}(e_{a_1} \otimes \cdots \otimes e_{a_n}) \\ &= -\Pi_{\mathcal{X}^{\circledcirc n}}(e_{a_1} \otimes \cdots \otimes e_{a_n}), \end{aligned} \tag{7.52}$$

and therefore

$$\Pi_{\mathcal{X}^{\circledcirc n}}(e_{a_1} \otimes \cdots \otimes e_{a_n}) = 0. \tag{7.53}$$

On the other hand, if $(a_1, \ldots, a_n) \in \Sigma^n$ is an n-tuple for which $a_1, \ldots, a_n$ are distinct elements of Σ, it must hold that

$$(a_{\pi(1)}, \ldots, a_{\pi(n)}) = (b_1, \ldots, b_n) \tag{7.54}$$

for some choice of a permutation $\pi \in S_n$ and an n-tuple $(b_1, \ldots, b_n) \in \Sigma^n$

satisfying $b_1 < \cdots < b_n$. One therefore has

$$\begin{aligned} \Pi_{\mathcal{X}^{\circledcirc n}}(e_{a_1} \otimes \cdots \otimes e_{a_n}) &= \Pi_{\mathcal{X}^{\circledcirc n}} W_\pi (e_{b_1} \otimes \cdots \otimes e_{b_n}) \\ &= \operatorname{sign}(\pi) \Pi_{\mathcal{X}^{\circledcirc n}}(e_{b_1} \otimes \cdots \otimes e_{b_n}) = \frac{\operatorname{sign}(\pi)}{\sqrt{n!}} u_{b_1,\ldots,b_n}. \end{aligned} \tag{7.55}$$

It therefore holds that

$$\operatorname{im}(\Pi_{\mathcal{X}^{\circledcirc n}}) \subseteq \operatorname{span}\{u_{a_1,\ldots,a_n} \,:\, (a_1,\ldots,a_n) \in \Sigma^n,\ a_1 < \cdots < a_n\}, \tag{7.56}$$

which completes the proof. □

By the previous proposition, one has that the dimension of the anti-symmetric subspace is equal to the number of n-tuples $(a_1,\ldots,a_n) \in \Sigma^n$ satisfying $a_1 < \cdots < a_n$. This number is equal to the number of subsets of Σ having n elements.

Corollary 7.9 *Let $\mathcal{X}$ be a complex Euclidean space and let n be a positive integer. It holds that*

$$\dim(\mathcal{X}^{\circledcirc n}) = \binom{\dim(\mathcal{X})}{n}. \tag{7.57}$$

7.1.2 The algebra of permutation-invariant operators

By its definition, the symmetric subspace $\mathcal{X}^{\circledcirc n}$ includes all vectors $x \in \mathcal{X}^{\otimes n}$ that are invariant under the action of W_π for each $\pi \in S_n$. One may consider a similar notion for operators, with the action $x \mapsto W_\pi x$ being replaced by the action

$$X \mapsto W_\pi X W_\pi^* \tag{7.58}$$

for each $X \in \mathrm{L}(\mathcal{X}^{\otimes n})$. The notation $\mathrm{L}(\mathcal{X})^{\circledcirc n}$ will be used to denote the set of operators X that are invariant under this action:

$$\mathrm{L}(\mathcal{X})^{\circledcirc n} = \{X \in \mathrm{L}(\mathcal{X}^{\otimes n}) \,:\, X = W_\pi X W_\pi^* \text{ for all } \pi \in S_n\}. \tag{7.59}$$

Similar to the analogous notion for vectors, one may denote this set as $\mathrm{L}(\mathcal{X}_1) \circledcirc \cdots \circledcirc \mathrm{L}(\mathcal{X}_n)$ when it is convenient to do this, under the assumption that the spaces $\mathcal{X}_1,\ldots,\mathcal{X}_n$ have been identified with a single space $\mathcal{X}$.

Assuming that $\mathsf{X}_1,\ldots,\mathsf{X}_n$ are registers sharing the same classical state set Σ, and identifying each of the spaces $\mathcal{X}_1,\ldots,\mathcal{X}_n$ with $\mathcal{X} = \mathbb{C}^\Sigma$, one observes that the density operator elements of the set $\mathrm{L}(\mathcal{X})^{\circledcirc n}$ represent states of the compound register $(\mathsf{X}_1,\ldots,\mathsf{X}_n)$ that are invariant under all permutations of the registers $\mathsf{X}_1,\ldots,\mathsf{X}_n$. Such states are said to be *exchangeable*.

Algebraic properties of the set $\mathrm{L}(\mathcal{X})^{\circledcirc n}$, along with a relationship between exchangeable states and permutation-invariant vectors, are described in the subsections that follow.

Vector space structure of the permutation-invariant operators

The notation $\mathrm{L}(\mathcal{X})^{\circledcirc n}$ is a natural choice for the space of all permutation-invariant operators; if one regards $\mathrm{L}(\mathcal{X})$ as a vector space, then $\mathrm{L}(\mathcal{X})^{\circledcirc n}$ indeed coincides with the symmetric subspace of the tensor product space $\mathrm{L}(\mathcal{X})^{\otimes n}$. The next proposition formalizes this connection and states some immediate consequences of the results of the previous section.

Proposition 7.10 *Let $\mathcal{X}$ be a complex Euclidean space, let n be a positive integer, and let $X \in \mathrm{L}(\mathcal{X}^{\otimes n})$. The following statements are equivalent:*

1. $X \in \mathrm{L}(\mathcal{X})^{\circledcirc n}$.
2. *For $V \in \mathrm{U}(\mathcal{X}^{\otimes n} \otimes \mathcal{X}^{\otimes n}, (\mathcal{X} \otimes \mathcal{X})^{\otimes n})$ being the isometry defined by the equation*

$$V \operatorname{vec}(Y_1 \otimes \cdots \otimes Y_n) = \operatorname{vec}(Y_1) \otimes \cdots \otimes \operatorname{vec}(Y_n) \tag{7.60}$$

   holding for all $Y_1, \ldots, Y_n \in \mathrm{L}(\mathcal{X})$, one has that

$$V \operatorname{vec}(X) \in (\mathcal{X} \otimes \mathcal{X})^{\circledcirc n}. \tag{7.61}$$

3. $X \in \operatorname{span}\{Y^{\otimes n} : Y \in \mathrm{L}(\mathcal{X})\}$.

Proof For each permutation $\pi \in S_n$, let

$$U_\pi \in \mathrm{U}\big((\mathcal{X} \otimes \mathcal{X})^{\otimes n}\big) \tag{7.62}$$

be the unitary operator defined by the equation

$$U_\pi(w_1 \otimes \cdots \otimes w_n) = w_{\pi^{-1}(1)} \otimes \cdots \otimes w_{\pi^{-1}(n)} \tag{7.63}$$

holding for all vectors $w_1, \ldots, w_n \in \mathcal{X} \otimes \mathcal{X}$. Each operator U_π is analogous to W_π, as defined in (7.3), but with the space $\mathcal{X}$ replaced by $\mathcal{X} \otimes \mathcal{X}$. It holds that

$$U_\pi = V(W_\pi \otimes W_\pi)V^* \tag{7.64}$$

for every $\pi \in S_n$, from which one may conclude that the first and second statements are equivalent.

Theorem 7.5 implies that

$$V \operatorname{vec}(X) \in (\mathcal{X} \otimes \mathcal{X})^{\circledcirc n} \tag{7.65}$$

if and only if

$$V\operatorname{vec}(X) \in \operatorname{span}\bigl\{\operatorname{vec}(Y)^{\otimes n} \,:\, Y \in \mathrm{L}(\mathcal{X})\bigr\}. \tag{7.66}$$

The containment (7.66) is equivalent to

$$\operatorname{vec}(X) \in \operatorname{span}\bigl\{\operatorname{vec}\bigl(Y^{\otimes n}\bigr) \,:\, Y \in \mathrm{L}(\mathcal{X})\bigr\}, \tag{7.67}$$

which in turn is equivalent to

$$X \in \operatorname{span}\bigl\{Y^{\otimes n} \,:\, Y \in \mathrm{L}(\mathcal{X})\bigr\}. \tag{7.68}$$

The second and third statements are therefore equivalent. □

Theorem 7.11 *Let $\mathcal{X}$ be a complex Euclidean space and let n be a positive integer. It holds that*

$$\mathrm{L}(\mathcal{X})^{\circledcirc n} = \operatorname{span}\bigl\{U^{\otimes n} \,:\, U \in \mathrm{U}(\mathcal{X})\bigr\}. \tag{7.69}$$

Proof Let Σ be the alphabet for which $\mathcal{X} = \mathbb{C}^{\Sigma}$, and let

$$D = \operatorname{Diag}(u) \tag{7.70}$$

be a diagonal operator, for an arbitrary choice of $u \in \mathcal{X}$. It holds that $u^{\otimes n} \in \mathcal{X}^{\circledcirc n}$, so by Theorem 7.5 one has that

$$u^{\otimes n} \in \operatorname{span}\bigl\{v^{\otimes n} \,:\, v \in \mathbb{T}^{\Sigma}\bigr\}, \tag{7.71}$$

for $\mathbb{T} = \bigl\{\alpha \in \mathbb{C} \,:\, |\alpha| = 1\bigr\}$ denoting the set of complex units. It is therefore possible to write

$$u^{\otimes n} = \sum_{b \in \Gamma} \beta_b v_b^{\otimes n} \tag{7.72}$$

for some choice of an alphabet Γ, vectors $\{v_b \,:\, b \in \Gamma\} \subset \mathbb{T}^{\Sigma}$, and complex numbers $\{\beta_b \,:\, b \in \Gamma\} \subset \mathbb{C}$. It follows that

$$D^{\otimes n} = \sum_{b \in \Gamma} \beta_b U_b^{\otimes n} \tag{7.73}$$

for $U_b \in \mathrm{U}(\mathcal{X})$ being the unitary operator defined as

$$U_b = \operatorname{Diag}(v_b) \tag{7.74}$$

for each $b \in \Gamma$.

Now, for an arbitrary operator $A \in \mathrm{L}(\mathcal{X})$, one may write $A = VDW$ for $V, W \in \mathrm{U}(\mathcal{X})$ being unitary operators and $D \in \mathrm{L}(\mathcal{X})$ being a diagonal operator, by Corollary 1.7 (to the singular value theorem). Invoking the argument above, one may assume that (7.73) holds, and therefore

$$A^{\otimes n} = \sum_{b \in \Gamma} \beta_b (V U_b W)^{\otimes n}, \tag{7.75}$$

for some choice of an alphabet Γ, complex numbers $\{\beta_b : b \in \Gamma\} \subset \mathbb{C}$, and diagonal unitary operators $\{U_b : b \in \Gamma\}$. As VU_bW is unitary for each $b \in \Gamma$, one has

$$A^{\otimes n} \in \operatorname{span}\{U^{\otimes n} : U \in \mathrm{U}(\mathcal{X})\}, \tag{7.76}$$

so by Proposition 7.10 it follows that

$$\mathrm{L}(\mathcal{X})^{\circledcirc n} \subseteq \operatorname{span}\{U^{\otimes n} : U \in \mathrm{U}(\mathcal{X})\}. \tag{7.77}$$

The reverse containment is immediate, so the theorem is proved. □

Symmetric purifications of exchangeable density operators

A density operator $\rho \in \mathrm{D}(\mathcal{X}^{\otimes n})$ is exchangeable if and only if $\rho \in \mathrm{L}(\mathcal{X})^{\circledcirc n}$, which is equivalent to

$$\rho = W_\pi \rho W_\pi^* \tag{7.78}$$

for every permutation $\pi \in S_n$. In operational terms, an exchangeable state ρ of a compound register $(\mathsf{X}_1, \ldots, \mathsf{X}_n)$, for n identical registers $\mathsf{X}_1, \ldots, \mathsf{X}_n$, is one that does not change if the contents of these n registers are permuted in an arbitrary way.

For every symmetric unit vector $u \in \mathcal{X}^{\circledcirc n}$, one has that the pure state uu^* is exchangeable, and naturally any convex combination of such states must be exchangeable as well. In general, this does not exhaust all possible exchangeable states. For instance, the completely mixed state in $\mathrm{D}(\mathcal{X}^{\otimes n})$ is exchangeable, but the image of the density operator corresponding to this state is generally not contained within the symmetric subspace.

There is, nevertheless, an interesting relationship between exchangeable states and symmetric pure states, which is that every exchangeable state can be purified in such a way that its purification lies within a larger symmetric subspace, in the sense described by the following theorem.

Theorem 7.12 *Let Σ and Γ be alphabets with $|\Gamma| \geq |\Sigma|$ and let n be a positive integer. Also let $\mathsf{X}_1, \ldots, \mathsf{X}_n$ be registers, each having classical state set Σ, let $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$ be registers, each having classical state set Γ, and let $\rho \in \mathrm{D}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n)$ be an exchangeable density operator. There exists a unit vector*

$$u \in (\mathcal{X}_1 \otimes \mathcal{Y}_1) \circledcirc \cdots \circledcirc (\mathcal{X}_n \otimes \mathcal{Y}_n) \tag{7.79}$$

such that

$$(uu^*)[\mathsf{X}_1, \ldots, \mathsf{X}_n] = \rho. \tag{7.80}$$

Proof Let $A \in \mathrm{U}(\mathbb{C}^{\Sigma}, \mathbb{C}^{\Gamma})$ be an arbitrarily chosen isometry, which one may regard as an element of $\mathrm{U}(\mathcal{X}_k, \mathcal{Y}_k)$ for any choice of $k \in \{1, \ldots, n\}$. Also let

$$\begin{aligned} V \in \mathrm{U}\big((\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n) \otimes (\mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n), \\ (\mathcal{X}_1 \otimes \mathcal{Y}_1) \otimes \cdots \otimes (\mathcal{X}_n \otimes \mathcal{Y}_n)\big) \end{aligned} \tag{7.81}$$

be the isometry defined by the equation

$$V \operatorname{vec}(B_1 \otimes \cdots \otimes B_n) = \operatorname{vec}(B_1) \otimes \cdots \otimes \operatorname{vec}(B_n), \tag{7.82}$$

holding for all choices of $B_1 \in \mathrm{L}(\mathcal{Y}_1, \mathcal{X}_1)$, …, $B_n \in \mathrm{L}(\mathcal{Y}_n, \mathcal{X}_n)$. Equivalently, this isometry is defined by the equation

$$\begin{aligned} V((x_1 \otimes \cdots \otimes x_n) \otimes (y_1 \otimes \cdots \otimes y_n)) \\ = (x_1 \otimes y_1) \otimes \cdots \otimes (x_n \otimes y_n), \end{aligned} \tag{7.83}$$

holding for all vectors $x_1 \in \mathcal{X}_1, \ldots, x_n \in \mathcal{X}_n$ and $y_1 \in \mathcal{Y}_1, \ldots, y_n \in \mathcal{Y}_n$.

Consider the vector

$$u = V \operatorname{vec}\big(\sqrt{\rho}(A^* \otimes \cdots \otimes A^*)\big) \in (\mathcal{X}_1 \otimes \mathcal{Y}_1) \otimes \cdots \otimes (\mathcal{X}_n \otimes \mathcal{Y}_n). \tag{7.84}$$

A calculation reveals that

$$(uu^*)[\mathsf{X}_1, \ldots, \mathsf{X}_n] = \rho, \tag{7.85}$$

and so it remains to prove that u is symmetric. Because ρ is exchangeable, one has

$$\big(W_\pi \sqrt{\rho} W_\pi^*\big)^2 = W_\pi \rho W_\pi^* = \rho \tag{7.86}$$

for every permutation $\pi \in \mathcal{S}_n$, and therefore

$$W_\pi \sqrt{\rho} W_\pi^* = \sqrt{\rho} \tag{7.87}$$

by the uniqueness of the square root. By Proposition 7.10, it therefore holds that

$$\sqrt{\rho} \in \operatorname{span}\{Y^{\otimes n} \,:\, Y \in \mathrm{L}(\mathbb{C}^{\Sigma})\}. \tag{7.88}$$

Consequently, one has

$$u \in \operatorname{span}\Big\{V \operatorname{vec}\Big(\big(YA^*\big)^{\otimes n}\Big) \,:\, Y \in \mathrm{L}(\mathbb{C}^{\Sigma})\Big\}, \tag{7.89}$$

and therefore

$$u \in \operatorname{span}\Big\{\operatorname{vec}\big(YA^*\big)^{\otimes n} \,:\, Y \in \mathrm{L}(\mathbb{C}^{\Sigma})\Big\}. \tag{7.90}$$

From this containment it is evident that

$$u \in (\mathcal{X}_1 \otimes \mathcal{Y}_1) \circledcirc \cdots \circledcirc (\mathcal{X}_n \otimes \mathcal{Y}_n), \tag{7.91}$$

which completes the proof. □

Von Neumann's double commutant theorem

To establish further properties of the set $\mathrm{L}(\mathcal{X})^{\otimes n}$, particularly ones relating to the operator structure of its elements, it is convenient to make use of a theorem known as *von Neumann's double commutant theorem*. This theorem is stated below, and its proof will make use of the following lemma.

Lemma 7.13 *Let $\mathcal{X}$ be a complex Euclidean space, let $\mathcal{V} \subseteq \mathcal{X}$ be a subspace of $\mathcal{X}$, and let $A \in \mathrm{L}(\mathcal{X})$ be an operator. The following two statements are equivalent:*

1. *It holds that both $A\mathcal{V} \subseteq \mathcal{V}$ and $A^*\mathcal{V} \subseteq \mathcal{V}$.*
2. *It holds that $[A, \Pi_{\mathcal{V}}] = 0$.*

Proof Assume first that statement 2 holds. If two operators commute, then their adjoints must also commute, and so one has the following for every vector $v \in \mathcal{V}$:

$$\begin{aligned} Av &= A\Pi_{\mathcal{V}}v = \Pi_{\mathcal{V}}Av \in \mathcal{V}, \\ A^*v &= A^*\Pi_{\mathcal{V}}v = \Pi_{\mathcal{V}}A^*v \in \mathcal{V}. \end{aligned} \tag{7.92}$$

It has been proved that statement 2 implies statement 1.

Now assume statement 1 holds. For every $v \in \mathcal{V}$, one has

$$\Pi_{\mathcal{V}}Av = Av = A\Pi_{\mathcal{V}}v, \tag{7.93}$$

by virtue of the fact that $Av \in \mathcal{V}$. For every $w \in \mathcal{X}$ with $w \perp \mathcal{V}$, it must hold that

$$\langle v, Aw \rangle = \langle A^*v, w \rangle = 0 \tag{7.94}$$

for every $v \in \mathcal{V}$, following from the assumption $A^*v \in \mathcal{V}$, and therefore $Aw \perp \mathcal{V}$. Consequently,

$$\Pi_{\mathcal{V}}Aw = 0 = A\Pi_{\mathcal{V}}w. \tag{7.95}$$

As every vector $u \in \mathcal{X}$ may be written as $u = v + w$ for some choice of $v \in \mathcal{V}$ and $w \in \mathcal{X}$ with $w \perp \mathcal{V}$, equations (7.93) and (7.95) imply

$$\Pi_{\mathcal{V}}Au = A\Pi_{\mathcal{V}}u \tag{7.96}$$

for every vector $u \in \mathcal{X}$, and therefore $\Pi_{\mathcal{V}}A = A\Pi_{\mathcal{V}}$. It has been proved that statement 1 implies statement 2, which completes the proof. □

Theorem 7.14 (Von Neumann's double commutant theorem) *Let $\mathcal{A}$ be a self-adjoint, unital subalgebra of $\mathrm{L}(\mathcal{X})$, for $\mathcal{X}$ being a complex Euclidean space. It holds that*

$$\mathrm{comm}(\mathrm{comm}(\mathcal{A})) = \mathcal{A}. \tag{7.97}$$

Proof It is immediate from the definition of the commutant that

$$\mathcal{A} \subseteq \operatorname{comm}(\operatorname{comm}(\mathcal{A})), \tag{7.98}$$

and so it remains to prove the reverse inclusion.

The key idea of the proof will be to consider the algebra $\mathrm{L}(\mathcal{X} \otimes \mathcal{X})$, and to make use of its relationships with $\mathrm{L}(\mathcal{X})$. Define $\mathcal{B} \subseteq \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$ as

$$\mathcal{B} = \{X \otimes \mathbb{1} \,:\, X \in \mathcal{A}\}, \tag{7.99}$$

and let Σ be the alphabet for which $\mathcal{X} = \mathbb{C}^{\Sigma}$. Every operator $Y \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$ may be written as

$$Y = \sum_{a,b \in \Sigma} Y_{a,b} \otimes E_{a,b} \tag{7.100}$$

for a unique choice of operators $\{Y_{a,b} \,:\, a, b \in \Sigma\} \subset \mathrm{L}(\mathcal{X})$. The condition

$$Y(X \otimes \mathbb{1}) = (X \otimes \mathbb{1})Y, \tag{7.101}$$

for any operator $X \in \mathrm{L}(\mathcal{X})$ and any operator Y having the form (7.100), is equivalent to $[Y_{a,b}, X] = 0$ for every choice of $a, b \in \Sigma$, and so it follows that

$$\operatorname{comm}(\mathcal{B}) = \left\{ \sum_{a,b \in \Sigma} Y_{a,b} \otimes E_{a,b} \,:\, \{Y_{a,b} \,:\, a, b \in \Sigma\} \subset \operatorname{comm}(\mathcal{A}) \right\}. \tag{7.102}$$

For a given operator $X \in \operatorname{comm}(\operatorname{comm}(\mathcal{A}))$, it is therefore evident that

$$X \otimes \mathbb{1} \in \operatorname{comm}\big(\operatorname{comm}(\mathcal{B})\big). \tag{7.103}$$

Now, define a subspace $\mathcal{V} \subseteq \mathcal{X} \otimes \mathcal{X}$ as

$$\mathcal{V} = \{\operatorname{vec}(X) \,:\, X \in \mathcal{A}\}, \tag{7.104}$$

and let $X \in \mathcal{A}$ be chosen arbitrarily. It holds that

$$(X \otimes \mathbb{1})\mathcal{V} \subseteq \mathcal{V}, \tag{7.105}$$

owing to the fact that $\mathcal{A}$ is an algebra. As $\mathcal{A}$ is self-adjoint, it follows that $X^* \in \mathcal{A}$, and therefore

$$(X^* \otimes \mathbb{1})\mathcal{V} \subseteq \mathcal{V}. \tag{7.106}$$

Lemma 7.13 therefore implies that

$$[X \otimes \mathbb{1}, \Pi_{\mathcal{V}}] = 0. \tag{7.107}$$

As $X \in \mathcal{A}$ was chosen arbitrarily, it follows that $\Pi_{\mathcal{V}} \in \operatorname{comm}(\mathcal{B})$.

Finally, let $X \in \operatorname{comm}(\operatorname{comm}(\mathcal{A}))$ be chosen arbitrarily. As was argued above, the inclusion (7.103) therefore holds, from which the commutation

relation (7.107) follows. The reverse implication of Lemma 7.13 implies the containment (7.105). In particular, given that the subalgebra $\mathcal{A}$ is unital, one has $\operatorname{vec}(\mathbb{1}) \in \mathcal{V}$, and therefore

$$\operatorname{vec}(X) = (X \otimes \mathbb{1}) \operatorname{vec}(\mathbb{1}) \in \mathcal{V}, \tag{7.108}$$

which implies $X \in \mathcal{A}$. The containment

$$\operatorname{comm}(\operatorname{comm}(\mathcal{A})) \subseteq \mathcal{A} \tag{7.109}$$

has therefore been proved, which completes the proof. □

Operator structure of the permutation-invariant operators

With von Neumann's double commutant theorem in hand, one is prepared to prove the following fundamental theorem, which concerns the operator structure of the set $\mathrm{L}(\mathcal{X})^{\otimes_{\mathrm{S}} n}$.

Theorem 7.15 *Let $\mathcal{X}$ be a complex Euclidean space, let n be a positive integer, and let $X \in \mathrm{L}(\mathcal{X}^{\otimes n})$ be an operator. The following statements are equivalent:*

1. *It holds that $[X, Y^{\otimes n}] = 0$ for all $Y \in \mathrm{L}(\mathcal{X})$.*
2. *It holds that $[X, U^{\otimes n}] = 0$ for all $U \in \mathrm{U}(\mathcal{X})$.*
3. *It holds that*

$$X = \sum_{\pi \in S_n} u(\pi) W_\pi \tag{7.110}$$

for some choice of a vector $u \in \mathbb{C}^{S_n}$.

Proof By Proposition 7.10 and Theorem 7.11, together with the bilinearity of the Lie bracket, the first and second statements are equivalent to the inclusion

$$X \in \operatorname{comm}\bigl(\mathrm{L}(\mathcal{X})^{\otimes_{\mathrm{S}} n}\bigr). \tag{7.111}$$

For the set $\mathcal{A} \subseteq \mathrm{L}(\mathcal{X}^{\otimes n})$ defined as

$$\mathcal{A} = \left\{ \sum_{\pi \in S_n} u(\pi) W_\pi \,:\, u \in \mathbb{C}^{S_n} \right\}, \tag{7.112}$$

one has that the third statement is equivalent to the inclusion $X \in \mathcal{A}$. To prove the theorem, it therefore suffices to demonstrate that

$$\mathcal{A} = \operatorname{comm}\bigl(\mathrm{L}(\mathcal{X})^{\otimes_{\mathrm{S}} n}\bigr). \tag{7.113}$$

For any operator $Z \in \mathrm{L}(\mathcal{X}^{\otimes n})$, it is evident from an inspection of (7.59)

that $Z \in \mathrm{L}(\mathcal{X})^{\otimes_{\mathrm{S}} n}$ if and only if $[Z, W_\pi] = 0$ for each $\pi \in S_n$. Again using the bilinearity of the Lie bracket, it follows that

$$\mathrm{L}(\mathcal{X})^{\otimes_{\mathrm{S}} n} = \mathrm{comm}(\mathcal{A}). \tag{7.114}$$

Finally, one observes that the set $\mathcal{A}$ forms a self-adjoint, unital subalgebra of $\mathrm{L}(\mathcal{X}^{\otimes n})$. By Theorem 7.14, one has

$$\mathrm{comm}\big(\mathrm{L}(\mathcal{X})^{\otimes_{\mathrm{S}} n}\big) = \mathrm{comm}(\mathrm{comm}(\mathcal{A})) = \mathcal{A}, \tag{7.115}$$

which establishes the relation (7.113), and therefore completes the proof. □

7.2 Unitarily invariant probability measures

Two probability measures having fundamental importance in the theory of quantum information are introduced in the present section: the *uniform spherical measure*, defined on the unit sphere $\mathcal{S}(\mathcal{X})$, and the *Haar measure*, defined on the set of unitary operators $\mathrm{U}(\mathcal{X})$, for every complex Euclidean space $\mathcal{X}$. These measures are closely connected, and may both be defined in simple and concrete terms based on the standard Gaussian measure on the real line (q.v. Section 1.2.1).

7.2.1 Uniform spherical measure and Haar measure

Definitions and basic properties of the uniform spherical measure and Haar measure are discussed below, starting with the uniform spherical measure.

Uniform spherical measure

Intuitively speaking, the uniform spherical measure provides a formalism through which one may consider a probability distribution over vectors in a complex Euclidean space that is uniform over the unit sphere. In more precise terms, the uniform spherical measure is a probability measure μ, defined on the Borel subsets of the unit sphere $\mathcal{S}(\mathcal{X})$ of a complex Euclidean space $\mathcal{X}$, that is invariant under the action of every unitary operator:

$$\mu(\mathcal{A}) = \mu(U\mathcal{A}) \tag{7.116}$$

for every $\mathcal{A} \in \mathrm{Borel}(\mathcal{S}(\mathcal{X}))$ and $U \in \mathrm{U}(\mathcal{X})$.[1] One concrete way of defining such a measure is as follows.

[1] Indeed, the measure μ is uniquely determined by these requirements. The fact that this is so will be verified through the use of the Haar measure, which is introduced below.

Definition 7.16 Let Σ be an alphabet, let $\{X_a : a \in \Sigma\} \cup \{Y_a : a \in \Sigma\}$ be a collection of independent and identically distributed standard normal random variables, and let $\mathcal{X} = \mathbb{C}^{\Sigma}$. Define a vector-valued random variable Z, taking values in $\mathcal{X}$, as

$$Z = \sum_{a \in \Sigma} (X_a + iY_a) e_a. \tag{7.117}$$

The *uniform spherical measure* μ on $\mathcal{S}(\mathcal{X})$ is the Borel probability measure

$$\mu : \text{Borel}(\mathcal{S}(\mathcal{X})) \to [0, 1] \tag{7.118}$$

defined as

$$\mu(\mathcal{A}) = \Pr\big(\alpha Z \in \mathcal{A} \text{ for some } \alpha > 0\big) \tag{7.119}$$

for every $\mathcal{A} \in \text{Borel}(\mathcal{S}(\mathcal{X}))$.

The fact that the uniform spherical measure μ is a well-defined Borel probability measure follows from three observations. First, one has that

$$\{x \in \mathcal{X} : \alpha x \in \mathcal{A} \text{ for some } \alpha > 0\} = \text{cone}(\mathcal{A}) \backslash \{0\} \tag{7.120}$$

is a Borel subset of $\mathcal{X}$ for every Borel subset $\mathcal{A}$ of $\mathcal{S}(\mathcal{X})$, which implies that μ is a well-defined function. Second, if $\mathcal{A}$ and $\mathcal{B}$ are disjoint Borel subsets of $\mathcal{S}(\mathcal{X})$, then $\text{cone}(\mathcal{A}) \backslash \{0\}$ and $\text{cone}(\mathcal{B}) \backslash \{0\}$ are also disjoint, from which it follows that μ is a measure. Finally, it holds that

$$\mu(\mathcal{S}(\mathcal{X})) = \Pr(Z \neq 0) = 1, \tag{7.121}$$

and therefore μ is a probability measure.

It is evident that this definition is independent of how one might choose to order the elements of the alphabet Σ. For this reason, the fundamentally interesting properties of the uniform spherical measure defined on $\mathcal{S}(\mathcal{X})$ will follow from the same properties of the uniform spherical measure on $\mathcal{S}(\mathbb{C}^n)$. In some cases, restricting one's attention to complex Euclidean spaces of the form $\mathbb{C}^n$ will offer conveniences, mostly concerning notational simplicity, that will therefore cause no loss of generality.

The unitary invariance of the uniform spherical measure follows directly from the rotational invariance of the standard Gaussian measure, as the proof of the following proposition reveals.

Proposition 7.17 *For every complex Euclidean space $\mathcal{X}$, the uniform spherical measure μ on $\mathcal{S}(\mathcal{X})$ is unitarily invariant:*

$$\mu(U\mathcal{A}) = \mu(\mathcal{A}) \tag{7.122}$$

for every $\mathcal{A} \in \text{Borel}(\mathcal{S}(\mathcal{X}))$ and $U \in \text{U}(\mathcal{X})$.

Proof Assume that Σ is the alphabet for which $\mathcal{X} = \mathbb{C}^{\Sigma}$, and let

$$\{X_a : a \in \Sigma\} \cup \{Y_a : a \in \Sigma\} \tag{7.123}$$

be a collection of independent and identically distributed standard normal random variables. Define vector-valued random variables X and Y, taking values in $\mathbb{R}^{\Sigma}$, as

$$X = \sum_{a \in \Sigma} X_a e_a \quad \text{and} \quad Y = \sum_{a \in \Sigma} Y_a e_a, \tag{7.124}$$

so that the vector-valued random variable Z referred to in Definition 7.16 may be expressed as $Z = X + iY$. To prove the proposition, it suffices to observe that Z and UZ are identically distributed for every unitary operator $U \in \mathrm{U}(\mathcal{X})$, for then one has that

$$\begin{aligned} \mu(U^{-1}\mathcal{A}) &= \Pr(\alpha UZ \in \mathcal{A} \text{ for some } \alpha > 0) \\ &= \Pr(\alpha Z \in \mathcal{A} \text{ for some } \alpha > 0) = \mu(\mathcal{A}) \end{aligned} \tag{7.125}$$

for every Borel subset $\mathcal{A}$ of $\mathcal{S}(\mathcal{X})$.

To verify that Z and UZ are identically distributed, for any choice of a unitary operator $U \in \mathrm{U}(\mathcal{X})$, note that

$$\begin{aligned} \begin{pmatrix} \Re(UZ) \\ \Im(UZ) \end{pmatrix} &= \begin{pmatrix} \Re(U) & -\Im(U) \\ \Im(U) & \Re(U) \end{pmatrix} \begin{pmatrix} \Re(Z) \\ \Im(Z) \end{pmatrix} \\ &= \begin{pmatrix} \Re(U) & -\Im(U) \\ \Im(U) & \Re(U) \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}, \end{aligned} \tag{7.126}$$

where $\Re(\cdot)$ and $\Im(\cdot)$ denote the entry-wise real and imaginary parts of operators and vectors, as a calculation reveals. The operator

$$\begin{pmatrix} \Re(U) & -\Im(U) \\ \Im(U) & \Re(U) \end{pmatrix} \tag{7.127}$$

is an orthogonal operator, while the vector-valued random variable $X \oplus Y$ is distributed with respect to the standard Gaussian measure on $\mathbb{R}^{\Sigma} \oplus \mathbb{R}^{\Sigma}$, and is therefore invariant under orthogonal transformations. It therefore follows that

$$X \oplus Y \quad \text{and} \quad \Re(UZ) \oplus \Im(UZ) \tag{7.128}$$

identically distributed, which implies that Z and UZ are also identically distributed. □

Haar measure

Along similar lines to the uniform spherical measure, a unitarily invariant Borel probability measure η, known as the *Haar measure*,[2] may be defined on the set of unitary operators $\mathrm{U}(\mathcal{X})$ acting on given complex Euclidean space $\mathcal{X}$. More specifically, this measure is invariant with respect to both left and right multiplication by every unitary operator:

$$\eta(U\mathcal{A}) = \eta(\mathcal{A}) = \eta(\mathcal{A}U) \tag{7.129}$$

for every choice of $\mathcal{A} \in \mathrm{Borel}(\mathrm{U}(\mathcal{X}))$ and $U \in \mathrm{U}(\mathcal{X})$.

Definition 7.18 Let Σ be an alphabet, let $\mathcal{X} = \mathbb{C}^{\Sigma}$, and let

$$\{X_{a,b} : a, b \in \Sigma\} \cup \{Y_{a,b} : a, b \in \Sigma\} \tag{7.130}$$

be a collection of independent and identically distributed standard normal random variables. Define an operator-valued random variable Z, taking values in $\mathrm{L}(\mathcal{X})$, as

$$Z = \sum_{a,b \in \Sigma} (X_{a,b} + iY_{a,b}) E_{a,b}. \tag{7.131}$$

The *Haar measure* η on $\mathrm{U}(\mathcal{X})$ is the Borel probability measure

$$\eta : \mathrm{Borel}(\mathrm{U}(\mathcal{X})) \to [0, 1] \tag{7.132}$$

defined as

$$\eta(\mathcal{A}) = \Pr\big(PZ \in \mathcal{A} \text{ for some } P \in \mathrm{Pd}(\mathcal{X})\big) \tag{7.133}$$

for every $\mathcal{A} \in \mathrm{Borel}(\mathrm{U}(\mathcal{X}))$.

As the following theorem states, the Haar measure, as just defined, is indeed a Borel probability measure.

Theorem 7.19 *Let $\eta : \mathrm{Borel}(\mathrm{U}(\mathcal{X})) \to [0, 1]$ be as in Definition 7.18, for any choice of a complex Euclidean space $\mathcal{X}$. It holds that η is a Borel probability measure.*

Proof For every $\mathcal{A} \in \mathrm{Borel}(\mathrm{U}(\mathcal{X}))$, define a set $\mathcal{R}(\mathcal{A}) \subseteq \mathrm{L}(\mathcal{X})$ as

$$\mathcal{R}(\mathcal{A}) = \{QU : Q \in \mathrm{Pd}(\mathcal{X}),\, U \in \mathcal{A}\}. \tag{7.134}$$

For any operator $X \in \mathrm{L}(\mathcal{X})$, one has that $PX \in \mathcal{A}$ for some $P \in \mathrm{Pd}(\mathcal{X})$ if and only if $X \in \mathcal{R}(\mathcal{A})$. To prove that η is a Borel measure, it therefore suffices

[2] The term *Haar measure* often refers to a more general notion, which is that of a measure defined on a certain class of groups that is invariant under the action of the group on which it is defined. The definition presented here is a restriction of this notion to the group of unitary operators acting on a given complex Euclidean space.

to prove that $\mathcal{R}(\mathcal{A})$ is a Borel subset of $\mathrm{L}(\mathcal{X})$ for every $\mathcal{A} \in \mathrm{Borel}(\mathrm{U}(\mathcal{X}))$, and that $\mathcal{R}(\mathcal{A})$ and $\mathcal{R}(\mathcal{B})$ are disjoint provided that $\mathcal{A}$ and $\mathcal{B}$ are disjoint.

The first of these requirements follows from the observation that the set $\mathrm{Pd}(\mathcal{X}) \times \mathcal{A}$ is a Borel subset of $\mathrm{Pd}(\mathcal{X}) \times \mathrm{U}(\mathcal{X})$, with respect to the product topology on the Cartesian product of these sets, together with the fact that operator multiplication is a continuous mapping.

For the second requirement, one observes that if

$$Q_0 U_0 = Q_1 U_1 \tag{7.135}$$

for some choice of $Q_0, Q_1 \in \mathrm{Pd}(\mathcal{X})$ and $U_0, U_1 \in \mathrm{U}(\mathcal{X})$, then it must hold that $Q_0 = Q_1 V$ for V being unitary. Therefore

$$Q_0^2 = Q_1 V V^* Q_1 = Q_1^2, \tag{7.136}$$

which implies that $Q_0 = Q_1$ by the fact that positive semidefinite operators have unique square roots. It therefore holds that $U_0 = U_1$. Consequently, if $\mathcal{R}(\mathcal{A}) \cap \mathcal{R}(\mathcal{B})$ is nonempty, then the same is true of $\mathcal{A} \cap \mathcal{B}$.

It remains to prove that η is a probability measure. Assume that Σ is the alphabet for which $\mathcal{X} = \mathbb{C}^\Sigma$, let

$$\{X_{a,b} \,:\, a, b \in \Sigma\} \cup \{Y_{a,b} \,:\, a, b \in \Sigma\} \tag{7.137}$$

be a collection of independent and identically distributed standard normal random variables, and define an operator-valued random variable

$$Z = \sum_{a,b \in \Sigma} (X_{a,b} + i Y_{a,b}) E_{a,b}\,, \tag{7.138}$$

as in Definition 7.18. It holds that $PZ \in \mathrm{U}(\mathcal{X})$ for some positive definite operator $P \in \mathrm{Pd}(\mathcal{X})$ if and only if Z is nonsingular, and therefore

$$\eta(\mathrm{U}(\mathcal{X})) = \Pr\big(\mathrm{Det}(Z) \neq 0\big). \tag{7.139}$$

An operator is singular if and only if its column vectors form a linearly dependent set, and therefore $\mathrm{Det}(Z) = 0$ if and only if there exists a symbol $b \in \Sigma$ such that

$$\sum_{a \in \Sigma} (X_{a,b} + i Y_{a,b}) e_a \in \mathrm{span}\left\{ \sum_{a \in \Sigma} (X_{a,c} + i Y_{a,c}) e_a \,:\, c \in \Sigma \backslash \{b\} \right\}. \tag{7.140}$$

The subspace referred to in this equation is necessarily a proper subspace of $\mathcal{X}$, because its dimension is at most $|\Sigma| - 1$, and therefore the event (7.140) occurs with probability zero. By the union bound, one has that $\mathrm{Det}(Z) = 0$ with probability zero, as is implied by Proposition 1.17, and therefore $\eta(\mathrm{U}(\mathcal{X})) = 1$. □

The following proposition establishes that the Haar measure is unitary invariant, in the sense specified by (7.129).

Proposition 7.20 *Let $\mathcal{X}$ be a complex Euclidean space. The Haar measure η on $\mathrm{U}(\mathcal{X})$ satisfies*

$$\eta(U\mathcal{A}) = \eta(\mathcal{A}) = \eta(\mathcal{A}U) \tag{7.141}$$

for every $\mathcal{A} \in \mathrm{Borel}(\mathrm{U}(\mathcal{X}))$ and $U \in \mathrm{U}(\mathcal{X})$.

Proof Assume that Σ is the alphabet for which $\mathcal{X} = \mathbb{C}^{\Sigma}$, let

$$\{X_{a,b} : a, b \in \Sigma\} \cup \{Y_{a,b} : a, b \in \Sigma\} \tag{7.142}$$

be a collection of independent and identically distributed standard normal random variables, and let

$$Z = \sum_{a,b \in \Sigma} (X_{a,b} + iY_{a,b}) E_{a,b}, \tag{7.143}$$

as in Definition 7.18.

Suppose that $\mathcal{A}$ is a Borel subset of $\mathrm{U}(\mathcal{X})$ and $U \in \mathrm{U}(\mathcal{X})$ is any unitary operator. To prove the left unitary invariance of η, it suffices to prove that Z and UZ are identically distributed, and to prove the right unitary invariance of η, it suffices to prove that Z and ZU are identically distributed, for then one has

$$\begin{aligned} \eta(U\mathcal{A}) &= \Pr\bigl(U^{-1}PZ \in \mathcal{A} \text{ for some } P \in \mathrm{Pd}(\mathcal{X})\bigr) \\ &= \Pr\bigl((U^{-1}PU)Z \in \mathcal{A} \text{ for some } P \in \mathrm{Pd}(\mathcal{X})\bigr) = \eta(\mathcal{A}) \end{aligned} \tag{7.144}$$

and

$$\begin{aligned} \eta(\mathcal{A}U) &= \Pr\bigl(PZU^{-1} \in \mathcal{A} \text{ for some } P \in \mathrm{Pd}(\mathcal{X})\bigr) \\ &= \Pr\bigl(PZ \in \mathcal{A} \text{ for some } P \in \mathrm{Pd}(\mathcal{X})\bigr) = \eta(\mathcal{A}). \end{aligned} \tag{7.145}$$

The fact that UZ, Z, and ZU are identically distributed follows, through essentially the same argument as the one used to prove Proposition 7.17, from the invariance of the standard Gaussian measure under orthogonal transformations. □

For every complex Euclidean space, one has that the Haar measure η on $\mathrm{U}(\mathcal{X})$ is the unique Borel probability measure that is both left and right unitarily invariant. Indeed, any Borel probability measure on $\mathrm{U}(\mathcal{X})$ that is either left unitarily invariant or right unitarily invariant must necessarily be equal to the Haar measure, as the following theorem reveals.

Theorem 7.21 *Let $\mathcal{X}$ be a complex Euclidean space and let*

$$\nu : \text{Borel}(\text{U}(\mathcal{X})) \to [0, 1] \tag{7.146}$$

be a Borel probability measure that possesses either of the following two properties:

1. *Left unitary invariance: $\nu(U\mathcal{A}) = \nu(\mathcal{A})$ for all Borel subsets $\mathcal{A} \subseteq \text{U}(\mathcal{X})$ and all unitary operators $U \in \text{U}(\mathcal{X})$.*
2. *Right unitary invariance: $\nu(\mathcal{A}U) = \nu(\mathcal{A})$ for all Borel subsets $\mathcal{A} \subseteq \text{U}(\mathcal{X})$ and all unitary operators $U \in \text{U}(\mathcal{X})$.*

It holds that ν is equal to the Haar measure $\eta : \text{Borel}(\text{U}(\mathcal{X})) \to [0, 1]$.

Proof It will be assumed that ν is left unitarily invariant; the case in which ν is right unitarily invariant is proved through a similar argument. Let $\mathcal{A}$ be an arbitrary Borel subset of $\text{U}(\mathcal{X})$, and let f denote the characteristic function of $\mathcal{A}$:

$$f(U) = \begin{cases} 1 & \text{if } U \in \mathcal{A} \\ 0 & \text{if } U \notin \mathcal{A} \end{cases} \tag{7.147}$$

for every $U \in \text{U}(\mathcal{X})$. One has that

$$\nu(\mathcal{A}) = \int f(U) \, \mathrm{d}\nu(U) = \int f(VU) \, \mathrm{d}\nu(U) \tag{7.148}$$

for every unitary operator $V \in \text{U}(\mathcal{X})$ by the left unitary invariance of ν. Integrating over all unitary operators V with respect to the Haar measure η yields

$$\nu(\mathcal{A}) = \iint f(VU) \, \mathrm{d}\nu(U) \, \mathrm{d}\eta(V) = \iint f(VU) \, \mathrm{d}\eta(V) \, \mathrm{d}\nu(U), \tag{7.149}$$

where the change in the order of integration is made possible by Fubini's theorem. By the right unitary invariance of Haar measure, it follows that

$$\nu(\mathcal{A}) = \iint f(V) \, \mathrm{d}\eta(V) \, \mathrm{d}\nu(U) = \int f(V) \, \mathrm{d}\eta(V) = \eta(\mathcal{A}). \tag{7.150}$$

As $\mathcal{A}$ was chosen arbitrarily, it follows that $\nu = \eta$, as required. □

The Haar measure and uniform spherical measure are closely related, as the following theorem indicates. The proof uses the same methodology as the proof of the previous theorem.

Theorem 7.22 *Let $\mathcal{X}$ be a complex Euclidean space, let μ denote the uniform spherical measure on $\mathcal{S}(\mathcal{X})$, and let η denote the Haar measure on $\mathrm{U}(\mathcal{X})$. For every $\mathcal{A} \in \mathrm{Borel}(\mathcal{S}(\mathcal{X}))$ and $x \in \mathcal{S}(\mathcal{X})$, it holds that*

$$\mu(\mathcal{A}) = \eta\big(\{U \in \mathrm{U}(\mathcal{X}) \,:\, Ux \in \mathcal{A}\}\big). \tag{7.151}$$

Proof Let $\mathcal{A}$ be any Borel subset of $\mathcal{S}(\mathcal{X})$ and let f denote the characteristic function of $\mathcal{A}$:

$$f(y) = \begin{cases} 1 & \text{if } y \in \mathcal{A} \\ 0 & \text{if } y \notin \mathcal{A} \end{cases} \tag{7.152}$$

for every $y \in \mathcal{S}(\mathcal{X})$. It holds that

$$\mu(\mathcal{A}) = \int f(y)\,\mathrm{d}\mu(y) = \int f(Uy)\,\mathrm{d}\mu(y) \tag{7.153}$$

for every $U \in \mathrm{U}(\mathcal{X})$, by the unitary invariance of the uniform spherical measure. Integrating over all $U \in \mathrm{U}(\mathcal{X})$ with respect to the Haar measure and changing the order of integration by means of Fubini's theorem yields

$$\mu(\mathcal{A}) = \iint f(Uy)\,\mathrm{d}\mu(y)\,\mathrm{d}\eta(U) = \iint f(Uy)\,\mathrm{d}\eta(U)\,\mathrm{d}\mu(y). \tag{7.154}$$

Now, for any fixed choice of unit vectors $x, y \in \mathcal{S}(\mathcal{X})$, one may choose a unitary operator $V \in \mathrm{U}(\mathcal{X})$ for which it holds that $Vy = x$. By the right unitary invariance of the Haar measure, one has

$$\int f(Uy)\,\mathrm{d}\eta(U) = \int f(UVy)\,\mathrm{d}\eta(U) = \int f(Ux)\,\mathrm{d}\eta(U). \tag{7.155}$$

Consequently,

$$\begin{aligned} \mu(\mathcal{A}) &= \iint f(Uy)\,\mathrm{d}\eta(U)\,\mathrm{d}\mu(y) = \iint f(Ux)\,\mathrm{d}\eta(U)\,\mathrm{d}\mu(y) \\ &= \int f(Ux)\,\mathrm{d}\eta(U) = \eta\big(\{U \in \mathrm{U}(\mathcal{X}) \,:\, Ux \in \mathcal{A}\}\big), \end{aligned} \tag{7.156}$$

as required. □

Noting that the proof of the previous theorem has not made use of any properties of the measure μ aside from the fact that it is normalized and unitarily invariant, one obtains the following corollary.

Corollary 7.23 *Let $\mathcal{X}$ be a complex Euclidean space and let*

$$\nu : \mathrm{Borel}(\mathcal{S}(\mathcal{X})) \to [0, 1] \tag{7.157}$$

be a Borel probability measure that is unitarily invariant: $\nu(U\mathcal{A}) = \nu(\mathcal{A})$ for every Borel subset $\mathcal{A} \subseteq \mathcal{S}(\mathcal{X})$. It holds that ν is equal to the uniform spherical measure $\mu : \mathrm{Borel}(\mathcal{S}(\mathcal{X})) \to [0, 1]$.

Evaluating integrals by means of symmetries

Some integrals defined with respect to the uniform spherical measure or Haar measure may be evaluated by considering the symmetries present in those integrals. For example, for Σ being any alphabet and μ denoting the uniform spherical measure on $\mathcal{S}(\mathbb{C}^{\Sigma})$, one has that

$$\int uu^* \mathrm{d}\mu(u) = \frac{\mathbb{1}}{|\Sigma|}. \tag{7.158}$$

This is so because the operator represented by the integral is necessarily positive semidefinite, has unit trace, and is invariant under conjugation by every unitary operator; $\mathbb{1}/|\Sigma|$ is the only operator having these properties.

The following lemma establishes a generalization of this fact, providing an alternative description of the projection onto the symmetric subspace defined in Section 7.1.1.

Lemma 7.24 *Let $\mathcal{X}$ be a complex Euclidean space, let n be a positive integer, and let μ denote the uniform spherical measure on $\mathcal{S}(\mathcal{X})$. It holds that*

$$\Pi_{\mathcal{X}^{\circledcirc n}} = \dim(\mathcal{X}^{\circledcirc n}) \int (uu^*)^{\otimes n} \mathrm{d}\mu(u). \tag{7.159}$$

Proof Let

$$P = \dim(\mathcal{X}^{\circledcirc n}) \int (uu^*)^{\otimes n} \, \mathrm{d}\mu(u), \tag{7.160}$$

and note first that

$$\mathrm{Tr}(P) = \dim(\mathcal{X}^{\circledcirc n}), \tag{7.161}$$

as μ is a normalized measure.

Next, by the unitary invariance of the uniform spherical measure, one has that $[P, U^{\otimes n}] = 0$ for every $U \in \mathrm{U}(\mathcal{X})$. By Theorem 7.15, it follows that

$$P = \sum_{\pi \in S_n} v(\pi) W_\pi \tag{7.162}$$

for some choice of a vector $v \in \mathbb{C}^{S_n}$. Using the fact that $u^{\otimes n} \in \mathcal{X}^{\circledcirc n}$ for every unit vector $u \in \mathbb{C}^{\Sigma}$, one necessarily has that

$$\Pi_{\mathcal{X}^{\circledcirc n}} P = P, \tag{7.163}$$

which implies

$$\begin{aligned} P &= \frac{1}{n!} \sum_{\sigma \in S_n} W_\sigma \sum_{\pi \in S_n} v(\pi) W_\pi = \frac{1}{n!} \sum_{\pi \in S_n} \sum_{\sigma \in S_n} v(\sigma^{-1}\pi) W_\pi \\ &= \frac{1}{n!} \sum_{\sigma \in S_n} v(\sigma) \sum_{\pi \in S_n} W_\pi = \sum_{\sigma \in S_n} v(\sigma) \Pi_{\mathcal{X}^{\otimes n}} \end{aligned} \tag{7.164}$$

by Proposition 7.1. By (7.161), one has

$$\sum_{\sigma \in S_n} v(\sigma) = 1, \tag{7.165}$$

and therefore $P = \Pi_{\mathcal{X}^{\otimes n}}$, as required. □

The following example represents a continuation of Example 6.10. Two channels that have a close connection to the classes of Werner states and isotropic states are analyzed based on properties of their symmetries.

Example 7.25 As in Example 6.10, let Σ be an alphabet, let $n = |\Sigma|$, and let $\mathcal{X} = \mathbb{C}^\Sigma$, and recall the four projection operators[3]

$$\Delta_0,\, \Delta_1,\, \Pi_0,\, \Pi_1 \in \operatorname{Proj}(\mathcal{X} \otimes \mathcal{X}) \tag{7.166}$$

defined in that example:

$$\Delta_0 = \frac{1}{n} \sum_{a,b \in \Sigma} E_{a,b} \otimes E_{a,b}, \tag{7.167}$$

$$\Delta_1 = \mathbb{1} \otimes \mathbb{1} - \frac{1}{n} \sum_{a,b \in \Sigma} E_{a,b} \otimes E_{a,b}\,, \tag{7.168}$$

$$\Pi_0 = \frac{1}{2}\mathbb{1} \otimes \mathbb{1} + \frac{1}{2} \sum_{a,b \in \Sigma} E_{a,b} \otimes E_{b,a}\,, \tag{7.169}$$

$$\Pi_1 = \frac{1}{2}\mathbb{1} \otimes \mathbb{1} - \frac{1}{2} \sum_{a,b \in \Sigma} E_{a,b} \otimes E_{b,a}\,. \tag{7.170}$$

Equivalently, one may write

$$\Delta_0 = \frac{1}{n}(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(W)\,, \qquad \Pi_0 = \frac{1}{2}\mathbb{1} \otimes \mathbb{1} + \frac{1}{2}W\,, \tag{7.171}$$

$$\Delta_1 = \mathbb{1} \otimes \mathbb{1} - \frac{1}{n}(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(W)\,, \quad \Pi_1 = \frac{1}{2}\mathbb{1} \otimes \mathbb{1} - \frac{1}{2}W\,, \tag{7.172}$$

[3] Using the notation introduced in Section 7.1.1, one may alternatively write $\Pi_0 = \Pi_{\mathcal{X} \otimes \mathcal{X}}$ and $\Pi_1 = \Pi_{\mathcal{X} \oslash \mathcal{X}}$. The notations Π_0 and Π_1 will be used within this example to maintain consistency with Example 6.10.

for $\mathrm{T}(X) = X^{\mathsf{T}}$ denoting the transpose mapping on $\mathrm{L}(\mathcal{X})$ and

$$W = \sum_{a,b \in \Sigma} E_{a,b} \otimes E_{b,a}\,, \tag{7.173}$$

which is the swap operator on $\mathcal{X} \otimes \mathcal{X}$. States of the form

$$\lambda \Delta_0 + (1-\lambda)\frac{\Delta_1}{n^2-1} \quad \text{and} \quad \lambda \frac{\Pi_0}{\binom{n+1}{2}} + (1-\lambda)\frac{\Pi_1}{\binom{n}{2}}\,, \tag{7.174}$$

for $\lambda \in [0,1]$, were introduced in Example 6.10 as *isotropic states* and *Werner states*, respectively.

Now, consider the channel $\Xi \in \mathrm{C}(\mathcal{X} \otimes \mathcal{X})$ defined as

$$\Xi(X) = \int (U \otimes U) X (U \otimes U)^* \, \mathrm{d}\eta(U) \tag{7.175}$$

for all $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$, for η denoting the Haar measure on $\mathrm{U}(\mathcal{X})$. By the unitary invariance of Haar measure, one has that $[\Xi(X), U \otimes U] = 0$ for every $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$ and $U \in \mathrm{U}(\mathcal{X})$. By Theorem 7.15 it holds that

$$\Xi(X) \in \operatorname{span}\{\mathbb{1} \otimes \mathbb{1}, W\} = \operatorname{span}\{\Pi_0, \Pi_1\}, \tag{7.176}$$

and it must therefore hold that

$$\Xi(X) = \alpha(X)\,\Pi_0 + \beta(X)\,\Pi_1 \tag{7.177}$$

for $\alpha(X), \beta(X) \in \mathbb{C}$ being complex numbers depending linearly on X. The channel Ξ is self-adjoint and satisfies $\Xi(\mathbb{1} \otimes \mathbb{1}) = \mathbb{1} \otimes \mathbb{1}$ and $\Xi(W) = W$, so that $\Xi(\Pi_0) = \Pi_0$ and $\Xi(\Pi_1) = \Pi_1$. The following two equations hold:

$$\begin{aligned} \alpha(X) &= \frac{1}{\binom{n+1}{2}} \langle \Pi_0, \Xi(X) \rangle = \frac{1}{\binom{n+1}{2}} \langle \Xi(\Pi_0), X \rangle = \frac{1}{\binom{n+1}{2}} \langle \Pi_0, X \rangle \\ \beta(X) &= \frac{1}{\binom{n}{2}} \langle \Pi_1, \Xi(X) \rangle = \frac{1}{\binom{n}{2}} \langle \Xi(\Pi_1), X \rangle = \frac{1}{\binom{n}{2}} \langle \Pi_1, X \rangle. \end{aligned} \tag{7.178}$$

It therefore follows that

$$\Xi(X) = \frac{1}{\binom{n+1}{2}} \langle \Pi_0, X \rangle \Pi_0 + \frac{1}{\binom{n}{2}} \langle \Pi_1, X \rangle \Pi_1. \tag{7.179}$$

It is evident from this expression that, on any density operator input, the output of Ξ is a Werner state, and moreover every Werner state is fixed by this channel. The channel Ξ is sometimes called a *Werner twirling channel*.

A different but closely related channel $\Lambda \in \mathrm{C}(\mathcal{X} \otimes \mathcal{X})$ is defined as

$$\Lambda(X) = \int \big(U \otimes \overline{U}\big) X \big(U \otimes \overline{U}\big)^* \, \mathrm{d}\eta(U) \tag{7.180}$$

for all $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$, where η again denotes the Haar measure on $\mathrm{U}(\mathcal{X})$.

An alternate expression of this channel may be obtained by making use of the analysis of the channel Ξ presented above. The first step of this process is to observe that Λ may be obtained by composing the channel Ξ with the partial transpose in the following way:

$$\Lambda = (\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \mathrm{T})\,\Xi\,(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \mathrm{T}). \tag{7.181}$$

Then, using the identities

$$\begin{aligned} (\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \mathrm{T})(\Pi_0) &= \frac{n+1}{2}\Delta_0 + \frac{1}{2}\Delta_1, \\ (\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \mathrm{T})(\Pi_1) &= -\frac{n-1}{2}\Delta_0 + \frac{1}{2}\Delta_1, \end{aligned} \tag{7.182}$$

one finds that

$$\Lambda(X) = \langle \Delta_0, X\rangle \Delta_0 + \frac{1}{n^2-1}\langle \Delta_1, X\rangle \Delta_1. \tag{7.183}$$

On any density operator input, the output of the channel Λ is an isotropic state, and moreover every isotropic state is fixed by Λ. The channel Λ is sometimes called an *isotropic twirling channel*.

It is evident from the specification of the channels Ξ and Λ that one has the following expressions, in which Φ_U denotes the unitary channel defined by $\Phi_U(X) = UXU^*$ for each $X \in \mathrm{L}(\mathcal{X})$:

$$\begin{aligned} \Xi &\in \mathrm{conv}\big\{\Phi_U \otimes \Phi_U \,:\, U \in \mathrm{U}(\mathcal{X})\big\}, \\ \Lambda &\in \mathrm{conv}\big\{\Phi_U \otimes \Phi_{\overline{U}} \,:\, U \in \mathrm{U}(\mathcal{X})\big\}. \end{aligned} \tag{7.184}$$

It follows that Ξ and Λ are mixed-unitary channels, and LOCC channels as well. Indeed, both channels can be implemented without communication—local operations and shared randomness are sufficient.

Finally, for any choice of orthogonal unit vectors $u, v \in \mathcal{X}$, the following equalities may be observed:

$$\begin{aligned} \big\langle \Pi_0, uu^* \otimes vv^* \big\rangle &= \frac{1}{2}, \quad \big\langle \Pi_1, uu^* \otimes vv^* \big\rangle = \frac{1}{2}, \\ \big\langle \Pi_0, uu^* \otimes uu^* \big\rangle &= 1, \quad \big\langle \Pi_1, uu^* \otimes uu^* \big\rangle = 0. \end{aligned} \tag{7.185}$$

Therefore, for every choice of $\alpha \in [0,1]$, one has

$$\Xi(uu^* \otimes (\alpha uu^* + (1-\alpha)vv^*)) = \frac{1+\alpha}{2}\frac{\Pi_0}{\binom{n+1}{2}} + \frac{1-\alpha}{2}\frac{\Pi_1}{\binom{n}{2}}. \tag{7.186}$$

As Ξ is a separable channel and

$$uu^* \otimes (\alpha uu^* + (1-\alpha)vv^*) \in \mathrm{SepD}(\mathcal{X} : \mathcal{X}) \tag{7.187}$$

is a separable state, for every $\alpha \in [0,1]$, it follows that the state (7.186) is also separable. Equivalently, the Werner state

$$\lambda \frac{\Pi_0}{\binom{n+1}{2}} + (1-\lambda)\frac{\Pi_1}{\binom{n}{2}} \tag{7.188}$$

is separable for all $\lambda \in [1/2, 1]$. The partial transpose of the state (7.188) is

$$\frac{2\lambda - 1}{n}\Delta_0 + \Big(1 - \frac{2\lambda - 1}{n}\Big)\frac{\Delta_1}{n^2 - 1}. \tag{7.189}$$

Assuming $\lambda \in [1/2, 1]$, the state (7.188) is separable, and therefore its partial transpose is also separable. It follows that the isotropic state

$$\lambda \Delta_0 + (1-\lambda)\frac{\Delta_1}{n^2 - 1} \tag{7.190}$$

is separable for all $\lambda \in [0, 1/n]$.

7.2.2 Applications of unitarily invariant measures

There are many applications of integration with respect to the uniform spherical measure and Haar measure in quantum information theory. Three examples are presented below, and some additional examples involving the phenomenon of *measure concentration* are presented in Section 7.3.2.

The quantum de Finetti theorem

Intuitively speaking, the quantum de Finetti theorem states that if the state of a collection of identical registers is exchangeable, then the reduced state of any comparatively small number of these registers must be close to a convex combination of identical product states. This theorem will first be stated and proved for symmetric pure states, and from this theorem a more general statement for arbitrary exchangeable states may be derived using Theorem 7.12.

Theorem 7.26 *Let Σ be an alphabet, let n be a positive integer, and let $\mathsf{X}_1, \ldots, \mathsf{X}_n$ be registers, each having classical state set Σ. Also let*

$$v \in \mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n \tag{7.191}$$

be a symmetric unit vector and let $k \in \{1, \ldots, n\}$. There exists a state

$$\tau \in \operatorname{conv}\Big\{(uu^*)^{\otimes k} \,:\, u \in \mathcal{S}\big(\mathbb{C}^{\Sigma}\big)\Big\} \tag{7.192}$$

such that

$$\big\| (vv^*)[\mathsf{X}_1, \ldots, \mathsf{X}_k] - \tau \big\|_1 \leq \frac{4k\big(|\Sigma| - 1\big)}{n+1}. \tag{7.193}$$

Proof It will be proved that the requirements of the theorem are satisfied by the operator

$$\tau = \binom{n + |\Sigma| - 1}{|\Sigma| - 1} \int \langle (uu^*)^{\otimes n}, vv^* \rangle (uu^*)^{\otimes k} \, \mathrm{d}\mu(u), \tag{7.194}$$

for μ denoting the uniform spherical measure on $\mathcal{S}(\mathbb{C}^\Sigma)$. The fact that τ is positive semidefinite is evident from its definition, and by Lemma 7.24, together with the assumption $v \in \mathcal{X}_1 \otimes\!\!\!\!\bigcirc \cdots \otimes\!\!\!\!\bigcirc \mathcal{X}_n$, one has that $\mathrm{Tr}(\tau) = 1$.

For the sake of establishing the bound (7.193), it is convenient to define

$$N_m = \binom{m + |\Sigma| - 1}{|\Sigma| - 1} \tag{7.195}$$

for every nonnegative integer m. The following bounds on the ratio between N_{n-k} and N_n hold:

$$\begin{aligned} 1 \geq \frac{N_{n-k}}{N_n} &= \frac{n - k + |\Sigma| - 1}{n + |\Sigma| - 1} \cdots \frac{n - k + 1}{n + 1} \\ &\geq \left(\frac{n - k + 1}{n + 1} \right)^{|\Sigma| - 1} \geq 1 - \frac{k\big(|\Sigma| - 1\big)}{n + 1}. \end{aligned} \tag{7.196}$$

For every unit vector $u \in \mathcal{S}(\mathbb{C}^\Sigma)$ and every positive integer m, define a projection operator

$$\Delta_{m,u} = (uu^*)^{\otimes m}, \tag{7.197}$$

and also define an operator $P_u \in \mathrm{Pos}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k)$ as

$$P_u = \mathrm{Tr}_{\mathcal{X}_{k+1} \otimes \cdots \otimes \mathcal{X}_n} \Big((\mathbb{1}_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \otimes \Delta_{n-k,u}) vv^* \Big). \tag{7.198}$$

By Lemma 7.24, together with the assumption $v \in \mathcal{X}_1 \otimes\!\!\!\!\bigcirc \cdots \otimes\!\!\!\!\bigcirc \mathcal{X}_n$, one has that

$$vv^* = N_{n-k} \int (\mathbb{1}_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \otimes \Delta_{n-k,u}) vv^* \mathrm{d}\mu(u), \tag{7.199}$$

and therefore

$$(vv^*)[\mathsf{X}_1, \ldots, \mathsf{X}_k] = N_{n-k} \int P_u \, \mathrm{d}\mu(u). \tag{7.200}$$

This density operator is to be compared with τ, which may be expressed as

$$\tau = N_n \int \Delta_{k,u} P_u \Delta_{k,u} \, \mathrm{d}\mu(u). \tag{7.201}$$

The primary goal of the remainder of the proof is to bound the trace norm of the operator

$$\frac{1}{N_{n-k}}(vv^*)[\mathsf{X}_1,\ldots,\mathsf{X}_k] - \frac{1}{N_n}\tau = \int \Big(P_u - \Delta_{k,u}P_u\Delta_{k,u}\Big)\,\mathrm{d}\mu(u), \tag{7.202}$$

as such a bound will lead directly to a bound on the trace norm of

$$(vv^*)[\mathsf{X}_1,\ldots,\mathsf{X}_k] - \tau. \tag{7.203}$$

The operator identity

$$A - BAB = A(\mathbb{1} - B) + (\mathbb{1} - B)A - (\mathbb{1} - B)A(\mathbb{1} - B), \tag{7.204}$$

which holds for any two square operators A and B acting on a given space, will be useful for this purpose. It holds that

$$\begin{aligned}\int \Delta_{k,u}P_u\,\mathrm{d}\mu(u) &= \int \mathrm{Tr}_{\mathcal{X}_{k+1}\otimes\cdots\otimes\mathcal{X}_n}\Big(\Delta_{n,u}vv^*\Big)\,\mathrm{d}\mu(u)\\ &= \frac{1}{N_n}(vv^*)[\mathsf{X}_1,\ldots,\mathsf{X}_k],\end{aligned} \tag{7.205}$$

and therefore

$$\int (\mathbb{1} - \Delta_{k,u})P_u\,\mathrm{d}\mu(u) = \left(\frac{1}{N_{n-k}} - \frac{1}{N_n}\right)(vv^*)[\mathsf{X}_1,\ldots,\mathsf{X}_k], \tag{7.206}$$

which implies

$$\left\|\int (\mathbb{1} - \Delta_{k,u})P_u\,\mathrm{d}\mu(u)\right\|_1 = \left(\frac{1}{N_{n-k}} - \frac{1}{N_n}\right). \tag{7.207}$$

By similar reasoning, one finds that

$$\left\|\int P_u(\mathbb{1} - \Delta_{k,u})\,\mathrm{d}\mu(u)\right\|_1 = \left(\frac{1}{N_{n-k}} - \frac{1}{N_n}\right). \tag{7.208}$$

Moreover, one has

$$\begin{aligned}&\left\|\int (\mathbb{1} - \Delta_{k,u})P_u(\mathbb{1} - \Delta_{k,u})\,\mathrm{d}\mu(u)\right\|_1\\ &\qquad= \mathrm{Tr}\left(\int (\mathbb{1} - \Delta_{k,u})P_u(\mathbb{1} - \Delta_{k,u})\,\mathrm{d}\mu(u)\right)\\ &\qquad= \mathrm{Tr}\left(\int (\mathbb{1} - \Delta_{k,u})P_u\,\mathrm{d}\mu(u)\right) = \left(\frac{1}{N_{n-k}} - \frac{1}{N_n}\right),\end{aligned} \tag{7.209}$$

and therefore, by the triangle inequality together with the identity (7.204), it follows that

$$\left\|\frac{1}{N_{n-k}}(vv^*)[\mathsf{X}_1,\ldots,\mathsf{X}_k] - \frac{1}{N_n}\tau\right\|_1 \leq 3\left(\frac{1}{N_{n-k}} - \frac{1}{N_n}\right). \tag{7.210}$$

Having established a bound on the trace norm of the operator (7.202), the theorem follows:

$$\begin{aligned}
&\Big\| (vv^*)[\mathsf{X}_1, \ldots, \mathsf{X}_k] - \tau \Big\|_1 \\
&\quad \leq N_{n-k} \left\| \frac{1}{N_{n-k}} (vv^*)[\mathsf{X}_1, \ldots, \mathsf{X}_k] - \frac{1}{N_n} \tau \right\|_1 \\
&\qquad + N_{n-k} \left\| \frac{1}{N_n} \tau - \frac{1}{N_{n-k}} \tau \right\|_1 \\
&\quad \leq 4 \left(1 - \frac{N_{n-k}}{N_n} \right) \\
&\quad \leq \frac{4k\big(|\Sigma| - 1\big)}{n+1},
\end{aligned} \tag{7.211}$$

as required. □

Corollary 7.27 (Quantum de Finetti theorem) *Let Σ be an alphabet, let n be a positive integer, and let $\mathsf{X}_1, \ldots, \mathsf{X}_n$ be registers sharing the same classical state set Σ. For every exchangeable density operator $\rho \in \mathrm{D}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n)$ and every positive integer $k \in \{1, \ldots, n\}$, there exists a density operator*

$$\tau \in \operatorname{conv}\big\{ \sigma^{\otimes k} \,:\, \sigma \in \mathrm{D}\big(\mathbb{C}^{\Sigma}\big) \big\} \tag{7.212}$$

such that

$$\big\| \rho[\mathsf{X}_1, \ldots, \mathsf{X}_k] - \tau \big\|_1 \leq \frac{4k\big(|\Sigma|^2 - 1\big)}{n+1}. \tag{7.213}$$

Proof Let $\mathsf{Y}_1, \ldots, \mathsf{Y}_n$ be registers, all sharing the classical state set Σ. By Theorem 7.12, there exists a symmetric unit vector

$$v \in (\mathcal{X}_1 \otimes \mathcal{Y}_1) \mathbin{\circledcirc} \cdots \mathbin{\circledcirc} (\mathcal{X}_n \otimes \mathcal{Y}_n), \tag{7.214}$$

representing a pure state of the compound register $((\mathsf{X}_1, \mathsf{Y}_1), \ldots, (\mathsf{X}_n, \mathsf{Y}_n))$, with the property that

$$(vv^*)[\mathsf{X}_1, \ldots, \mathsf{X}_n] = \rho. \tag{7.215}$$

By Theorem 7.26, there exists a density operator

$$\xi \in \operatorname{conv}\big\{ (uu^*)^{\otimes k} \,:\, u \in \mathcal{S}\big(\mathbb{C}^{\Sigma} \otimes \mathbb{C}^{\Sigma}\big) \big\}, \tag{7.216}$$

representing a state of the compound register $((\mathsf{X}_1, \mathsf{Y}_1), \ldots, (\mathsf{X}_k, \mathsf{Y}_k))$, such that

$$\big\| (vv^*)[(\mathsf{X}_1, \mathsf{Y}_1), \ldots, (\mathsf{X}_k, \mathsf{Y}_k)] - \xi \big\|_1 \leq \frac{4k\big(|\Sigma|^2 - 1\big)}{n+1}. \tag{7.217}$$

Taking $\tau = \xi[\mathsf{X}_1, \ldots, \mathsf{X}_k]$, one has that

$$\tau \in \operatorname{conv}\{\sigma^{\otimes k} \,:\, \sigma \in \mathrm{D}(\mathbb{C}^{\Sigma})\}, \tag{7.218}$$

and the required bound

$$\begin{aligned} \bigl\|\rho[\mathsf{X}_1, \ldots, \mathsf{X}_k] - \tau\bigr\|_1 &\leq \bigl\|(vv^*)[(\mathsf{X}_1, \mathsf{Y}_1), \ldots, (\mathsf{X}_k, \mathsf{Y}_k)] - \xi\bigr\|_1 \\ &\leq \frac{4k\bigl(|\Sigma|^2 - 1\bigr)}{n+1} \end{aligned} \tag{7.219}$$

follows by the monotonicity of the trace norm under partial tracing. □

Optimal cloning of pure quantum states

Let Σ be an alphabet, let n and m be positive integers with $n \leq m$, and let $\mathsf{X}_1, \ldots, \mathsf{X}_m$ be registers, all sharing the same classical state Σ. In the task of *cloning*, one assumes that the state of $(\mathsf{X}_1, \ldots, \mathsf{X}_n)$ is given by

$$\rho^{\otimes n} \in \mathrm{D}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n), \tag{7.220}$$

for some choice of $\rho \in \mathrm{D}(\mathbb{C}^{\Sigma})$, and the goal is to transform $(\mathsf{X}_1, \ldots, \mathsf{X}_n)$ into $(\mathsf{X}_1, \ldots, \mathsf{X}_m)$ in such a way that the resulting state of this register is as close as possible to

$$\rho^{\otimes m} \in \mathrm{D}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_m). \tag{7.221}$$

One may consider the quality with which a given channel

$$\Phi \in \mathrm{C}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n, \mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_m) \tag{7.222}$$

performs this task in a variety of specific ways. For example, one might measure the closeness of $\Phi(\rho^n)$ to ρ^m with respect to the trace norm, some other norm, or the fidelity function; and one might consider the average closeness over some distribution on the possible choices of ρ, or consider the worst case over all ρ or over some subset of possible choices for ρ. It is most typical that one assumes ρ is a pure state—the mixed state case is more complicated and has very different characteristics from the pure state case.

The specific variant of the cloning task that will be considered here is that one aims to choose a channel of the form (7.222) so as to maximize the minimum fidelity

$$\alpha(\Phi) = \inf_{u \in \mathcal{S}(\mathbb{C}^{\Sigma})} \mathrm{F}\bigl(\Phi\bigl((uu^*)^{\otimes n}\bigr), (uu^*)^{\otimes m}\bigr) \tag{7.223}$$

over all pure states $\rho = uu^*$. The following theorem establishes an upper bound on this quantity, and states that this bound is achieved for some choice of a channel Φ.

Theorem 7.28 (Werner) *Let $\mathcal{X}$ be a complex Euclidean space and let n and m be positive integers with $n \leq m$. For every channel*

$$\Phi \in \mathrm{C}(\mathcal{X}^{\otimes n}, \mathcal{X}^{\otimes m}) \tag{7.224}$$

it holds that

$$\inf_{u \in \mathcal{S}(\mathcal{X})} \left\langle \Phi\big((uu^*)^{\otimes n}\big), (uu^*)^{\otimes m} \right\rangle \leq \frac{N_n}{N_m}, \tag{7.225}$$

where

$$N_k = \binom{k + \dim(\mathcal{X}) - 1}{\dim(\mathcal{X}) - 1} \tag{7.226}$$

for each positive integer k. Moreover, there exists a channel Φ of the above form for which equality is achieved in (7.225).

Remark In the case that $n = 1$ and $m = 2$, one has

$$\frac{N_1}{N_2} = \frac{2}{\dim(\mathcal{X}) + 1}, \tag{7.227}$$

which is strictly less than 1 if $\dim(\mathcal{X}) \geq 2$. Theorem 7.28 therefore provides a quantitative form of the *no-cloning theorem*, which states that it is not possible to create a perfect copy of an unknown quantum state (aside from the trivial case of one-dimensional systems).

Proof The infimum on the left-hand side of (7.225) can be no larger than the average with respect to the uniform spherical measure on $\mathcal{S}(\mathcal{X})$:

$$\begin{aligned} &\inf_{u \in \mathcal{S}(\mathcal{X})} \left\langle \Phi\big((uu^*)^{\otimes n}\big), (uu^*)^{\otimes m} \right\rangle \\ &\qquad \leq \int \left\langle \Phi\big((uu^*)^{\otimes n}\big), (uu^*)^{\otimes m} \right\rangle \mathrm{d}\mu(u). \end{aligned} \tag{7.228}$$

As $(uu^*)^{\otimes n} \leq \Pi_{\mathcal{X}^{\otimes n}}$ for every $u \in \mathcal{S}(\mathcal{X})$, it follows that

$$\begin{aligned} \int \left\langle \Phi\big((uu^*)^{\otimes n}\big), (uu^*)^{\otimes m} \right\rangle \mathrm{d}\mu(u) &\leq \int \left\langle \Phi\big(\Pi_{\mathcal{X}^{\otimes n}}\big), (uu^*)^{\otimes m} \right\rangle \mathrm{d}\mu(u) \\ &= \frac{1}{N_m} \left\langle \Phi\big(\Pi_{\mathcal{X}^{\otimes n}}\big), \Pi_{\mathcal{X}^{\otimes m}} \right\rangle \leq \frac{1}{N_m} \operatorname{Tr}\big(\Phi\big(\Pi_{\mathcal{X}^{\otimes n}}\big)\big) = \frac{N_n}{N_m}. \end{aligned} \tag{7.229}$$

This establish the required bound (7.225).

It remains to prove that there exists a channel

$$\Phi \in \mathrm{C}(\mathcal{X}^{\otimes n}, \mathcal{X}^{\otimes m}) \tag{7.230}$$

for which equality is achieved in (7.225). Define

$$\Phi(X) = \frac{N_n}{N_m} \Pi_{\mathcal{X}^{\otimes m}} \Big(X \otimes \mathbb{1}_{\mathcal{X}}^{\otimes (m-n)} \Big) \Pi_{\mathcal{X}^{\otimes m}} + \Big\langle \mathbb{1}_{\mathcal{X}}^{\otimes n} - \Pi_{\mathcal{X}^{\otimes n}}, X \Big\rangle \, \sigma \tag{7.231}$$

for all $X \in \mathrm{L}(\mathcal{X}^{\otimes n})$, where $\sigma \in \mathrm{D}(\mathcal{X}^{\otimes m})$ is an arbitrary density operator. It is evident that Φ is completely positive, and the fact that Φ preserves trace follows from the observation

$$\Big(\mathbb{1}_{\mathrm{L}(\mathcal{X})}^{\otimes n} \otimes \mathrm{Tr}_{\mathcal{X}}^{\otimes (m-n)} \Big) (\Pi_{\mathcal{X}^{\otimes m}}) = \frac{N_m}{N_n} \, \Pi_{\mathcal{X}^{\otimes n}}. \tag{7.232}$$

A direct calculation reveals that

$$\big\langle (uu^*)^{\otimes m}, \Phi\big((uu^*)^{\otimes n}\big) \big\rangle = \frac{N_n}{N_m} \tag{7.233}$$

for every unit vector $u \in \mathcal{S}(\mathcal{X})$, which completes the proof. □

Example 7.29 The channel described in Example 2.33 is an optimal cloning channel, achieving equality in (7.225) for the case $\mathcal{X} = \mathbb{C}^2$, $n = 1$, and $m = 2$.

Unital channels near the completely depolarizing channel

The final example of an application of unitarily invariant measures in the theory of quantum information to be presented in this section demonstrates that all unital channels sufficiently close to the completely depolarizing channel must be mixed-unitary channels. The following lemma will be used to demonstrate this fact.

Lemma 7.30 *Let $\mathcal{X}$ be a complex Euclidean space having dimension $n \geq 2$, let η denote the Haar measure on $\mathrm{U}(\mathcal{X})$, and let $\Omega \in \mathrm{C}(\mathcal{X})$ denote the completely depolarizing channel defined with respect to the space $\mathcal{X}$. The map $\Xi \in \mathrm{CP}(\mathcal{X} \otimes \mathcal{X})$ defined as*

$$\Xi(X) = \int \langle \mathrm{vec}(U)\,\mathrm{vec}(U)^*, X \rangle \, \mathrm{vec}(U)\,\mathrm{vec}(U)^* \, \mathrm{d}\eta(U) \tag{7.234}$$

for every $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$ is given by

$$\Xi = \frac{1}{n^2 - 1} (\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})} - \Omega \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})} - \mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Omega + n^2 \Omega \otimes \Omega). \tag{7.235}$$

Proof Let $V \in \mathrm{U}(\mathcal{X} \otimes \mathcal{X} \otimes \mathcal{X} \otimes \mathcal{X})$ be the permutation operator defined by the equation

$$V \operatorname{vec}(Y \otimes Z) = \operatorname{vec}(Y) \otimes \operatorname{vec}(Z), \tag{7.236}$$

holding for all $Y, Z \in \mathrm{L}(\mathcal{X})$. Alternatively, this operator may be defined by the equation

$$V(x_1 \otimes x_2 \otimes x_3 \otimes x_4) = x_1 \otimes x_3 \otimes x_2 \otimes x_4 \tag{7.237}$$

holding for all $x_1, x_2, x_3, x_4 \in \mathcal{X}$. As V is its own inverse, one has

$$V\big(\operatorname{vec}(Y) \otimes \operatorname{vec}(Z)\big) = \operatorname{vec}(Y \otimes Z) \tag{7.238}$$

for all $Y, Z \in \mathrm{L}(\mathcal{X})$. For every choice of maps $\Phi_0, \Phi_1 \in \mathrm{T}(\mathcal{X})$, it holds that

$$V J(\Phi_0 \otimes \Phi_1) V^* = J(\Phi_0) \otimes J(\Phi_1). \tag{7.239}$$

Now, the Choi representation of Ξ is given by

$$J(\Xi) = \int \operatorname{vec}(U) \operatorname{vec}(U)^* \otimes \operatorname{vec}\big(\overline{U}\big) \operatorname{vec}\big(\overline{U}\big)^* \mathrm{d}\eta(U), \tag{7.240}$$

and therefore

$$V J(\Xi) V^* = \int \operatorname{vec}\big(U \otimes \overline{U}\big) \operatorname{vec}\big(U \otimes \overline{U}\big)^* \mathrm{d}\eta(U). \tag{7.241}$$

This operator is the Choi representation of the isotropic twirling channel

$$\Lambda(X) = \int \big(U \otimes \overline{U}\big) X \big(U \otimes \overline{U}\big)^* \, \mathrm{d}\eta(U) \tag{7.242}$$

defined in Example 7.25. From the analysis presented in that example, it follows that

$$\begin{aligned} V J(\Xi) V^* &= \frac{1}{n^2} J(\mathbb{1}_{\mathrm{L}(\mathcal{X})}) \otimes J(\mathbb{1}_{\mathrm{L}(\mathcal{X})}) \\ &\quad + \frac{1}{n^2 - 1} \Big(n J(\Omega) - \frac{1}{n} J(\mathbb{1}_{\mathrm{L}(\mathcal{X})}) \Big) \otimes \Big(n J(\Omega) - \frac{1}{n} J(\mathbb{1}_{\mathrm{L}(\mathcal{X})}) \Big). \end{aligned} \tag{7.243}$$

By expanding the expression (7.243) and making use of the identity (7.239), one obtains (7.235), as required. □

Theorem 7.31 *Let $\mathcal{X}$ be a complex Euclidean space with dimension $n \geq 2$, let $\Omega \in \mathrm{C}(\mathcal{X})$ denote the completely depolarizing channel defined with respect to the space $\mathcal{X}$, and let $\Phi \in \mathrm{C}(\mathcal{X})$ be a unital channel. The channel*

$$\frac{n^2 - 2}{n^2 - 1} \Omega + \frac{1}{n^2 - 1} \Phi \tag{7.244}$$

is a mixed-unitary channel.

Proof Let $\Psi \in \mathrm{CP}(\mathcal{X})$ be the map defined as

$$\Psi(X) = \int \big\langle \mathrm{vec}(U)\,\mathrm{vec}(U)^*, J(\Phi)\big\rangle\, UXU^*\,\mathrm{d}\eta(U), \tag{7.245}$$

for η being the Haar measure on $\mathrm{U}(\mathcal{X})$. It holds that

$$\int \mathrm{vec}(U)\,\mathrm{vec}(U)^*\,\mathrm{d}\eta(U) = \frac{1}{n}\mathbb{1}_{\mathcal{X}\otimes\mathcal{X}}, \tag{7.246}$$

and therefore

$$\int \big\langle \mathrm{vec}(U)\,\mathrm{vec}(U)^*, J(\Phi)\big\rangle\,\mathrm{d}\eta(U) = \frac{1}{n}\,\mathrm{Tr}(J(\Phi)) = 1. \tag{7.247}$$

It follows that the mapping Ψ is a mixed-unitary channel.

By Lemma 7.30, one has $J(\Psi) = \Xi(J(\Phi))$ for $\Xi \in \mathrm{CP}(\mathcal{X}\otimes\mathcal{X})$ being defined as

$$\Xi = \frac{1}{n^2-1}\big(\mathbb{1}_{\mathrm{L}(\mathcal{X})}\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})} - \Omega\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})} - \mathbb{1}_{\mathrm{L}(\mathcal{X})}\otimes\Omega + n^2\Omega\otimes\Omega\big). \tag{7.248}$$

By the assumption that Φ is a unital channel, one has

$$\begin{aligned}(\Omega\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})})(J(\Phi)) &= (\mathbb{1}_{\mathrm{L}(\mathcal{X})}\otimes\Omega)(J(\Phi))\\ &= (\Omega\otimes\Omega)(J(\Phi)) = \frac{\mathbb{1}_{\mathcal{X}}\otimes\mathbb{1}_{\mathcal{X}}}{n},\end{aligned} \tag{7.249}$$

and therefore

$$J(\Psi) = \frac{1}{n^2-1}J(\Phi) + \frac{n^2-2}{n(n^2-1)}\mathbb{1}_{\mathcal{X}}\otimes\mathbb{1}_{\mathcal{X}}. \tag{7.250}$$

This is equivalent to Ψ being equal to (7.244), and therefore completes the proof. □

Corollary 7.32 *Let $\mathcal{X}$ be a complex Euclidean space having dimension $n \geq 2$, let $\Omega \in \mathrm{C}(\mathcal{X})$ denote the completely depolarizing channel defined with respect to the space $\mathcal{X}$, and let $\Phi \in \mathrm{T}(\mathcal{X})$ be a Hermitian-preserving, trace-preserving, and unital map satisfying*

$$\|J(\Omega) - J(\Phi)\| \leq \frac{1}{n(n^2-1)}. \tag{7.251}$$

It holds that Φ is a mixed-unitary channel.

Proof Define a map $\Psi \in \mathrm{T}(\mathcal{X})$ as

$$\Psi = (n^2-1)\Phi - (n^2-2)\Omega. \tag{7.252}$$

It holds that Ψ is trace preserving and unital. Moreover, one has

$$\begin{aligned} J(\Psi) &= (n^2 - 1)(J(\Phi) - J(\Omega)) + J(\Omega) \\ &= (n^2 - 1)(J(\Phi) - J(\Omega)) + \frac{1}{n}\mathbb{1}_{\mathcal{X}\otimes\mathcal{X}}, \end{aligned} \tag{7.253}$$

which, by the assumptions of the corollary, implies that Ψ is completely positive. By Theorem 7.31 it follows that

$$\frac{n^2 - 2}{n^2 - 1}\Omega + \frac{1}{n^2 - 1}\Psi = \Phi \tag{7.254}$$

is a mixed-unitary channel, which completes the proof. □

7.3 Measure concentration and it applications

The unitarily invariant measures introduced in the previous section exhibit a phenomenon known as *measure concentration*.[4] For the uniform spherical measure μ defined on the unit sphere of a complex Euclidean space $\mathcal{X}$, this phenomenon is reflected by the fact that, for every Lipschitz continuous function $f : \mathcal{S}(\mathcal{X}) \to \mathbb{R}$, the subset of $\mathcal{S}(\mathcal{X})$ on which f differs significantly from its average value (or, alternatively, any of its median values) must have relatively small measure. This phenomenon becomes more and more pronounced as the dimension of $\mathcal{X}$ grows.

Measure concentration is particularly useful in the theory of quantum information when used in the context of the *probabilistic method*. Various objects of interest, such as channels possessing certain properties, may be shown to exist by considering random choices of these object (typically based on the uniform spherical measure or Haar measure), followed by an analysis that demonstrates that the randomly chosen object possesses the property of interest with a nonzero probability. This method has been used successfully to demonstrate the existence of several interesting classes of objects for which explicit constructions are not known.

The present section explains this methodology, with its primary goal being to prove that the minimum output entropy of quantum channels is non-additive. Toward this goal, concentration bounds are established for uniform spherical measures, leading to an asymptotically strong form of a theorem known as *Dvoretzky's theorem*.

[4] Measure concentration is not limited to the measures introduced in the previous section—it is a more general phenomenon. For the purposes of this book, however, it will suffice to consider measure concentration with respect to those particular measures.

7.3.1 Lévy's lemma and Dvoretzky's theorem

This subsection establishes facts concerning the concentration of measure phenomenon mentioned previously, for the measures defined in the previous section. A selection of bounds will be presented, mainly targeted toward a proof of Dvoretzky's theorem, which concerns the existence of a relatively large subspace $\mathcal{V}$ of a given complex Euclidean space $\mathcal{X}$ on which a given Lipschitz function $f : \mathcal{S}(\mathcal{X}) \to \mathbb{R}$ does not deviate significantly from its mean or median values with respect to the uniform spherical measure.

Concentration bounds for Gaussian measure

In order to prove concentration bounds for the uniform spherical measure, with respect to a given complex Euclidean space $\mathcal{X}$, it is helpful to begin by proving an analogous result for the standard Gaussian measure on $\mathbb{R}^n$. Theorem 7.33, which is stated and proved below, establishes a result of this form that serves as a starting point for the concentration bounds to follow.

In the statements of the theorems representing concentration bounds to be presented below, including Theorem 7.33, it will be necessary to refer to certain universal real number constants. Such constants will, as a general convention, be denoted δ, δ_1, δ_2, etc., and must be chosen to be sufficiently small for the various theorems to hold. Although the optimization of these absolute constants should not be seen as being necessarily uninteresting or unimportant, this goal will be considered as being secondary in this book. Suitable values for these constants will be given in each case, but in some cases these values have been selected to simplify expressions and proofs rather than to optimize their values.

Theorem 7.33 *There exists a positive real number $\delta_1 > 0$ for which the following holds. For every choice of a positive integer n, independent and identically distributed standard normal random variables $X_1, \ldots, X_n$, a κ-Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$, and a positive real number $\varepsilon > 0$, it holds that*

$$\Pr\bigl(f(X_1, \ldots, X_n) - \mathrm{E}(f(X_1, \ldots, X_n)) \geq \varepsilon\bigr) \leq \exp\left(-\frac{\delta_1 \varepsilon^2}{\kappa^2}\right). \tag{7.255}$$

Remark One may take $\delta_1 = 2/\pi^2$.

The proof of Theorem 7.33 will make use of the two lemmas that follow. The first lemma is a fairly standard smoothing argument that will allow for basic multivariate calculus to be applied in the proof of the theorem.

Lemma 7.34 *Let n be a positive integer, let $f : \mathbb{R}^n \to \mathbb{R}$ be a κ-Lipschitz function, and let $\varepsilon > 0$ be a positive real number. There exists a differentiable κ-Lipschitz function $g : \mathbb{R}^n \to \mathbb{R}$ such that $|f(x) - g(x)| \leq \varepsilon$ for every $x \in \mathbb{R}^n$.*

Proof For every $\delta > 0$, define a function $g_\delta : \mathbb{R}^n \to \mathbb{R}$ as

$$g_\delta(x) = \int f(x + \delta z) \, \mathrm{d}\gamma_n(z) \tag{7.256}$$

for all $x \in \mathbb{R}^n$, where γ_n denotes the standard Gaussian measure on $\mathbb{R}^n$. It will be proved that setting $g = g_\delta$ for a suitable choice of δ satisfies the requirements of the lemma.

First, by the assumption that f is κ-Lipschitz, it holds that

$$\begin{aligned} |f(x) - g_\delta(x)| &\leq \int |f(x) - f(x + \delta z)| \, \mathrm{d}\gamma_n(z) \\ &\leq \delta\kappa \int \|z\| \, \mathrm{d}\gamma_n(z) \leq \delta\kappa\sqrt{n} \end{aligned} \tag{7.257}$$

for all $x \in \mathbb{R}^n$ and $\delta > 0$. The last inequality in (7.257) makes use of (1.279) in Chapter 1. At this point, one may fix

$$\delta = \frac{\varepsilon}{\kappa\sqrt{n}} \tag{7.258}$$

and $g = g_\delta$, so that $|f(x) - g(x)| \leq \varepsilon$ for every $x \in \mathbb{R}^n$.

Next, it holds that g is κ-Lipschitz, as the following calculation shows:

$$\begin{aligned} |g(x) - g(y)| &\leq \int |f(x + \delta z) - f(y + \delta z)| \, \mathrm{d}\gamma_n(z) \\ &\leq \int \kappa \|x - y\| \, \mathrm{d}\gamma_n(z) = \kappa\|x - y\|, \end{aligned} \tag{7.259}$$

for every $x, y \in \mathbb{R}^n$.

It remains to prove that g is differentiable. Using the definition of the standard Gaussian measure, one may calculate that the gradient of g at an arbitrary point $x \in \mathbb{R}^n$ is given by

$$\nabla g(x) = \frac{1}{\delta} \int f(x + \delta z) z \, \mathrm{d}\gamma_n(z). \tag{7.260}$$

The fact that the integral on the right-hand side of (7.260) exists follows from the inequality

$$\begin{aligned} &\int \| f(x + \delta z) z \| \, \mathrm{d}\gamma_n(z) \\ &\quad \leq \int \| f(x + \delta z) z - f(x) z \| \, \mathrm{d}\gamma_n(z) + \int \| f(x) z \| \, \mathrm{d}\gamma_n(z) \\ &\leq \kappa\delta \int \|z\|^2 \, \mathrm{d}\gamma_n(z) + |f(x)| \int \|z\| \, \mathrm{d}\gamma_n(z) \leq \kappa\delta n + |f(x)|\sqrt{n}. \end{aligned} \tag{7.261}$$

Moreover, it holds that $\nabla g(x)$ is a continuous function of x (and in fact is Lipschitz continuous), as

$$\begin{aligned} \|\nabla g(x) - \nabla g(y)\| &\leq \frac{1}{\delta} \int |f(x + \delta z) - f(y + \delta z)| \|z\| \, \mathrm{d}\gamma_n(z) \\ &\leq \frac{\kappa}{\delta} \|x - y\| \sqrt{n}. \end{aligned} \tag{7.262}$$

As $\nabla g(x)$ is a continuous function of x, it follows that g is differentiable, which completes the proof. □

The second lemma establishes that the random variable $f(X_1, \ldots, X_n)$, for independent and normally distributed random variables $X_1, \ldots, X_n$ and a differentiable κ-Lipschitz function f, does not deviate too much from an independent copy of itself.

Lemma 7.35 *Let n be a positive integer, let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function satisfying $\|\nabla f(x)\| \leq \kappa$ for every $x \in \mathbb{R}^n$, let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be independent and identically distributed standard normal random variables, and define vector-valued random variables*

$$X = (X_1, \ldots, X_n) \quad \text{and} \quad Y = (Y_1, \ldots, Y_n). \tag{7.263}$$

For every real number $\lambda \in \mathbb{R}$, it holds that

$$\mathrm{E}\big(\exp(\lambda f(X) - \lambda f(Y))\big) \leq \exp\left(\frac{\lambda^2 \pi^2 \kappa^2}{8}\right). \tag{7.264}$$

Proof First, define a function $g_{x,y} : \mathbb{R} \to \mathbb{R}$, for every choice of vectors $x, y \in \mathbb{R}^n$, as follows:

$$g_{x,y}(\theta) = f(\sin(\theta)x + \cos(\theta)y). \tag{7.265}$$

Applying the chain rule for differentiation, one finds that

$$g'_{x,y}(\theta) = \big\langle \nabla f(\sin(\theta)x + \cos(\theta)y), \cos(\theta)x - \sin(\theta)y \big\rangle \tag{7.266}$$

for every $x, y \in \mathbb{R}^n$ and $\theta \in \mathbb{R}$. By the fundamental theorem of calculus, it therefore follows that

$$\begin{aligned} f(x) - f(y) &= g_{x,y}(\pi/2) - g_{x,y}(0) = \int_0^{\frac{\pi}{2}} g'_{x,y}(\theta) \mathrm{d}\theta \\ &= \int_0^{\frac{\pi}{2}} \big\langle \nabla f(\sin(\theta)x + \cos(\theta)y), \cos(\theta)x - \sin(\theta)y \big\rangle \, \mathrm{d}\theta. \end{aligned} \tag{7.267}$$

Next, define a random variable Z_θ, for each $\theta \in [0, \pi/2]$, as

$$Z_\theta = \big\langle \nabla f(\sin(\theta)X + \cos(\theta)Y), \cos(\theta)X - \sin(\theta)Y \big\rangle. \tag{7.268}$$

By (7.267), it follows that

$$\mathrm{E}\big(\exp(\lambda f(X) - \lambda f(Y))\big) = \mathrm{E}\Big(\exp\Big(\lambda \int_0^{\frac{\pi}{2}} Z_\theta \,\mathrm{d}\theta\Big)\Big). \tag{7.269}$$

By Jensen's inequality, one has

$$\mathrm{E}\Big(\exp\Big(\lambda \int_0^{\frac{\pi}{2}} Z_\theta \,\mathrm{d}\theta\Big)\Big) \leq \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \mathrm{E}\Big(\exp\Big(\frac{\pi\lambda}{2} Z_\theta\Big)\Big) \,\mathrm{d}\theta. \tag{7.270}$$

Finally, one arrives at a key step of the proof: the observation that each of the random variables Z_θ is identically distributed, as a consequence of the invariance of Gaussian measure under orthogonal transformations. That is, one has the following equality of vector-valued random variables:

$$\begin{pmatrix} \sin(\theta)X + \cos(\theta)Y \\ \cos(\theta)X - \sin(\theta)Y \end{pmatrix} = \begin{pmatrix} \sin(\theta)\mathbb{1} & \cos(\theta)\mathbb{1} \\ \cos(\theta)\mathbb{1} & -\sin(\theta)\mathbb{1} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}. \tag{7.271}$$

As the distribution of $(X, Y) = (X_1, \ldots, X_n, Y_1, \ldots, Y_n)$ is invariant under orthogonal transformations, it follows that the distribution of Z_θ does not depend on θ. Consequently,

$$\frac{2}{\pi} \int_0^{\frac{\pi}{2}} \mathrm{E}\Big(\exp\Big(\frac{\pi\lambda}{2} Z_\theta\Big)\Big) \mathrm{d}\theta = \mathrm{E}\Big(\exp\Big(\frac{\pi\lambda}{2} Z_0\Big)\Big). \tag{7.272}$$

This quantity can be evaluated using the Gaussian integral equation (1.268), yielding

$$\mathrm{E}\Big(\exp\Big(\frac{\pi\lambda}{2} Z_0\Big)\Big) = \mathrm{E}\Big(\exp\Big(\frac{\pi^2\lambda^2}{8} \|\nabla f(Y)\|^2\Big)\Big). \tag{7.273}$$

As it is to be assumed that $\|\nabla f(x)\| \leq \kappa$ for all $x \in \mathbb{R}^n$, the required bound is obtained as a result of (7.269), (7.270), (7.272), and (7.273). □

Proof of Theorem 7.33 Let X be a vector-valued random variable, defined as $X = (X_1, \ldots, X_n)$, and let $\lambda > 0$ be a positive real number to be specified shortly. By Markov's inequality, one has

$$\begin{aligned} &\Pr\big(f(X) - \mathrm{E}(f(X)) \geq \varepsilon\big) \\ &\quad = \Pr\big(\exp(\lambda f(X) - \lambda\,\mathrm{E}(f(X))) \geq \exp(\lambda\varepsilon)\big) \\ &\quad \leq \exp(-\lambda\varepsilon)\,\mathrm{E}\big(\exp(\lambda f(X) - \lambda\,\mathrm{E}(f(X)))\big). \end{aligned} \tag{7.274}$$

By introducing a new random variable $Y = (Y_1, \ldots, Y_n)$, which is to be independent and identically distributed to X, one finds that

$$\mathrm{E}\big(\exp(\lambda f(X) - \lambda\,\mathrm{E}(f(X)))\big) \leq \mathrm{E}\big(\exp(\lambda f(X) - \lambda f(Y))\big) \tag{7.275}$$

by Jensen's inequality. Combining the two previous inequalities yields

$$\Pr\big(f(X) - \mathrm{E}(f(X)) \geq \varepsilon\big) \leq \exp(-\lambda\varepsilon)\,\mathrm{E}\big(\exp(\lambda f(X) - \lambda f(Y))\big). \quad (7.276)$$

Assume first that f is differentiable, so that $\|\nabla f(x)\| \leq \kappa$ for all $x \in \mathbb{R}^n$ by the assumption that f is κ-Lipschitz. By Lemma 7.35, it follows that

$$\exp(-\lambda\varepsilon)\,\mathrm{E}\big(\exp(\lambda f(X) - \lambda f(Y))\big) \leq \exp\biggl(-\lambda\varepsilon + \frac{\lambda^2\pi^2\kappa^2}{8}\biggr). \quad (7.277)$$

Setting $\lambda = 4\varepsilon/(\pi^2\kappa^2)$, and combining (7.276) with (7.277), yields

$$\Pr\big(f(X) - \mathrm{E}(f(X)) \geq \varepsilon\big) \leq \exp\biggl(-\frac{2\varepsilon^2}{\pi^2\kappa^2}\biggr), \quad (7.278)$$

which is the bound claimed in the statement of the theorem (for $\delta_1 = 2/\pi^2$).

Finally, suppose that f is κ-Lipschitz, but not necessarily differentiable. By Lemma 7.34, for every $\zeta \in (0, \varepsilon/2)$ there exists a differentiable κ-Lipschitz function $g : \mathbb{R}^n \to \mathbb{R}$ satisfying $|f(x) - g(x)| \leq \zeta$ for every $x \in \mathbb{R}^n$, and therefore

$$\Pr\big(f(X) - \mathrm{E}(f(X)) \geq \varepsilon\big) \leq \Pr\big(g(X) - \mathrm{E}(g(X)) \geq \varepsilon - 2\zeta\big). \quad (7.279)$$

Applying the above analysis to g in place of f therefore yields

$$\Pr\big(f(X) - \mathrm{E}(f(X)) \geq \varepsilon\big) \leq \exp\biggl(-\frac{2(\varepsilon - 2\zeta)^2}{\pi^2\kappa^2}\biggr). \quad (7.280)$$

As this inequality holds for every $\zeta \in (0, \varepsilon/2)$, the theorem follows. □

The following example illustrates the application of Theorem 7.33 to the Euclidean norm. The analysis to be presented in this example is relevant to the discussion of the uniform spherical measure to be discussed shortly.

Example 7.36 Let n be a positive integer and define $f(x) = \|x\|$ for each $x \in \mathbb{R}^n$. It is an immediate consequence of the triangle inequality that f is 1-Lipschitz:

$$\big|f(x) - f(y)\big| = \big|\|x\| - \|y\|\big| \leq \|x - y\| \quad (7.281)$$

for all $x, y \in \mathbb{R}^n$. The mean value of $f(X_1, \ldots, X_n)$, for $X_1, \ldots, X_n$ being independent and identically distributed standard normal random variables, has the following closed-form expression (q.v. Section 1.2.2):

$$\mathrm{E}\big(f(X_1, \ldots, X_n)\big) = \frac{\sqrt{2}\,\Gamma\bigl(\frac{n+1}{2}\bigr)}{\Gamma\bigl(\frac{n}{2}\bigr)}. \quad (7.282)$$

From this expression, an analysis reveals that

$$\mathrm{E}\big(f(X_1,\ldots,X_n)\big) = \upsilon_n\sqrt{n}, \tag{7.283}$$

where $\upsilon_1, \upsilon_2, \upsilon_3, \ldots$ is a strictly increasing sequence that begins

$$\upsilon_1 = \sqrt{\frac{2}{\pi}}, \quad \upsilon_2 = \frac{\sqrt{\pi}}{2}, \quad \upsilon_3 = \sqrt{\frac{8}{3\pi}}, \quad \ldots \tag{7.284}$$

and converges to 1 in the limit as n goes to infinity.

For any positive real number $\varepsilon > 0$, one may conclude the following two bounds from Theorem 7.33:

$$\begin{aligned} \Pr\big(\big\|(X_1,\ldots,X_n)\big\| \leq (\nu_n - \varepsilon)\sqrt{n}\big) &\leq \exp(-\delta_1\varepsilon^2 n), \\ \Pr\big(\big\|(X_1,\ldots,X_n)\big\| \geq (\nu_n + \varepsilon)\sqrt{n}\big) &\leq \exp(-\delta_1\varepsilon^2 n). \end{aligned} \tag{7.285}$$

Consequently, one has

$$\Pr\big(\big|\big\|(X_1,\ldots,X_n)\big\| - \nu_n\sqrt{n}\big| \geq \varepsilon\sqrt{n}\big) \leq 2\exp(-\delta_1\varepsilon^2 n). \tag{7.286}$$

This bound illustrates that the Euclidean norm of a Gaussian-random vector $x \in \mathbb{R}^n$ is tightly concentrated around its mean value $\upsilon_n\sqrt{n}$.

Concentration bounds for uniform spherical measure

The uniform spherical measure may be derived from the standard Gaussian measure, as described in Section 7.2.1, so it is not unreasonable to expect that Theorem 7.33 might lead to an analogous fact holding for the uniform spherical measure. Indeed this is the case, as the theorems below establish.

The first theorem concerns the deviation of a Lipschitz random variable, defined with respect to the uniform spherical measure, from its mean value.

Theorem 7.37 (Lévy's lemma, mean value form) *There exists a positive real number $\delta_2 > 0$ for which the following holds. For every κ-Lipschitz random variable $X : \mathcal{S}(\mathcal{X}) \to \mathbb{R}$, distributed with respect to the uniform spherical measure μ on $\mathcal{S}(\mathcal{X})$ for a given complex Euclidean space $\mathcal{X}$, and every positive real number $\varepsilon > 0$, it holds that*

$$\begin{aligned} \Pr\big(X - \mathrm{E}(X) \geq \varepsilon\big) &\leq 2\exp\bigg(-\frac{\delta_2\varepsilon^2 n}{\kappa^2}\bigg), \\ \Pr\big(X - \mathrm{E}(X) \leq -\varepsilon\big) &\leq 2\exp\bigg(-\frac{\delta_2\varepsilon^2 n}{\kappa^2}\bigg), \end{aligned} \tag{7.287}$$

and

$$\Pr\big(|X - \mathrm{E}(X)| \geq \varepsilon\big) \leq 3\exp\bigg(-\frac{\delta_2\varepsilon^2 n}{\kappa^2}\bigg), \tag{7.288}$$

where $n = \dim(\mathcal{X})$.

Remark One may take $\delta_2 = 1/(25\pi)$.

The proof of Lemma 7.37 will make use of the following lemma, which provides a simple mechanism for extending a Lipschitz function defined on the unit sphere of $\mathbb{C}^n$ to a Lipschitz function defined on all of $\mathbb{R}^{2n}$.

Lemma 7.38 *Let n be a positive integer and let $f : \mathcal{S}(\mathbb{C}^n) \to \mathbb{R}$ be a κ-Lipschitz function that is neither strictly positive nor strictly negative. Define a function $g : \mathbb{R}^{2n} \to \mathbb{R}$ as*

$$g(x \oplus y) = \begin{cases} \|x + iy\| f\left(\frac{x+iy}{\|x+iy\|}\right) & \text{if } x + iy \neq 0 \\ 0 & \text{if } x + iy = 0 \end{cases} \tag{7.289}$$

for all $x, y \in \mathbb{R}^n$. It holds that g is a (3κ)-Lipschitz function.

Proof By the assumption that f is neither strictly positive nor strictly negative, one has that for every unit vector $u \in \mathbb{C}^n$, there must exist a unit vector $v \in \mathbb{C}^n$ such that $f(u)f(v) \leq 0$. This in turn implies

$$|f(u)| \leq |f(u) - f(v)| \leq \kappa \|u - v\| \leq 2\kappa, \tag{7.290}$$

by the assumption that f is κ-Lipschitz.

Now suppose that $x_0, y_0, x_1, y_1 \in \mathbb{R}^n$ are vectors. If it is the case that $x_0 + iy_0 = 0$ and $x_1 + iy_1 = 0$, then it is immediate that

$$|g(x_0 \oplus y_0) - g(x_1 \oplus y_1)| = 0. \tag{7.291}$$

If it holds that $x_0 + iy_0 \neq 0$ and $x_1 + iy_1 = 0$, then (7.290) implies

$$|g(x_0 \oplus y_0) - g(x_1 \oplus y_1)| = |g(x_0 \oplus y_0)| \leq 2\kappa \|x_0 + iy_0\| = 2\kappa \|x_0 \oplus y_0\|. \tag{7.292}$$

A similar bound holds for the case in which $x_0 + iy_0 = 0$ and $x_1 + iy_1 \neq 0$.

Finally, suppose that $x_0 + iy_0$ and $x_1 + iy_1$ are both nonzero. Write

$$z_0 = x_0 + iy_0 \quad \text{and} \quad z_1 = x_1 + iy_1, \tag{7.293}$$

and set

$$\alpha_0 = \frac{1}{\|z_0\|} \quad \text{and} \quad \alpha_1 = \frac{1}{\|z_1\|}. \tag{7.294}$$

This implies that both $\alpha_0 z_0$ and $\alpha_1 z_1$ are unit vectors. There is no loss of generality in assuming $\alpha_0 \leq \alpha_1$; the case in which $\alpha_1 \leq \alpha_0$ is handled in a symmetric manner. By the triangle inequality, one has

$$\begin{aligned} |g(x_0 \oplus y_0) - g(x_1 \oplus y_1)| &= |\|z_0\| f(\alpha_0 z_0) - \|z_1\| f(\alpha_1 z_1)| \\ &\leq |f(\alpha_0 z_0)| \|z_0 - z_1\| + \|z_1\| |f(\alpha_0 z_0) - f(\alpha_1 z_1)|. \end{aligned} \tag{7.295}$$

Using (7.290), one finds that the first term in the final expression of (7.295) is bounded as follows:

$$|f(\alpha_0 z_0)|\,\|z_0 - z_1\| \leq 2\kappa\|z_0 - z_1\| = 2\kappa\|x_0 \oplus y_0 - x_1 \oplus y_1\|. \tag{7.296}$$

To bound the second term, it may first be noted that

$$\|z_1\|\,|f(\alpha_0 z_0) - f(\alpha_1 z_1)| \leq \kappa\|z_1\|\,\|\alpha_0 z_0 - \alpha_1 z_1\|, \tag{7.297}$$

again by the assumption that f is κ-Lipschitz. Given that $0 < \alpha_0 \leq \alpha_1$, together with the fact that $\alpha_0 z_0$ and $\alpha_1 z_1$ are unit vectors, one finds that

$$\|\alpha_0 z_0 - \alpha_1 z_1\| \leq \|\alpha_1 z_0 - \alpha_1 z_1\| = \frac{\|z_0 - z_1\|}{\|z_1\|}, \tag{7.298}$$

and therefore

$$\kappa\|z_1\|\,\|\alpha_0 z_0 - \alpha_1 z_1\| \leq \kappa\|z_0 - z_1\| = \kappa\|x_0 \oplus y_0 - x_1 \oplus y_1\|. \tag{7.299}$$

It follows that

$$|g(x_0 \oplus y_0) - g(x_1 \oplus y_1)| \leq 3\kappa\|x_0 \oplus y_0 - x_1 \oplus y_1\|. \tag{7.300}$$

It has therefore been established that g is (3κ)-Lipschitz, as required. □

Proof of Theorem 7.37 The random variable $X - \mathrm{E}(X)$ has mean value 0, and is therefore neither strictly positive nor strictly negative. As X is κ-Lipschitz, so too is $X - \mathrm{E}(X)$, and so it follows that

$$\big|X - \mathrm{E}(X)\big| \leq 2\kappa, \tag{7.301}$$

as argued in the first paragraph of the proof of Lemma 7.38. The inequalities (7.287) and (7.288) therefore hold trivially when $\varepsilon > 2\kappa$. For this reason it will be assumed that $\varepsilon \leq 2\kappa$ for the remainder of the proof. It will also be assumed that $\mathcal{X} = \mathbb{C}^n$, for n being an arbitrary positive integer, which will simplify the notation used throughout the proof, and which causes no loss of generality.

Define a function $g : \mathbb{R}^{2n} \to \mathbb{R}$ as

$$g(y \oplus z) = \begin{cases} \|y + iz\|\Big(X\Big(\frac{y+iz}{\|y+iz\|}\Big) - \mathrm{E}(X)\Big) & \text{if } y + iz \neq 0 \\ 0 & \text{if } y + iz = 0 \end{cases} \tag{7.302}$$

for all $y, z \in \mathbb{R}^n$, which is a (3κ)-Lipschitz function by Lemma 7.38. Let $Y = (Y_1, \ldots, Y_n)$ and $Z = (Z_1, \ldots, Z_n)$ be vector-valued random variables, for $Y_1, \ldots, Y_n$ and $Z_1, \ldots, Z_n$ being independent and identically distributed standard normal random variables, and define a random variable

$$W = g(Y \oplus Z). \tag{7.303}$$

As $X - \mathrm{E}(X)$ has mean value 0, it is evident that $\mathrm{E}(W) = 0$ as well. Finally, by considering the definition of the uniform spherical measure, one finds that

$$\Pr(X - \mathrm{E}(X) \geq \varepsilon) = \Pr\bigl(W \geq \varepsilon \|Y + iZ\|\bigr). \tag{7.304}$$

The probability (7.304) may be upper-bounded through the use of the union bound:

$$\Pr(X - \mathrm{E}(X) \geq \varepsilon) \leq \Pr\Bigl(W \geq \varepsilon\lambda\sqrt{2n}\Bigr) + \Pr\Bigl(\|Y + iZ\| \leq \lambda\sqrt{2n}\Bigr) \tag{7.305}$$

for every choice of $\lambda > 0$. By Theorem 7.33 it holds that

$$\Pr\Bigl(W \geq \varepsilon\lambda\sqrt{2n}\Bigr) \leq \exp\left(-\frac{2\delta_1\varepsilon^2\lambda^2 n}{9\kappa^2}\right), \tag{7.306}$$

and, as established in Example 7.36, it holds that

$$\Pr\Bigl(\|Y + iZ\| \leq \lambda\sqrt{2n}\Bigr) \leq \exp\Bigl(-2\delta_1(\upsilon_{2n} - \lambda)^2 n\Bigr). \tag{7.307}$$

Setting

$$\lambda = \frac{3\kappa\upsilon_{2n}}{3\kappa + \varepsilon} \tag{7.308}$$

yields

$$\Pr(X \geq \mathrm{E}(X) + \varepsilon) \leq 2\exp\left(-\frac{2\delta_1\varepsilon^2\upsilon_{2n}^2 n}{(3\kappa + \varepsilon)^2}\right) \leq 2\exp\left(-\frac{\delta_1\pi\varepsilon^2 n}{50\kappa^2}\right), \tag{7.309}$$

where the second inequality makes use of the assumption $\varepsilon \leq 2\kappa$, along with the observation that $\upsilon_{2n} \geq \upsilon_2 = \sqrt{\pi}/2$. As one may take $\delta_1 = 2/\pi^2$ in Theorem 7.33, the first inequality is therefore proved for $\delta_2 = 1/(25\pi)$.

The second and third inequalities may be proved in essentially the same manner. In particular, one has

$$\begin{aligned} &\Pr(X - \mathrm{E}(X) \leq -\varepsilon) \\ &\quad \leq \Pr\Bigl(W \leq -\varepsilon\lambda\sqrt{2n}\Bigr) + \Pr\Bigl(\|Y + iZ\| \leq \lambda\sqrt{2n}\Bigr) \end{aligned} \tag{7.310}$$

and

$$\begin{aligned} &\Pr(|X - \mathrm{E}(X)| \geq \varepsilon) \\ &\qquad \leq \Pr\Bigl(W \geq \varepsilon\lambda\sqrt{2n}\Bigr) + \Pr\Bigl(W \leq -\varepsilon\lambda\sqrt{2n}\Bigr) \\ &\qquad\quad + \Pr\Bigl(\|Y + iZ\| \leq \lambda\sqrt{2n}\Bigr), \end{aligned} \tag{7.311}$$

and again setting $\lambda = 3\kappa\upsilon_{2n}/(3\kappa + \varepsilon)$ yields the required bounds. □

The second theorem on measure concentration for the uniform spherical measure, stated and proved below, is similar in spirit to Theorem 7.37, but it is concerned with the deviation of a Lipschitz random variable from its *median value*—or, more generally, from any of its *central values*—rather than its mean value. The next definition makes precise the notions of a median value and a central value of a random variable, after which the theorem is stated and proved.

Definition 7.39 Let X be a random variable and let β be a real number. It is said that β is a *median value* of X if

$$\Pr(X \geq \beta) \geq \frac{1}{2} \quad \text{and} \quad \Pr(X \leq \beta) \geq \frac{1}{2}, \tag{7.312}$$

and it is said that β is a *central value* of X if

$$\Pr(X \geq \beta) \geq \frac{1}{4} \quad \text{and} \quad \Pr(X \leq \beta) \geq \frac{1}{4}. \tag{7.313}$$

Theorem 7.40 (Lévy's lemma, central value form) *There exists a positive real number $\delta_3 > 0$ for which the following holds. For every complex Euclidean space $\mathcal{X}$, every κ-Lipschitz random variable*

$$X : \mathcal{S}(\mathcal{X}) \to \mathbb{R}, \tag{7.314}$$

distributed with respect to the uniform spherical measure μ on $\mathcal{S}(\mathcal{X})$, every central value β of X, and every positive real number $\varepsilon > 0$, it holds that

$$\Pr\bigl(|X - \beta| \geq \varepsilon\bigr) \leq 8 \exp\biggl(-\frac{\delta_3 \varepsilon^2 n}{\kappa^2}\biggr), \tag{7.315}$$

where $n = \dim(\mathcal{X})$.

Remark One may take $\delta_3 = 1/(100\pi)$.

Proof Let

$$\zeta = \sqrt{\frac{\ln(8)\kappa^2}{\delta_2 n}}, \tag{7.316}$$

for δ_2 being any positive real number for which Theorem 7.37 holds. By that theorem, one may conclude that the following two inequalities hold for every positive real number $\alpha > 0$:

$$\Pr\bigl(X - \mathrm{E}(X) \geq \zeta + \alpha\bigr) \leq 2 \exp\biggl(-\frac{\delta_2 (\zeta + \alpha)^2 n}{\kappa^2}\biggr) < \frac{1}{4}, \tag{7.317}$$

$$\Pr\bigl(X - \mathrm{E}(X) \leq -(\zeta + \alpha)\bigr) \leq 2 \exp\biggl(-\frac{\delta_2 (\zeta + \alpha)^2 n}{\kappa^2}\biggr) < \frac{1}{4}. \tag{7.318}$$

From these inequalities, one concludes that $|\mathrm{E}(X) - \beta| \leq \zeta$.

Now suppose that ε is a given positive real number. If it is the case that $\varepsilon \geq 2\zeta$, then Theorem 7.37 implies

$$\begin{aligned} \Pr\bigl(|X - \beta| \geq \varepsilon\bigr) &\leq \Pr\bigl(|X - \mathrm{E}(X)| \geq \varepsilon - \zeta\bigr) \\ &\leq \Pr\Bigl(|X - \mathrm{E}(X)| \geq \frac{\varepsilon}{2}\Bigr) \leq 3\exp\Bigl(-\frac{\delta_2 \varepsilon^2 n}{4\kappa^2}\Bigr). \end{aligned} \tag{7.319}$$

On the other hand, if $\varepsilon < 2\zeta$, then one has

$$\exp\Bigl(-\frac{\delta_2 \varepsilon^2 n}{4\kappa^2}\Bigr) > \exp\Bigl(-\frac{\delta_2 \zeta^2 n}{\kappa^2}\Bigr) = \frac{1}{8}, \tag{7.320}$$

so it must trivially hold that

$$\Pr\bigl(|X - \beta| \geq \varepsilon\bigr) \leq 8\exp\Bigl(-\frac{\delta_2 \varepsilon^2 n}{4\kappa^2}\Bigr). \tag{7.321}$$

The required bound (7.315) therefore holds in both cases, provided one takes $\delta_3 \leq \delta_2/4$. As Theorem 7.37 holds for $\delta_2 = 1/(25\pi)$, the bound (7.315) holds for $\delta_3 = 1/(100\pi)$. □

Dvoretzky's theorem

Dvoretzky's theorem, which plays a key role in the section following this one, establishes that a Lipschitz random variable, defined with respect to the uniform spherical measure for a given complex Euclidean space $\mathcal{X}$, must remain close to its central values everywhere on the unit sphere $\mathcal{S}(\mathcal{V})$, for some choice of a subspace $\mathcal{V} \subseteq \mathcal{X}$ having relatively large dimension. There are, in fact, multiple variants and generalizations of Dvoretzky's theorem; the variant to be considered in this book is specific to the unitary invariant measures defined previously in the present chapter, and is applicable to *phase-invariant* functions, which are defined as follows.

Definition 7.41 Let $f : \mathcal{S}(\mathcal{X}) \to \mathbb{R}$ be a function, for a complex Euclidean space $\mathcal{X}$. The function f is said to be a *phase-invariant* function if it holds that $f(x) = f(e^{i\theta}x)$ for all $x \in \mathcal{S}(\mathcal{X})$ and $\theta \in \mathbb{R}$.

Theorem 7.42 (Dvoretzky's theorem) *There exists a positive real number $\delta > 0$ for which the following holds. Let $X : \mathcal{S}(\mathcal{X}) \to \mathbb{R}$ be a κ-Lipschitz, phase-invariant random variable, distributed with respect to the uniform spherical measure μ on $\mathcal{S}(\mathcal{X})$ for a given complex Euclidean space $\mathcal{X}$ of dimension n, let β be a central value of X, let $\varepsilon > 0$ and $\zeta > 0$ be positive real numbers, and let $\mathcal{V} \subseteq \mathcal{X}$ be a subspace with*

$$1 \leq \dim(\mathcal{V}) \leq \frac{\delta \varepsilon^2 \zeta^2 n}{\kappa^2}. \tag{7.322}$$

For each unit vector $v \in \mathcal{V}$, define a random variable $Y_v : \mathrm{U}(\mathcal{X}) \to \mathbb{R}$, distributed with respect to the Haar measure on $\mathrm{U}(\mathcal{X})$, as

$$Y_v(U) = X(Uv) \tag{7.323}$$

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

$$\Pr\big(|Y_v - \beta| \leq \varepsilon \text{ for every } v \in \mathcal{S}(\mathcal{V})\big) \geq 1 - \zeta. \tag{7.324}$$

Remark One may take $\delta = 1/(160000\pi)$.

The proof of Theorem 7.42 will make use of the two lemmas that follow.

Lemma 7.43 *Let $\mathcal{X}$ be a complex Euclidean space of dimension $n \geq 2$ and let $f : \mathcal{S}(\mathcal{X}) \to \mathbb{R}$ be a κ-Lipschitz, phase-invariant function. For every unit vector $u \in \mathcal{S}(\mathcal{X})$, define a random variable $X_u : \mathrm{U}(\mathcal{X}) \to \mathbb{R}$, distributed with respect to the Haar measure η on $\mathrm{U}(\mathcal{X})$, as*

$$X_u(U) = f(Uu) \tag{7.325}$$

for all $U \in \mathrm{U}(\mathcal{X})$. For any pair of linearly independent unit vectors $u, v \in \mathcal{X}$ and every positive real number $\varepsilon > 0$, it holds that

$$\Pr\big(|X_u - X_v| \geq \varepsilon\big) \leq 3 \exp\biggl(-\frac{\delta_2 \varepsilon^2 (n-1)}{\kappa^2 \|u - v\|^2}\biggr), \tag{7.326}$$

for any positive real number δ_2 satisfying the requirements of Theorem 7.37.

Proof The lemma will first be proved in the special case in which $\langle u, v \rangle$ is a nonnegative real number. First, define

$$\lambda = \frac{1 + \langle u, v \rangle}{2}, \tag{7.327}$$

which satisfies $1/2 \leq \lambda < 1$ by the assumption that $\langle u, v \rangle$ is nonnegative and u and v are linearly independent. Set

$$x = \frac{u + v}{2\sqrt{\lambda}} \quad \text{and} \quad y = \frac{u - v}{2\sqrt{1 - \lambda}}, \tag{7.328}$$

so that x and y are orthonormal unit vectors for which

$$\begin{aligned} u &= \sqrt{\lambda} x + \sqrt{1 - \lambda} y, \\ v &= \sqrt{\lambda} x - \sqrt{1 - \lambda} y. \end{aligned} \tag{7.329}$$

Next, let $\mathcal{Y}$ be any complex Euclidean space having dimension $n - 1$ and let $V \in \mathrm{U}(\mathcal{Y}, \mathcal{X})$ be any isometry for which $x \perp \mathrm{im}(V)$. For every $U \in \mathrm{U}(\mathcal{X})$,

define a random variable $Y_U : \mathcal{S}(\mathcal{Y}) \to \mathbb{R}$, distributed with respect to the uniform spherical measure μ on $\mathcal{S}(\mathcal{Y})$, as

$$Y_U(w) = f\Big(U\Big(\sqrt{\lambda}x + \sqrt{1-\lambda}Vw\Big)\Big) - f\Big(U\Big(\sqrt{\lambda}x - \sqrt{1-\lambda}Vw\Big)\Big) \tag{7.330}$$

for every $w \in \mathcal{S}(\mathcal{Y})$. Using the triangle inequality, along with the fact that

$$\|u - v\| = 2\sqrt{1-\lambda}, \tag{7.331}$$

one may verify that each Y_U is $(\kappa\|u-v\|)$-Lipschitz and satisfies $\mathrm{E}(Y_U) = 0$. By Lévy's lemma (Theorem 7.37), it therefore holds that

$$\Pr\big(|Y_U| \geq \varepsilon\big) \leq 3\exp\Big(-\frac{\delta_2 \varepsilon^2 (n-1)}{\kappa^2 \|u-v\|^2}\Big), \tag{7.332}$$

for every $U \in \mathrm{U}(\mathcal{X})$ and every $\varepsilon > 0$.

Finally, define a random variable $Z : \mathrm{U}(\mathcal{X}) \times \mathcal{S}(\mathcal{Y}) \to \mathbb{R}$, distributed with respect to the product measure $\eta \times \mu$, as

$$Z(U, w) = Y_U(w) \tag{7.333}$$

for all $U \in \mathrm{U}(\mathcal{X})$ and $w \in \mathcal{S}(\mathcal{Y})$. Because the uniform spherical measure and Haar measure are both unitary invariant, it follows that Z and $X_u - X_v$ are identically distributed. It therefore holds that

$$\begin{aligned} \Pr\big(|X_u - X_v| \geq \varepsilon\big) &= \Pr\big(|Z| \geq \varepsilon\big) \\ &= \int \Pr\big(|Y_U| \geq \varepsilon\big)\, \mathrm{d}\eta(U) \leq 3\exp\Big(-\frac{\delta_2 \varepsilon^2 (n-1)}{\kappa^2 \|u-v\|^2}\Big), \end{aligned} \tag{7.334}$$

which proves the lemma in the case that $\langle u, v\rangle$ is a nonnegative real number.

In the situation in which $\langle u, v\rangle$ is not a nonnegative real number, one may choose $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ so that $\langle u, \alpha v\rangle$ is a nonnegative real number. By the assumption that f is phase invariant, it holds that $X_v = X_{\alpha v}$, and therefore

$$\begin{aligned} \Pr\big(|X_u - X_v| \geq \varepsilon\big) &= \Pr\big(|X_u - X_{\alpha v}| \geq \varepsilon\big) \\ &\leq 3\exp\Big(-\frac{\delta_2 \varepsilon^2 (n-1)}{\kappa^2 \|u-\alpha v\|^2}\Big), \end{aligned} \tag{7.335}$$

by the analysis above. As it necessarily holds that $\|u - \alpha v\| \leq \|u - v\|$, it follows that

$$\Pr\big(|X_u - X_v| \geq \varepsilon\big) \leq 3\exp\Big(-\frac{\delta_2 \varepsilon^2 (n-1)}{\kappa^2 \|u-v\|^2}\Big) \tag{7.336}$$

for every $\varepsilon > 0$, which completes the proof. □

The next lemma bounds the mean value of the maximum of a collection of nonnegative random variables satisfying a property reminiscent of the bounds obtained for the concentration results presented above.

Lemma 7.44 *Let $N \geq 2$ be a positive integer, let K and θ be positive real numbers, and let $Y_1, \ldots, Y_N$ be nonnegative random variables for which*

$$\Pr(Y_k \geq \lambda) \leq K \exp(-\theta \lambda^2) \tag{7.337}$$

for every $k \in \{1, \ldots, N\}$ and every $\lambda \geq 0$. It holds that

$$\mathrm{E}(\max\{Y_1, \ldots, Y_N\}) \leq \sqrt{\frac{\ln(N)}{\theta}} + \frac{K}{\sqrt{2\theta}}. \tag{7.338}$$

Proof As the random variables $Y_1, \ldots, Y_N$ take only nonnegative values, one may write

$$\mathrm{E}(\max\{Y_1, \ldots, Y_N\}) = \int_0^\infty \Pr(\max\{Y_1, \ldots, Y_N\} \geq \lambda) \, \mathrm{d}\lambda. \tag{7.339}$$

Splitting the integral into two parts, and making use of the fact that the probability of any event is at most 1, yields

$$\begin{aligned} &\mathrm{E}(\max\{Y_1, \ldots, Y_N\}) \\ &\quad \leq \sqrt{\frac{\ln(N)}{\theta}} + \int_{\sqrt{\frac{\ln(N)}{\theta}}}^\infty \Pr(\max\{Y_1, \ldots, Y_N\} \geq \lambda) \, \mathrm{d}\lambda. \end{aligned} \tag{7.340}$$

By the union bound, together with the assumption (7.337) on $Y_1, \ldots, Y_N$, one has

$$\int_{\sqrt{\frac{\ln(N)}{\theta}}}^\infty \Pr(\max\{Y_1, \ldots, Y_N\} \geq \lambda) \, \mathrm{d}\lambda \leq KN \int_{\sqrt{\frac{\ln(N)}{\theta}}}^\infty \exp(-\theta\lambda^2) \, \mathrm{d}\lambda. \tag{7.341}$$

As $\ln(2) > 1/2$, it holds that $\lambda\sqrt{2\theta} > 1$ for every choice of λ satisfying

$$\lambda \geq \sqrt{\frac{\ln(N)}{\theta}}, \tag{7.342}$$

and therefore

$$\int_{\sqrt{\frac{\ln(N)}{\theta}}}^\infty \exp(-\theta\lambda^2) \, \mathrm{d}\lambda \leq \int_{\sqrt{\frac{\ln(N)}{\theta}}}^\infty \lambda\sqrt{2\theta} \exp(-\theta\lambda^2) \, \mathrm{d}\lambda = \frac{1}{N\sqrt{2\theta}}. \tag{7.343}$$

The required inequality now follows from (7.340), (7.341), and (7.343). □

Proof of Theorem 7.42 It will be proved that any choice of $\delta > 0$ satisfying

$$\delta \leq \left(\frac{8}{\sqrt{\delta_3}} + \frac{64}{\sqrt{\delta_2}} \right)^{-2}, \tag{7.344}$$

for δ_2 and δ_3 being positive real numbers that satisfy the requirements of Theorem 7.37 and Theorem 7.40, respectively, fulfills the requirements of the theorem. Taking $\delta_2 = 1/(25\pi)$ and $\delta_3 = 1/(100\pi)$, one has that

$$\delta = \frac{1}{160000\pi} \tag{7.345}$$

satisfies the requirement (7.344). The theorem is trivial in the case $n = 1$, as the phase invariance of X implies that X is constant in this case, and for this reason it will be assumed that $n \geq 2$ for the remainder of the proof.

By Markov's inequality, one has

$$\begin{aligned} &\Pr\bigl(\sup\{|Y_v - \beta| \,:\, v \in \mathcal{S}(\mathcal{V})\} \leq \varepsilon\bigr) \\ &\qquad \geq 1 - \frac{\mathrm{E}\bigl(\sup\{|Y_v - \beta| \,:\, v \in \mathcal{S}(\mathcal{V})\}\bigr)}{\varepsilon}, \end{aligned} \tag{7.346}$$

so the theorem will follow from a demonstration that

$$\mathrm{E}\bigl(\sup\{|Y_v - \beta| \,:\, v \in \mathcal{S}(\mathcal{V})\}\bigr) \leq \zeta\varepsilon. \tag{7.347}$$

Let $m = \dim(\mathcal{V})$, and for each nonnegative integer $k \in \mathbb{N}$, let $\mathcal{N}_k$ be a minimal (2^{-k+1})-net for $\mathcal{S}(\mathcal{V})$. It is evident that $|\mathcal{N}_0| = 1$, and for every $k \in \mathbb{N}$ it holds that

$$|\mathcal{N}_k| \leq \bigl(1 + 2^k\bigr)^{2m} \leq 4^{(k+1)m} \tag{7.348}$$

by Theorem 1.8. For each $v \in \mathcal{S}(\mathcal{V})$ and $k \in \mathbb{N}$, fix $z_k(v) \in \mathcal{N}_k$ to be any element of the set $\mathcal{N}_k$ for which the distance to v is minimized, which implies that

$$\|v - z_k(v)\| \leq 2^{-k+1}. \tag{7.349}$$

One may observe that $z_0 = z_0(v)$ is independent of v, as there is a single element in the set $\mathcal{N}_0$, and also that

$$\lim_{k\to\infty} z_k(v) = v \tag{7.350}$$

for every $v \in \mathcal{S}(\mathcal{V})$.

Next, observe that

$$X(Uv) = X(Uz_0) + \sum_{k=0}^{\infty} \Bigl(X\bigl(Uz_{k+1}(v)\bigr) - X\bigl(Uz_k(v)\bigr) \Bigr), \tag{7.351}$$

for every $v \in \mathcal{S}(\mathcal{V})$ and $U \in \mathrm{U}(\mathcal{X})$; this fact may be verified by telescoping

the sum and making use of (7.350), along with the continuity of X. It follows that

$$Y_v = Y_{z_0} + \sum_{k=0}^{\infty} (Y_{z_{k+1}(v)} - Y_{z_k(v)}) \tag{7.352}$$

for every $v \in \mathcal{S}(\mathcal{V})$. By the triangle inequality, one therefore has

$$\begin{aligned} &\sup\{|Y_v - \beta| \,:\, v \in \mathcal{S}(\mathcal{V})\} \\ &\quad \leq |Y_{z_0} - \beta| + \sup\left\{\sum_{k=0}^{\infty} |Y_{z_{k+1}(v)} - Y_{z_k(v)}| \,:\, v \in \mathcal{S}(\mathcal{V})\right\}. \end{aligned} \tag{7.353}$$

The expected value of the two terms on the right-hand side of this inequality will be bounded separately.

The expected value of the first term $|Y_{z_0} - \beta|$ will be considered first. The random variable Y_{z_0} is identically distributed to X, so it follows by Theorem 7.40 that

$$\Pr(|Y_{z_0} - \beta| \geq \lambda) = \Pr(|X - \beta| \geq \lambda) \leq 8 \exp\left(-\frac{\delta_3 \lambda^2 n}{\kappa^2}\right) \tag{7.354}$$

for every $\lambda \geq 0$. This implies that

$$\begin{aligned} \mathrm{E}(|Y_{z_0} - \beta|) &= \int_0^{\infty} \Pr(|Y_{z_0} - \beta| \geq \lambda) \, \mathrm{d}\lambda \\ &\leq 8 \int_0^{\infty} \exp\left(-\frac{\delta_3 \lambda^2 n}{\kappa^2}\right) \mathrm{d}\lambda = 4\sqrt{\frac{\pi \kappa^2}{\delta_3 n}} < \frac{8\kappa}{\sqrt{\delta_3 n}}. \end{aligned} \tag{7.355}$$

It remains to bound the expected value of the second term on the right-hand side of (7.353). It holds that

$$\|z_{k+1}(v) - z_k(v)\| \leq \|z_{k+1}(v) - v\| + \|v - z_k(v)\| < 2^{-k+2} \tag{7.356}$$

for all $v \in \mathcal{S}(\mathcal{V})$ and all $k \in \mathbb{N}$, and therefore

$$\begin{aligned} &\sup\left\{\sum_{k=0}^{\infty} |Y_{z_{k+1}(v)} - Y_{z_k(v)}| \,:\, v \in \mathcal{S}(\mathcal{V})\right\} \\ &\qquad \leq \sum_{k=0}^{\infty} \max\Big\{|Y_x - Y_y| \,:\, (x, y) \in \mathcal{M}_k\Big\} \end{aligned} \tag{7.357}$$

where

$$\mathcal{M}_k = \Big\{(x, y) \in \mathcal{N}_{k+1} \times \mathcal{N}_k, \, \|x - y\| < 2^{-k+2}\Big\}. \tag{7.358}$$

By Lemma 7.43, it holds that

$$\Pr\Big(|Y_x - Y_y| \geq \varepsilon\Big) \leq 3\exp\bigg(-\frac{\delta_2 \varepsilon^2 (n-1)}{\kappa^2 \|x-y\|^2}\bigg) \tag{7.359}$$

for every pair of linearly independent vectors $x, y \in \mathcal{S}(\mathcal{V})$, for δ_2 being any positive real number for which Theorem 7.37 holds. (By the assumption that X is phase-invariant, one has $Y_x = Y_y$ if $x, y \in \mathcal{S}(\mathcal{V})$ are linearly dependent.) For each choice of $k \in \mathbb{N}$, it therefore follows from Lemma 7.44 that

$$\mathrm{E}\Big(\max\Big\{|Y_x - Y_y| \,:\, (x,y) \in \mathcal{M}_k\Big\}\Big) \leq \sqrt{\frac{\ln(N)}{\theta}} + \frac{3}{\sqrt{2\theta}} \tag{7.360}$$

for

$$\theta = \frac{4^k \delta_2 (n-1)}{16\kappa^2} \quad \text{and} \quad N = |\mathcal{M}_k| < 16^{(k+2)m}. \tag{7.361}$$

The remainder of the proof consists of routine calculations showing that the required bound is achieved. Using the bound

$$\sqrt{\ln(N)} \leq \sqrt{\log(N)} < 2\sqrt{(k+2)m}, \tag{7.362}$$

summing over all $k \in \mathbb{N}$, and making use of the summations

$$\sum_{k=0}^{\infty} 2^{-k}\sqrt{k+2} \,<\, \frac{7}{2} \quad \text{and} \quad \sum_{k=0}^{\infty} 2^{-k} \,=\, 2, \tag{7.363}$$

one concludes that

$$\sum_{k=0}^{\infty} \mathrm{E}\Big(\max\Big\{|Y_x - Y_y| : (x,y) \in \mathcal{M}_k\Big\}\Big) < \frac{64\kappa}{\sqrt{\delta_2}}\sqrt{\frac{m}{n}}. \tag{7.364}$$

By (7.353), (7.355), and (7.364), it follows that

$$\mathrm{E}\big(\sup\{|Y_v - \beta| \,:\, v \in \mathcal{S}(\mathcal{V})\}\big) < \left(\frac{8}{\sqrt{\delta_3}} + \frac{64}{\sqrt{\delta_2}}\right)\kappa\sqrt{\frac{m}{n}}. \tag{7.365}$$

Under the assumption that

$$m \leq \frac{\delta \varepsilon^2 \zeta^2 n}{\kappa^2}, \tag{7.366}$$

for δ satisfying (7.344), it therefore holds that

$$\mathrm{E}\big(\sup\{|Y_v - \beta| \,:\, v \in \mathcal{S}(\mathcal{V})\}\big) < \zeta\varepsilon, \tag{7.367}$$

which completes the proof. □

7.3.2 Applications of measure concentration

Two applications of the results on measure concentration discussed in the previous subsection will now be presented. The first is a demonstration that most pure states of a pair of registers are highly entangled, and the second is a proof that the minimum output entropy of channels is non-additive in general. The two applications are related, with the second depending on the first.

Most pure states are highly entangled

Suppose that $\mathcal{X}$ and $\mathcal{Y}$ are complex Euclidean spaces, and suppose further that the dimensions $n = \dim(\mathcal{X})$ and $m = \dim(\mathcal{Y})$ of these spaces satisfy $n \leq m$. For some choices of a unit vector $u \in \mathcal{X} \otimes \mathcal{Y}$, it holds that

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = \omega, \tag{7.368}$$

for $\omega = \mathbb{1}/n$ denoting the completely mixed state with respect to $\mathcal{X}$. Of course, not every unit vector $u \in \mathcal{X} \otimes \mathcal{Y}$ satisfies this equation (unless $n = 1$); but as n grows, the equation holds approximately for an increasingly large portion of the set $\mathcal{S}(\mathcal{X} \otimes \mathcal{Y})$.

The following lemma establishes one specific fact along these lines, in which an approximation with respect to the 2-norm distance between states is considered. The proof makes use of Lévy's lemma (Theorem 7.37), along with calculations of integrals involving the uniform spherical measure.

Lemma 7.45 *There exists a positive real number K_0 with the following property. For complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$ of dimensions $n = \dim(\mathcal{X})$ and $m = \dim(\mathcal{Y})$, and for*

$$X : \mathcal{S}(\mathcal{X} \otimes \mathcal{Y}) \to \mathbb{R} \tag{7.369}$$

being a random variable, distributed with respect to the uniform spherical measure on $\mathcal{S}(\mathcal{X} \otimes \mathcal{Y})$ and defined as

$$X(u) = \left\| \mathrm{Tr}_{\mathcal{Y}}(uu^*) - \omega \right\|_2 \tag{7.370}$$

for $\omega = \mathbb{1}/n$, it holds that

$$\Pr\left(X \geq \frac{K_0}{\sqrt{m}} \right) < 4^{-n}. \tag{7.371}$$

Proof It will be proved that the lemma holds for $K_0 = \sqrt{12/\delta_2} + 1$, for δ_2 being any positive real number satisfying the requirements of the mean value form of Lévy's lemma (Theorem 7.37).

The random variable X may alternatively be defined as

$$X(\operatorname{vec}(A)) = \big\| AA^* - \omega \big\|_2 \tag{7.372}$$

for every operator $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ satisfying $\|A\|_2 = 1$. The triangle inequality implies that

$$\big| X(\operatorname{vec}(A)) - X(\operatorname{vec}(B)) \big| \leq \big\| AA^* - BB^* \big\|_2. \tag{7.373}$$

Again using the triangle inequality, along with the fact that the 2-norm is submultiplicative, one has

$$\begin{aligned} \big\| AA^* - BB^* \big\|_2 &\leq \big\| AA^* - AB^* \big\|_2 + \big\| AB^* - BB^* \big\|_2 \\ &\leq \big(\|A\|_2 + \|B\|_2\big)\|A - B\|_2 \leq 2\|A - B\|_2, \end{aligned} \tag{7.374}$$

for all $A, B \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ with $\|A\|_2 = \|B\|_2 = 1$. It therefore holds that X is 2-Lipschitz.

Next, it will be proved that

$$\mathrm{E}(X) \leq \frac{1}{\sqrt{m}}. \tag{7.375}$$

This bound follows from Jensen's inequality,

$$\big(\mathrm{E}(X)\big)^2 \leq \mathrm{E}(X^2), \tag{7.376}$$

along with an evaluation of $\mathrm{E}(X^2)$. To evaluate this expectation, observe first that

$$\big\| \operatorname{Tr}_{\mathcal{Y}}(uu^*) - \omega \big\|_2^2 = \operatorname{Tr}\Big(\big(\operatorname{Tr}_{\mathcal{Y}}(uu^*)\big)^2 \Big) - \frac{1}{n}. \tag{7.377}$$

For every vector $u \in \mathcal{X} \otimes \mathcal{Y}$, it holds that

$$\operatorname{Tr}\Big(\big(\operatorname{Tr}_{\mathcal{Y}}(uu^*)\big)^2 \Big) = \big\langle V, uu^* \otimes uu^* \big\rangle, \tag{7.378}$$

for $V \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{X} \otimes \mathcal{Y})$ being the operator defined as

$$V(x_0 \otimes y_0 \otimes x_1 \otimes y_1) = x_1 \otimes y_0 \otimes x_0 \otimes y_1 \tag{7.379}$$

for all vectors $x_0, x_1 \in \mathcal{X}$ and $y_0, y_1 \in \mathcal{Y}$. Equivalently, for Σ and Γ denoting the alphabets for which $\mathcal{X} = \mathbb{C}^{\Sigma}$ and $\mathcal{Y} = \mathbb{C}^{\Gamma}$, one may write

$$V = \sum_{\substack{a,b \in \Sigma \\ c,d \in \Gamma}} E_{a,b} \otimes E_{c,c} \otimes E_{b,a} \otimes E_{d,d}. \tag{7.380}$$

Integrating with respect to the uniform spherical measure yields

$$\begin{aligned} \mathrm{E}(X^2) &= \int \langle V, uu^* \otimes uu^* \rangle \, \mathrm{d}\mu(u) - \frac{1}{n} \\ &= \frac{1}{\binom{nm+1}{2}} \big\langle V, \Pi_{(\mathcal{X}\otimes\mathcal{Y})_{\circledS}(\mathcal{X}\otimes\mathcal{Y})} \big\rangle - \frac{1}{n}. \end{aligned} \tag{7.381}$$

A case analysis reveals that

$$\begin{aligned} &\big\langle E_{a,b} \otimes E_{c,c} \otimes E_{b,a} \otimes E_{d,d}, \Pi_{(\mathcal{X}\otimes\mathcal{Y})_{\circledS}(\mathcal{X}\otimes\mathcal{Y})} \big\rangle \\ &\quad = \begin{cases} 1 & \text{if } a = b \text{ and } c = d \\ \frac{1}{2} & \text{if } (a = b \text{ and } c \neq d) \text{ or } (a \neq b \text{ and } c = d) \\ 0 & \text{if } a \neq b \text{ and } c \neq d. \end{cases} \end{aligned} \tag{7.382}$$

Performing the required arithmetic yields

$$\mathrm{E}(X^2) = \frac{n+m}{nm+1} - \frac{1}{n} < \frac{1}{m}, \tag{7.383}$$

and therefore (7.375) has been established.

Finally, by the mean value form of Lévy's lemma (Theorem 7.37), one has

$$\Pr\left(X \geq \frac{K_0}{\sqrt{m}} \right) \leq 2 \exp\left(-\frac{\delta_2 (K_0 - 1)^2 n}{4} \right). \tag{7.384}$$

For $K_0 = \sqrt{12/\delta_2} + 1$, one has

$$2 \exp\left(-\frac{\delta_2 (K_0 - 1)^2 n}{4} \right) = 2 \exp(-3n) < 4^{-n}, \tag{7.385}$$

which completes the proof. □

If $\mathrm{Tr}_{\mathcal{Y}}(uu^*)$ is approximately equal to the completely mixed state ω, for a given unit vector $u \in \mathcal{X} \otimes \mathcal{Y}$, then it is reasonable to expect that the entanglement entropy $\mathrm{H}\big(\mathrm{Tr}_{\mathcal{Y}}(uu^*)\big)$ of the pure state represented by u will be approximately equal to its maximum possible value $\log(\dim(\mathcal{X}))$, depending on the particular notions of approximate equality under consideration. The following lemma establishes a lower bound on the von Neumann entropy that allows a precise implication along these lines to be made when combined with Lemma 7.45.

Lemma 7.46 *Let $\mathcal{X}$ be a complex Euclidean space and let $n = \dim(\mathcal{X})$. For every density operator $\rho \in \mathrm{D}(\mathcal{X})$ it holds that*

$$\mathrm{H}(\rho) \geq \log(n) - \frac{n}{\ln(2)} \big\| \rho - \omega \big\|_2^2, \tag{7.386}$$

where $\omega = \mathbb{1}/n$ denotes the completely mixed state with respect to $\mathcal{X}$.

Proof It holds that $\ln(\alpha) \leq \alpha - 1$ for all $\alpha > 0$, and therefore

$$\begin{aligned} \frac{n}{\ln(2)} \bigl\| \rho - \omega \bigr\|_2^2 &= \frac{n \operatorname{Tr}(\rho^2) - 1}{\ln(2)} \\ &\geq \log\bigl(n \operatorname{Tr}(\rho^2)\bigr) = \log(n) + \log\bigl(\operatorname{Tr}(\rho^2)\bigr). \end{aligned} \tag{7.387}$$

The logarithm function is concave, and therefore one has

$$-\operatorname{H}(p) = \sum_{a \in \Sigma} p(a) \log(p(a)) \leq \log\left(\sum_{a \in \Sigma} p(a)^2 \right) \tag{7.388}$$

for every alphabet Σ and every probability vector $p \in \mathcal{P}(\Sigma)$. Consequently,

$$-\operatorname{H}(\rho) \leq \log\bigl(\operatorname{Tr}(\rho^2)\bigr), \tag{7.389}$$

and therefore

$$\frac{n}{\ln(2)} \bigl\| \rho - \omega \bigr\|_2^2 \geq \log(n) - \operatorname{H}(\rho), \tag{7.390}$$

which is equivalent to the required inequality. □

As a consequence of Lemmas 7.45 and 7.46, it follows that most bipartite pure states have an entanglement entropy that is close to this quantity's maximum possible value.

Theorem 7.47 *There exists a positive real number K with the following property. For every choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, and for $X : \mathcal{S}(\mathcal{X} \otimes \mathcal{Y}) \to \mathbb{R}$ being a random variable, distributed with respect to the uniform spherical measure on $\mathcal{S}(\mathcal{X} \otimes \mathcal{Y})$ and defined as*

$$X(u) = \operatorname{H}\bigl(\operatorname{Tr}_{\mathcal{Y}}(uu^*)\bigr) \tag{7.391}$$

for every $u \in \mathcal{S}(\mathcal{X} \otimes \mathcal{Y})$, it holds that

$$\Pr\left(X \leq \log(n) - \frac{Kn}{m} \right) < 4^{-n}, \tag{7.392}$$

for $n = \dim(\mathcal{X})$ and $m = \dim(\mathcal{Y})$.

Proof It will be proved that the theorem holds for $K = K_0^2 / \ln(2)$, where K_0 is any positive real number that satisfies the requirements of Lemma 7.45.

Define a random variable $Y : \mathcal{S}(\mathcal{X} \otimes \mathcal{Y}) \to \mathbb{R}$, distributed with respect to the uniform spherical measure, as

$$Y(u) = \bigl\| \operatorname{Tr}_{\mathcal{Y}}(uu^*) - \omega \bigr\|_2 \tag{7.393}$$

for every $u \in \mathcal{S}(\mathcal{X} \otimes \mathcal{Y})$. If a given unit vector $u \in \mathcal{X} \otimes \mathcal{Y}$ satisfies

$$Y(u) < \frac{K_0}{\sqrt{m}}, \tag{7.394}$$

then

$$X(u) > \log(n) - \frac{n}{\ln(2)} \frac{K_0^2}{m} = \log(n) - \frac{Kn}{m} \tag{7.395}$$

by Lemma 7.46. One therefore has that

$$\Pr\left(X > \log(n) - \frac{Kn}{m} \right) \geq \Pr\left(Y < \frac{K_0}{\sqrt{m}} \right) > 1 - 4^{-n} \tag{7.396}$$

by Lemma 7.45. This bound is equivalent to (7.392), which completes the proof. □

Counter-example to the additivity of minimum output entropy

The minimum output entropy of a channel is, as the following definition states explicitly, the minimum value of the von Neumann entropy that can be obtained by evaluating that channel on a quantum state input.

Definition 7.48 Let $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be a channel, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. The *minimum output entropy* of Φ is defined as

$$\mathrm{H}_{\min}(\Phi) = \min\bigl\{ \mathrm{H}(\Phi(\rho)) \,:\, \rho \in \mathrm{D}(\mathcal{X}) \bigr\}. \tag{7.397}$$

It follows from the concavity of the von Neumann entropy function that the minimum output entropy $\mathrm{H}_{\min}(\Phi)$ of a given channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ is achieved by a pure state:

$$\mathrm{H}_{\min}(\Phi) = \min\bigl\{ \mathrm{H}(\Phi(uu^*)) \,:\, u \in \mathcal{S}(\mathcal{X}) \bigr\}. \tag{7.398}$$

It was a long-standing conjecture that the minimum output entropy is additive with respect to tensor products of channels. The following theorem demonstrates that this is, in fact, not the case.

Theorem 7.49 (Hastings) *There exist complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$ and channels $\Phi, \Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ such that*

$$\mathrm{H}_{\min}(\Phi \otimes \Psi) < \mathrm{H}_{\min}(\Phi) + \mathrm{H}_{\min}(\Psi). \tag{7.399}$$

A high-level overview of the proof of Theorem 7.49 is as follows. For each choice of a positive integer n, one may consider complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ with

$$\dim(\mathcal{X}) = n^2, \quad \dim(\mathcal{Y}) = n, \quad \text{and} \quad \dim(\mathcal{Z}) = n^2. \tag{7.400}$$

It will be proved, for a sufficiently large choice of n, that there exists an isometry $V \in \mathrm{U}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ for which the channels $\Phi, \Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ defined as

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}(VXV^*) \quad \text{and} \quad \Psi(X) = \mathrm{Tr}_{\mathcal{Z}}(\overline{V} X V^{\mathsf{T}}) \tag{7.401}$$

for all $X \in \mathrm{L}(\mathcal{X})$ yield the strict inequality (7.399). The existence of a suitable isometry V is proved using the probabilistic method: for any fixed isometry $V_0 \in \mathrm{U}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$, the set of all unitary operators $U \in \mathrm{U}(\mathcal{Y} \otimes \mathcal{Z})$ for which the isometry $V = UV_0$ possesses the required property will be shown to have positive measure, with respect to the Haar measure on $\mathrm{U}(\mathcal{Y} \otimes \mathcal{Z})$.

The proof of Theorem 7.49 will make use of the lemmas that follow. The first lemma provides an upper bound on the minimum output entropy of the tensor product $\Phi \otimes \Psi$ for two channels Φ and Ψ defined as in (7.401).

Lemma 7.50 *Let n be a positive integer and let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be complex Euclidean spaces with $\dim(\mathcal{X}) = n^2$, $\dim(\mathcal{Y}) = n$, and $\dim(\mathcal{Z}) = n^2$. Let $V \in \mathrm{U}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ be an isometry, and define channels $\Phi, \Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ as*

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(VXV^*\big) \quad \textit{and} \quad \Psi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(\overline{V}XV^{\mathsf{T}}\big) \tag{7.402}$$

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

$$\mathrm{H}_{\min}\big(\Phi \otimes \Psi\big) \leq 2\log(n) - \frac{\log(n) - 2}{n}. \tag{7.403}$$

Proof Define pure states $\tau \in \mathrm{D}(\mathcal{X} \otimes \mathcal{X})$ and $\sigma \in \mathrm{D}(\mathcal{Y} \otimes \mathcal{Y})$ as follows:

$$\tau = \frac{\mathrm{vec}(\mathbb{1}_{\mathcal{X}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*}{n^2} \quad \text{and} \quad \sigma = \frac{\mathrm{vec}(\mathbb{1}_{\mathcal{Y}})\,\mathrm{vec}(\mathbb{1}_{\mathcal{Y}})^*}{n}. \tag{7.404}$$

A calculation reveals that

$$\big\langle \sigma, (\Phi \otimes \Psi)(\tau) \big\rangle = \frac{1}{n^3} \big\| \mathrm{Tr}_{\mathcal{Y}}(VV^*) \big\|_2^2. \tag{7.405}$$

In greater detail, supposing that $\mathcal{Y} = \mathbb{C}^{\Sigma}$, one has

$$\begin{aligned}
&\big\langle \sigma, (\Phi \otimes \Psi)(\tau) \big\rangle \\
&= \frac{1}{n} \sum_{a,b \in \Sigma} \Big\langle V^*(E_{a,b} \otimes \mathbb{1}_{\mathcal{Z}})V \otimes V^{\mathsf{T}}(E_{a,b} \otimes \mathbb{1}_{\mathcal{Z}})\overline{V}, \tau \Big\rangle \\
&= \frac{1}{n^3} \sum_{a,b \in \Sigma} \mathrm{Tr}\Big(\big(V^*(E_{b,a} \otimes \mathbb{1}_{\mathcal{Z}})V\big) \big(V^*(E_{a,b} \otimes \mathbb{1}_{\mathcal{Z}})V\big) \Big) \\
&= \frac{1}{n^3} \big\| \mathrm{Tr}_{\mathcal{Y}}(VV^*) \big\|_2^2.
\end{aligned} \tag{7.406}$$

The operator $\mathrm{Tr}_{\mathcal{Y}}(VV^*)$ is positive semidefinite, and has trace equal to n^2 and rank at most n^2, so it follows that its 2-norm squared must be at least n^2. Consequently, one has

$$\lambda_1\big((\Phi \otimes \Psi)(\tau)\big) \geq \big\langle \sigma, (\Phi \otimes \Psi)(\tau) \big\rangle \geq \frac{1}{n}. \tag{7.407}$$

Now, under the constraint that a given density operator $\rho \in \mathrm{D}(\mathcal{Y} \otimes \mathcal{Y})$ has largest eigenvalue at least $1/n$, it holds that the von Neumann entropy $\mathrm{H}(\rho)$ is maximized when this largest eigenvalue is equal to $1/n$ and all other eigenvalues are equal:

$$\mathrm{H}(\rho) \leq \Big(1 - \frac{1}{n}\Big) \log(n^2 - 1) + \mathrm{H}\Big(\frac{1}{n}, 1 - \frac{1}{n}\Big). \tag{7.408}$$

Because $\ln(\alpha) \geq 1 - 1/\alpha$ for all positive α, one finds that

$$\mathrm{H}(\lambda, 1 - \lambda) \leq -\lambda \log(\lambda) + \frac{\lambda}{\ln(2)} \leq -\lambda \log(\lambda) + 2\lambda \tag{7.409}$$

for all $\lambda \in [0, 1]$, and therefore

$$\mathrm{H}(\rho) \leq 2 \log(n) - \frac{\log(n) - 2}{n}. \tag{7.410}$$

As this inequality holds for $\rho = (\Phi \otimes \Psi)(\tau)$ the proof is complete. □

The remaining lemmas required for the proof of Theorem 7.49 are used to establish a lower bound on the quantity $\mathrm{H}_{\min}(\Phi) + \mathrm{H}_{\min}(\Psi)$, for some choice of channels Φ and Ψ taking the form (7.401). The first lemma is concerned with the modification of a random variable that is Lipschitz on a compact subset of its domain, yielding one that is Lipschitz everywhere.

Lemma 7.51 *Let $\mathcal{X}$ be a complex Euclidean space, let $X : \mathcal{S}(\mathcal{X}) \to \mathbb{R}$ be a continuous random variable, distributed with respect to the uniform spherical measure μ on $\mathcal{S}(\mathcal{X})$, and let $\mathcal{A} \subseteq \mathcal{S}(\mathcal{X})$ be a compact subset of $\mathcal{S}(\mathcal{X})$ satisfying $\mu(\mathcal{A}) \geq 3/4$. Let κ be a positive real number such that*

$$|X(x) - X(y)| \leq \kappa \|x - y\| \tag{7.411}$$

for all $x, y \in \mathcal{A}$, and define a new random variable $Y : \mathcal{S}(\mathcal{X}) \to \mathbb{R}$, distributed with respect to μ, as

$$Y(x) = \min_{y \in \mathcal{A}} \big(X(y) + \kappa \|x - y\|\big) \tag{7.412}$$

for all $x \in \mathcal{S}(\mathcal{X})$. The following statements hold:

1. *Y is κ-Lipschitz.*
2. *For every $x \in \mathcal{A}$, one has that $X(x) = Y(x)$.*
3. *Every median value of Y is a central value of X.*

Proof The first statement holds regardless of the behavior of X on points in $\mathcal{A}$. Consider any two vectors $x_0, x_1 \in \mathcal{S}(\mathcal{X})$, and let $y_0, y_1 \in \mathcal{A}$ satisfy

$$Y(x_0) = X(y_0) + \kappa \|x_0 - y_0\| \quad \text{and} \quad Y(x_1) = X(y_1) + \kappa \|x_1 - y_1\|. \tag{7.413}$$

That is, y_0 and y_1 achieve the minimum values that define the function Y on x_0 and x_1, respectively. It must therefore hold that

$$X(y_0) + \kappa\|x_0 - y_0\| \leq X(y_1) + \kappa\|x_0 - y_1\|, \tag{7.414}$$

which implies

$$Y(x_0) - Y(x_1) \leq \kappa\|x_0 - y_1\| - \kappa\|x_1 - y_1\| \leq \kappa\|x_0 - x_1\|. \tag{7.415}$$

The inequality

$$Y(x_1) - Y(x_0) \leq \kappa\|x_0 - x_1\| \tag{7.416}$$

is proved through the same argument by exchanging the indices 0 and 1. It therefore holds that

$$|Y(x_0) - Y(x_1)| \leq \kappa\|x_0 - x_1\|, \tag{7.417}$$

so Y is κ-Lipschitz.

Next, consider any vector $x \in \mathcal{A}$. By the assumptions of the lemma, one has

$$|X(x) - X(y)| \leq \kappa\|x - y\| \tag{7.418}$$

for every $y \in \mathcal{A}$, and therefore

$$Y(x) - X(x) = \min_{y \in \mathcal{A}} \Big(X(y) - X(x) + \kappa\|x - y\| \Big) \geq 0. \tag{7.419}$$

On the other hand, because one may choose $y = x$ when considering the minimum, it holds that $Y(x) \leq X(x)$. It follows that $X(x) = Y(x)$, which establishes the second statement.

Finally, let $\alpha \in \mathbb{R}$ be a median value of Y, so that

$$\Pr(Y \geq \alpha) \geq \frac{1}{2} \quad \text{and} \quad \Pr(Y \leq \alpha) \geq \frac{1}{2}. \tag{7.420}$$

Define a random variable $Z : \mathcal{S}(\mathcal{X}) \to [0, 1]$, again distributed with respect to μ, as

$$Z(x) = \begin{cases} 1 & \text{if } x \in \mathcal{A} \\ 0 & \text{if } x \notin \mathcal{A}, \end{cases} \tag{7.421}$$

so that $\Pr(Z = 0) \leq 1/4$. By the union bound, one has

$$\Pr(Y < \alpha \text{ or } Z = 0) \leq \frac{3}{4}, \tag{7.422}$$

and therefore

$$\Pr(X \geq \alpha) \geq \Pr(Y \geq \alpha \text{ and } Z = 1) \geq \frac{1}{4}. \tag{7.423}$$

By similar reasoning,

$$\Pr(X \leq \alpha) \geq \Pr(Y \leq \alpha \text{ and } Z = 1) \geq \frac{1}{4}. \tag{7.424}$$

This implies that α is a central value of X, which completes the proof. □

The next lemma is, in some sense, the heart of the proof of Theorem 7.49. It establishes the existence of an isometry $V \in \mathrm{U}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ that may be taken in the definition (7.401) of the channels Φ and Ψ to obtain the inequality (7.399) for a sufficiently large value of n. It is proved through the use of Dvoretzky's theorem.

Lemma 7.52 *There exists a real number $K > 0$ for which the following statement holds. For every choice of a positive integer n, and for $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ being complex Euclidean spaces with*

$$\dim(\mathcal{X}) = n^2, \quad \dim(\mathcal{Y}) = n, \quad \text{and} \quad \dim(\mathcal{Z}) = n^2, \tag{7.425}$$

there exists an isometry $V \in \mathrm{U}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ such that

$$\left\| \mathrm{Tr}_{\mathcal{Z}}(Vxx^*V^*) - \omega \right\|_2 \leq \frac{K}{n} \tag{7.426}$$

for every unit vector $x \in \mathcal{S}(\mathcal{X})$, where $\omega = \mathbb{1}/n$ denotes the completely mixed state with respect to $\mathcal{Y}$.

Proof Let δ be a positive real number that satisfies the requirements of Dvoretzky's theorem (Theorem 7.42) and let K_0 be a positive real number satisfying the requirements of Lemma 7.45. It will be proved that the lemma holds for

$$K = K_0 + 6\sqrt{\frac{K_0 + 1}{\delta}} + \frac{18}{\delta}. \tag{7.427}$$

Assume, for the remainder of the proof, that a positive integer n and complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ satisfying (7.425) have been fixed. Let $\mathcal{V}$ be an arbitrary subspace of $\mathcal{Y} \otimes \mathcal{Z}$ having dimension n^2. Throughout the proof, μ will denote the uniform spherical measure on $\mathcal{S}(\mathcal{Y} \otimes \mathcal{Z})$, and η will denote the Haar measure on $\mathrm{U}(\mathcal{Y} \otimes \mathcal{Z})$.

The first step of the proof is the specification of a collection of random variables; an analysis of these random variables follows their specification. First, let

$$X, Y : \mathcal{S}(\mathcal{Y} \otimes \mathcal{Z}) \to \mathbb{R} \tag{7.428}$$

be random variables, distributed with respect to the uniform spherical

measure μ and defined as follows:

$$X(u) = \sqrt{\bigl\|\mathrm{Tr}_{\mathcal{Z}}(uu^*)\bigr\|} \quad \text{and} \quad Y(u) = \bigl\|\mathrm{Tr}_{\mathcal{Z}}(uu^*) - \omega\bigr\|_2 \tag{7.429}$$

for all $u \in \mathcal{S}(\mathcal{Y} \otimes \mathcal{Z})$. Next, let

$$K_1 = \sqrt{K_0 + 1} + \frac{3}{\sqrt{\delta}} \quad \text{and} \quad \kappa = \frac{2K_1}{\sqrt{n}}, \tag{7.430}$$

define a set

$$\mathcal{A} = \left\{ u \in \mathcal{S}(\mathcal{Y} \otimes \mathcal{Z}) \,:\, X(u) \leq \frac{K_1}{\sqrt{n}} \right\}, \tag{7.431}$$

and define a random variable $Z : \mathcal{S}(\mathcal{Y} \otimes \mathcal{Z}) \to \mathbb{R}$, also distributed with respect to the uniform spherical measure μ, as

$$Z(u) = \min_{v \in \mathcal{A}} \bigl(Y(v) + \kappa \| u - v \| \bigr) \tag{7.432}$$

for every $u \in \mathcal{S}(\mathcal{Y} \otimes \mathcal{Z})$. It is evident from their specifications that X, Y, and Z are phase-invariant random variables. Finally, for each unit vector $v \in \mathcal{S}(\mathcal{V})$, define random variables

$$P_v, Q_v, R_v : \mathrm{U}(\mathcal{Y} \otimes \mathcal{Z}) \to \mathbb{R}, \tag{7.433}$$

distributed with respect to the Haar measure η on $\mathrm{U}(\mathcal{Y} \otimes \mathcal{Z})$, as

$$P_v(U) = X(Uv), \quad Q_v(U) = Y(Uv), \quad \text{and} \quad R_v(U) = Z(Uv), \tag{7.434}$$

for every $U \in \mathrm{U}(\mathcal{Y} \otimes \mathcal{Z})$.

When analyzing the random variables that have just been defined, it is helpful to begin with the observation that

$$X(\mathrm{vec}(A)) = \|A\| \quad \text{and} \quad Y(\mathrm{vec}(A)) = \bigl\|AA^* - \omega\bigr\|_2 \tag{7.435}$$

for every operator $A \in \mathrm{L}(\mathcal{Z}, \mathcal{Y})$ satisfying $\|A\|_2 = 1$. It is immediate from the first of these expressions, along with the inequality $\|A\| \leq \|A\|_2$, that X is 1-Lipschitz. Also, given that

$$\|A\|^2 = \bigl\|AA^*\bigr\| \leq \bigl\|AA^* - \omega\bigr\| + \|\omega\| \leq \bigl\|AA^* - \omega\bigr\|_2 + \frac{1}{n} \tag{7.436}$$

for every operator $A \in \mathrm{L}(\mathcal{Z}, \mathcal{Y})$, one necessarily has that

$$X^2 \leq Y + \frac{1}{n}. \tag{7.437}$$

By Lemma 7.45, one may therefore conclude that

$$\Pr\left(X \leq \sqrt{\frac{K_0 + 1}{n}} \right) \geq \Pr\left(Y \leq \frac{K_0}{n} \right) > \frac{3}{4}. \tag{7.438}$$

Dvoretzky's theorem (Theorem 7.42) will be applied twice in the proof, with the first application concerning the random variables X and P_v for each $v \in \mathcal{S}(\mathcal{V})$. By (7.438), it follows that every central value of X is at most

$$\sqrt{\frac{K_0 + 1}{n}}. \tag{7.439}$$

Setting

$$\varepsilon = \frac{3}{\sqrt{\delta n}} \quad \text{and} \quad \zeta = \frac{1}{3} \tag{7.440}$$

in Dvoretzky's theorem yields

$$\Pr\left(P_v \leq \frac{K_1}{\sqrt{n}} \text{ for every } v \in \mathcal{S}(\mathcal{V})\right) \geq \frac{2}{3}, \tag{7.441}$$

by virtue of the fact that $\dim(\mathcal{V}) = \delta\varepsilon^2\zeta^2 \dim(\mathcal{Y} \otimes \mathcal{Z})$.

The second application of Dvoretzky's theorem concerns Z and R_v for each $v \in \mathcal{S}(\mathcal{V})$. Before applying Dvoretzky's theorem, however, the implications of Lemma 7.51 to the random variables Y and Z will be considered. First, note that

$$\mu(\mathcal{A}) = \Pr\left(X \leq \frac{K_1}{\sqrt{n}}\right) \geq \Pr\left(X \leq \sqrt{\frac{K_0 + 1}{n}}\right) > \frac{3}{4}. \tag{7.442}$$

Second, for any choice of vectors $u, v \in \mathcal{A}$, one may write $u = \operatorname{vec}(A)$ and $v = \operatorname{vec}(B)$ for $A, B \in \mathrm{L}(\mathcal{Z}, \mathcal{Y})$ satisfying $\|A\|_2 = \|B\|_2 = 1$, so that

$$\|A\| = X(\operatorname{vec}(A)) \leq \frac{K_1}{\sqrt{n}} \quad \text{and} \quad \|B\| = X(\operatorname{vec}(B)) \leq \frac{K_1}{\sqrt{n}}. \tag{7.443}$$

This implies that

$$\begin{aligned} |Y(u) - Y(v)| &= \Big| \|AA^* - \omega\|_2 - \|BB^* - \omega\|_2 \Big| \\ &\leq \|AA^* - BB^*\|_2 \leq (\|A\| + \|B\|)\|A - B\|_2 \leq \kappa \|u - v\|. \end{aligned} \tag{7.444}$$

It therefore follows from Lemma 7.51 that Z is κ-Lipschitz, Z and Y agree everywhere on $\mathcal{A}$, and every median value of Z is a central value of Y. By (7.438), every central value of Y is at most K_0/n, and therefore the same upper bound applies to every median value of Z. Setting

$$\varepsilon = \frac{3\kappa}{\sqrt{\delta n}} \quad \text{and} \quad \zeta = \frac{1}{3} \tag{7.445}$$

and applying Dvoretzky's theorem therefore yields

$$\Pr\left(R_v \leq \frac{K}{n} \text{ for all } v \in \mathcal{S}(\mathcal{V})\right) \geq \frac{2}{3}, \tag{7.446}$$

by virtue of the fact that

$$\dim(\mathcal{V}) = \frac{\delta\varepsilon^2\zeta^2}{\kappa^2}\dim(\mathcal{Y}\otimes\mathcal{Z}). \tag{7.447}$$

Finally, consider the random variables Y and Q_v for each $v \in \mathcal{S}(\mathcal{V})$. For every vector $u \in \mathcal{S}(\mathcal{Y}\otimes\mathcal{Z})$, one has either $u \in \mathcal{A}$ or $u \notin \mathcal{A}$; and if it holds that $u \in \mathcal{A}$, then $Y(u) = Z(u)$. Consequently, if it holds that $Y(u) > K/n$ for a given choice of $u \in \mathcal{S}(\mathcal{Y}\otimes\mathcal{Z})$, then it must hold that

$$Z(u) > \frac{K}{n} \quad \text{or} \quad X(u) > \frac{K_1}{\sqrt{n}} \tag{7.448}$$

(or both). By the union bound, one concludes that

$$\begin{aligned}
&\Pr\Bigg(Q_v > \frac{K}{n} \text{ for some } v \in \mathcal{S}(\mathcal{V})\Bigg) \\
&\qquad \leq \Pr\Bigg(R_v > \frac{K}{n} \text{ for some } v \in \mathcal{S}(\mathcal{V})\Bigg) \\
&\qquad\quad + \Pr\Bigg(P_v > \frac{K_1}{\sqrt{n}} \text{ for some } v \in \mathcal{S}(\mathcal{V})\Bigg).
\end{aligned} \tag{7.449}$$

By (7.441) and (7.446), it follows that

$$\Pr\Bigg(Q_v \leq \frac{K}{n} \text{ for all } v \in \mathcal{S}(\mathcal{V})\Bigg) \geq \frac{1}{3} > 0. \tag{7.450}$$

By (7.450), one concludes that there exists a unitary operator U for which $Q_v(U) \leq K/n$ for all $v \in \mathcal{S}(\mathcal{V})$. Taking $V_0 \in \mathrm{U}(\mathcal{X}, \mathcal{Y}\otimes\mathcal{Z})$ to be any linear isometry for which $\mathrm{im}(V_0) = \mathcal{V}$, one therefore has

$$\left\| \mathrm{Tr}_{\mathcal{Z}}(UV_0xx^*V_0^*U^*) - \omega \right\|_2 \leq \frac{K}{n} \tag{7.451}$$

for every unit vector $x \in \mathcal{S}(\mathcal{X})$. Taking $V = UV_0$, the lemma is proved. □

Finally, a proof of Theorem 7.49 is to be presented. The proof is made quite straightforward through the use of Lemmas 7.50 and 7.52.

Proof of Theorem 7.49 Let $K > 0$ be a real number for which Lemma 7.52 holds, and choose n to be a positive integer satisfying

$$\log(n) > \frac{2K^2}{\ln(2)} + 2. \tag{7.452}$$

For $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ being complex Euclidean spaces with $\dim(\mathcal{X}) = n^2$, $\dim(\mathcal{Y}) = n$, and $\dim(\mathcal{Z}) = n^2$, it follows (by Lemma 7.52) that there exists an isometry $V \in \mathrm{U}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ such that

$$\left\| \mathrm{Tr}_{\mathcal{Z}}(Vxx^*V^*) - \frac{\mathbb{1}_{\mathcal{Y}}}{n} \right\|_2 \leq \frac{K}{n} \tag{7.453}$$

for every unit vector $x \in \mathcal{S}(\mathcal{X})$. By Lemma 7.46, one therefore has that

$$\mathrm{H}\big(\mathrm{Tr}_{\mathcal{Z}}(Vxx^*V^*)\big) \geq \log(n) - \frac{K^2}{n \ln(2)} \tag{7.454}$$

for every $x \in \mathcal{S}(\mathcal{X})$. Replacing V by the entry-wise complex conjugate of V results in the same bound:

$$\mathrm{H}\big(\mathrm{Tr}_{\mathcal{Z}}(\overline{V}xx^*V^{\mathsf{T}})\big) \geq \log(n) - \frac{K^2}{n \ln(2)} \tag{7.455}$$

for every $x \in \mathcal{S}(\mathcal{X})$.

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

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}(VXV^*) \quad \text{and} \quad \Psi(X) = \mathrm{Tr}_{\mathcal{Z}}(\overline{V}XV^{\mathsf{T}}) \tag{7.456}$$

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

$$\mathrm{H}_{\min}(\Phi) = \mathrm{H}_{\min}(\Psi) \geq \log(n) - \frac{K^2}{n \ln(2)}, \tag{7.457}$$

and therefore

$$\mathrm{H}_{\min}(\Phi) + \mathrm{H}_{\min}(\Psi) \geq 2\log(n) - \frac{2K^2}{n \ln(2)}. \tag{7.458}$$

On the other hand, Lemma 7.50 implies that

$$\mathrm{H}_{\min}(\Phi \otimes \Psi) \leq 2\log(n) - \frac{\log(n) - 2}{n}. \tag{7.459}$$

Consequently,

$$\begin{aligned} &\mathrm{H}_{\min}(\Phi \otimes \Psi) - \big(\mathrm{H}_{\min}(\Phi) + \mathrm{H}_{\min}(\Psi)\big) \\ &\qquad = \frac{2K^2}{n \ln(2)} - \frac{\log(n) - 2}{n} < 0, \end{aligned} \tag{7.460}$$

which completes the proof. □

7.4 Exercises

Exercise 7.1 For every positive integer $n \geq 2$, define a unital channel $\Phi_n \in \mathrm{C}(\mathbb{C}^n)$ as

$$\Phi_n(X) = \frac{1}{n-1}\operatorname{Tr}(X)\mathbb{1}_n - \frac{1}{n-1}X^{\mathsf{T}} \tag{7.461}$$

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 a mixed-unitary channel when n is even. (Observe that this exercise is complementary to Exercise 4.2.)

Exercise 7.2 Let n and m be positive integers with $n < m$, and consider the set $\mathrm{U}(\mathbb{C}^n, \mathbb{C}^m)$ of all isometries from $\mathbb{C}^n$ to $\mathbb{C}^m$.

(a) Prove that there exists a Borel probability measure

$$\nu : \mathrm{Borel}\big(\mathrm{U}(\mathbb{C}^n, \mathbb{C}^m)\big) \to [0, 1] \tag{7.462}$$

for which it holds that

$$\nu(\mathcal{A}) = \nu(U\mathcal{A}V) \tag{7.463}$$

for every choice of a Borel subset $\mathcal{A} \in \mathrm{Borel}(\mathrm{U}(\mathbb{C}^n, \mathbb{C}^m))$ and unitary operators $U \in \mathrm{U}(\mathbb{C}^m)$ and $V \in \mathrm{U}(\mathbb{C}^n)$.

(b) Prove that if

$$\mu : \mathrm{Borel}\big(\mathrm{U}(\mathbb{C}^n, \mathbb{C}^m)\big) \to [0, 1] \tag{7.464}$$

is a Borel probability measure on $\mathrm{U}(\mathbb{C}^n, \mathbb{C}^m)$ satisfying

$$\mu(\mathcal{A}) = \mu(U\mathcal{A}) \tag{7.465}$$

for every for every choice of a Borel subset $\mathcal{A} \in \mathrm{Borel}(\mathrm{U}(\mathbb{C}^n, \mathbb{C}^m))$ and a unitary operator $U \in \mathrm{U}(\mathbb{C}^m)$, then it must hold that $\mu = \nu$, where ν is the measure defined by a correct solution to part (a).

Exercise 7.3 Let $\mathcal{X}$ be a complex Euclidean space, let $n = \dim(\mathcal{X})$, and define a mapping $\Phi \in \mathrm{CP}(\mathcal{X})$ as

$$\Phi(X) = n \int \langle uu^*, X \rangle uu^* \, \mathrm{d}\mu(u) \tag{7.466}$$

for all $X \in \mathrm{L}(\mathcal{X})$, where μ denotes the uniform spherical measure on $\mathcal{S}(\mathcal{X})$. Give a simple, closed-form expression for Φ.

Exercise 7.4 Let $\mathcal{X}$ be a complex Euclidean space, let $n = \dim(\mathcal{X})$, and define a channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{X} \otimes \mathcal{X})$ as

$$\Phi(X) = n \int \langle uu^*, X \rangle uu^* \otimes uu^* \, \mathrm{d}\mu(u) \tag{7.467}$$

for all $X \in \mathrm{L}(\mathcal{X})$, where μ denotes the uniform spherical measure on $\mathcal{S}(\mathcal{X})$. Give a closed-form expression for the minimum cloning fidelity

$$\alpha(\Phi) = \inf_{v \in \mathcal{S}(\mathcal{X})} \mathrm{F}\big(\Phi(vv^*), vv^* \otimes vv^*\big) \tag{7.468}$$

obtained through the use of Φ. (Observe that Φ is a sub-optimal cloning channel, in the sense of Theorem 7.28, aside from the trivial case in which $\dim(\mathcal{X}) = 1$.)

Exercise 7.5 Prove that there exists a positive real number K with the following property. For every positive integer n and every nonnegative κ-Lipschitz random variable

$$X : \mathcal{S}(\mathbb{C}^n) \to [0, \infty), \tag{7.469}$$

distributed with respect to the uniform spherical measure on $\mathcal{S}(\mathbb{C}^n)$, one has that

$$\mathrm{E}(X^2) - \mathrm{E}(X)^2 \leq \frac{K\kappa^2}{n}. \tag{7.470}$$

Exercise 7.6 Prove that there exist positive real numbers $K, \delta > 0$ for which the following statement holds. For every choice of a complex Euclidean space $\mathcal{X}$, a κ-Lipschitz nonnegative random variable

$$X : \mathcal{S}(\mathcal{X}) \to [0, \infty), \tag{7.471}$$

distributed with respect to the uniform spherical measure μ on $\mathcal{S}(\mathcal{X})$, and every positive real number $\varepsilon > 0$, it holds that

$$\Pr\left(\left|X - \sqrt{\mathrm{E}(X^2)}\right| \geq \varepsilon\right) \leq K \exp\left(-\frac{\delta\varepsilon^2 n}{\kappa^2}\right). \tag{7.472}$$

The fact established by a correct solution to Exercise 7.5 is useful for proving this result. (Observe that a correct solution to this problem establishes a variant of Lévy's lemma in which concentration occurs around the root-mean-squared value of a nonnegative random variable, as opposed to its mean or central values.)

7.5 Bibliographic remarks

Permutation-invariant vectors and operators are commonly studied objects in multilinear algebra, which is the subject of the books of Greub (1978) and Marcus (1973, 1975), among others. These concepts and generalizations of them are also relevant to the subject of representation theory, as explained in the book of Goodman and Wallach (1998), for instance. Theorem 7.14 is a finite-dimensional form of the double commutant theorem, also known as the bicommutant theorem, proved by von Neumann (1930).

The existence of unitarily invariant measures on both the unit sphere and the set of unitary operators in a complex Euclidean space is implied by a much more general construction due to Haar (1933). Von Neumann (1933) proved the uniqueness of the measures constructed by Haar, with their two papers appearing consecutively in the same journal. This work was further generalized by Weil (1979) and others. Due to the generality of these notions, many books that include a discussion of Haar measure do not consider the specialized definitions of uniform spherical measure or Haar measure (for unitary operators in finite dimensions) of the sort that has been presented in this chapter. Definitions of this type are, however, fairly standard in random matrix theory. These definitions are rooted in the work of Dyson (1962a,b,c) and Diaconis and Shahshahani (1987), and a more broad overview of random matrix theory may be found in the book of Mehta (2004).

The Werner twirling channel, defined in Example 7.25, was introduced by Werner (1989) in the same paper, mentioned in the previous chapter, that introduced the states now known as Werner states. Theorem 7.28 on optimal cloning of pure states is also due to Werner (1998). The original no-cloning theorem is generally attributed to Wootters and Zurek (1982) and Deiks (1982), although an equivalent statement and proof appear in an earlier paper of Park (1970). Although not published until 1983, a paper of Wiesner (1983) proposing a scheme for unforgeable money based on quantum information, relying implicitly on the assumption that quantum states cannot be cloned, was allegedly written in the late 1960s.

Multiple versions of the quantum de Finetti theorem are known. These theorems are so-named because they generalize theorems in combinatorics and probability theory originally found in the work of de Finetti (1937). A quantum information-theoretic variant of de Finetti's eponymous theorem was first proved by Hudson and Moody (1976) in 1976. Caves, Fuchs, and Schack (2002) later gave a simpler proof of this theorem. Like the original de Finetti theorem, this was a qualitative result regarding the behavior of

an infinite number of identical systems. A finite quantum formulation of de Finetti's theorem, closer in spirit to classical results due to Diaconis and Freedman (1980), was proved by König and Renner (2005). Theorems 7.12 and 7.26 and Corollary 7.27 were proved by Christandl, König, Mitchison, and Renner (2007), who improved on the error bounds and generalized the results obtained by König and Renner.

Theorem 7.31 and Corollary 7.32 are due to Watrous (2009a).

Readers interested in learning more about the phenomenon of measure concentration are referred to the books of Ledoux (2001) and Milman and Schechtman (1986). Theorems 7.37 and 7.40 are variants of a theorem due to Lévy (1951). The proofs of these theorems appearing in this chapter have mostly followed those in Appendix V of Milman and Schechtman's book (which are partially based on a technique due to Maurey and Pisier (1976)). Multiple formulations of Dvoretzky's theorem are known, with the original having been proved by Dvoretzky around 1960 (Dvoretzky, 1961). Milman (1971) gave a proof of Dvoretzky's theorem in 1971 based on the measure concentration phenomenon, which he was the first to explicitly identify.

To prove Theorem 7.49 on the non-additivity of the minimum output entropy, a particularly sharp version of Dvoretzky's theorem (as stated in Theorem 7.42) is evidently required. The proof of this theorem, as well as its application to Theorem 7.49, is due to Aubrun, Szarek, and Werner (2011). The proof makes essential use of the *chaining method* of Talagrand (2006).

There are several known applications of the concentration of measure phenomenon to quantum information theory, the first of which were due to Hayden, Leung, Shor, and Winter (2004), Bennett, Hayden, Leung, Shor, and Winter (2005), and Harrow, Hayden, and Leung (2004). Theorem 7.47 is a variant of a theorem due to Hayden, Leung, and Winter (2006).

Theorem 7.49 was proved by Hastings (2009), based in part on Hayden and Winter's disproof of the so-called *maximal p-norm multiplicativity conjecture* shortly before (Hayden and Winter, 2008). As suggested above, the proof of Theorem 7.49 that has been presented in this chapter is due to Aubrun, Szarek, and Werner (2011). The implications of Hastings discovery to the study of channel capacities is discussed in the next chapter.