set $\mathrm{D}(\mathcal{H})$, the extremum is achieved on a pure state, which significantly reduces the dimension of the problem.

As opposed to pure states, which are extremal, the "most central" element in $\mathrm{D}(\mathcal{H})$ is the state $\mathrm{I}/\dim\mathcal{H}$, which is called the *maximally mixed state* and denoted by ρ_* when there is no ambiguity. We also note that the set of states on $\mathcal{H}$ which are diagonal with respect to a given orthonormal basis $(e_i)_{i\in I}$ naturally identifies with the set of classical states on I.

EXERCISE 2.1. Describe states which belong to the boundary of $\mathrm{D}(\mathcal{H})$.

EXERCISE 2.2 (Every state is an average of pure states). Show that every state $\rho \in \mathrm{D}(\mathbb{C}^d)$ can be written as $\frac{1}{d}(|\psi_1\rangle\langle\psi_1| + \cdots + |\psi_d\rangle\langle\psi_d|)$ for some unit vectors $\psi_1, \dots, \psi_d$ in $\mathbb{C}^d$.

2.1.2. The Bloch ball $\mathrm{D}(\mathbb{C}^2)$. The situation for $d = 2$ is very special. Let $\rho \in \mathsf{M}_2^{\mathrm{sa}}$, with $\operatorname{Tr}\rho = 1$. Then ρ has two eigenvalues, which can be written as $1/2-\lambda$ and $1/2+\lambda$ for some $\lambda \in \mathbb{R}$. Moreover, $\rho \geqslant 0$ if and only if $|\lambda| \leqslant 1/2$. On the other hand, we have

$$\|\rho - \rho_*\|_{\mathrm{HS}} = \sqrt{2}|\lambda|.$$

Therefore, ρ is a state if and only if $\|\rho - \rho_*\|_{\mathrm{HS}} \leqslant 1/\sqrt{2}$. What we have proved is that, inside the space of trace one self-adjoint operators, the set of states is a Euclidean ball centered at ρ_* and with radius $1/\sqrt{2}$. This ball is called the *Bloch ball* and its boundary is called the *Bloch sphere*. Once we introduce the *Pauli matrices*

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{2.2}$$

a convenient orthonormal basis (with respect to the Hilbert–Schmidt inner product) in $\mathsf{M}_2^{\mathrm{sa}}$ is

$$\left(\frac{1}{\sqrt{2}}\mathrm{I}, \frac{1}{\sqrt{2}}\sigma_x, \frac{1}{\sqrt{2}}\sigma_y, \frac{1}{\sqrt{2}}\sigma_z\right). \tag{2.3}$$

A very useful consequence of $\mathrm{D}(\mathbb{C}^2)$ being a ball is the fact—mentioned already in Section 1.2.1—that the cone $\mathcal{PSD}(\mathbb{C}^2)$ is isomorphic (or even isometric in the appropriate sense) to the Lorentz cone $\mathcal{L}_4$. A popular explicit isomorphism, inducing the so-called *spinor map* (see Appendix C), is given by

$$\mathbb{R}^4 \ni \mathbf{x} = (t, x, y, z) \mapsto \begin{bmatrix} t+z & x-iy \\ x+iy & t-z \end{bmatrix} = X \in \mathsf{M}_2^{\mathrm{sa}}. \tag{2.4}$$

The formula for X can be rewritten in terms of the Pauli matrices (2.2) as

$$X = t\,\mathrm{I} + x\sigma_x + y\sigma_y + z\sigma_z, \tag{2.5}$$

and so a convenient expression for it is $X = \mathbf{x}\cdot\sigma$, where σ is a shorthand for $(\mathrm{I}, \sigma_x, \sigma_y, \sigma_z)$, and "$\cdot$" is a "formal dot product". Since $\{\mathrm{I}, \sigma_x, \sigma_y, \sigma_z\}$ is a multiple of the orthonormal basis (2.3) of $\mathsf{M}_2^{\mathrm{sa}}$, it follows that the map given by (2.4) is likewise a multiple of isometry (with respect to the Euclidean metric in the domain and the Hilbert–Schmidt metric in the range). Next, it is readily verified that

$$\frac{1}{2}\operatorname{Tr} X = t, \qquad \det X = t^2 - x^2 - y^2 - z^2 =: q(\mathbf{x}), \tag{2.6}$$

where q is the quadratic form of the Minkowski spacetime, which confirms that $X \in \mathcal{PSD}(\mathbb{C}^2)$ iff $\mathbf{x} \in \mathcal{L}_4$. The isomorphism $\mathbf{x} \mapsto \mathbf{x}\cdot\sigma$ will be useful in understanding automorphisms of the cones $\mathcal{L}_4$ and $\mathcal{PSD}(\mathbb{C}^2)$, and when proving Størmer's theorem in Section 2.4.5.

When $d > 2$, the set $\mathrm{D}(\mathbb{C}^d)$ is no longer a ball, but rather the non-commutative analogue of a simplex. Its symmetrization (see Section 4.1.2)

$$\mathrm{D}(\mathbb{C}^d)_{\mathcal{S}} = \mathrm{conv}\left(\mathrm{D}(\mathbb{C}^d) \cup -\mathrm{D}(\mathbb{C}^d)\right) = \{A \in \mathsf{M}_d^{\mathrm{sa}} \ : \ \|A\|_1 \leqslant 1\}$$

is $S_1^{d,\mathrm{sa}}$, the unit ball of the self-adjoint part of the 1-Schatten space (see Section 1.3.2).

One way to quantify the fact that the set $\mathrm{D}(\mathbb{C}^d)$ is different from a ball when $d > 2$ is to compute the radius of its inscribed and circumscribed Hilbert–Schmidt balls. The former equals $1/\sqrt{d(d-1)}$ while the latter is $\sqrt{(d-1)/d}$ (the same values as for the set Δ_{d-1} of classical states on $\{1,\dots,d\}$, and for the same reasons). In other words, if we denote by $B(\rho_*, r)$ the ball centered at ρ_* and with Hilbert–Schmidt radius r inside the hyperplane $H_1 = \{\mathrm{Tr}(\cdot) = 1\} \subset \mathsf{M}_d^{\mathrm{sa}}$, we have

$$B\left(\rho_*, \frac{1}{\sqrt{d(d-1)}}\right) \subset \mathrm{D}(\mathbb{C}^d) \subset B\left(\rho_*, \sqrt{\frac{d-1}{d}}\right), \tag{2.7}$$

and these values—differing by the factor of $d-1$—are the best possible.

Exercise 2.3 (The Bloch sphere is a sphere). Show that the matrix X given by (2.5) has eigenvalues 1 and -1 if and only if $t = 0$ and $x^2 + y^2 + z^2 = 1$.

Exercise 2.4 (Composition rules for Pauli matrices). Verify the composition rules for Pauli matrices. (i) $\sigma_a^2 = \mathrm{I}$. (ii) If a, b, c are all different, then $\sigma_a\sigma_b = i\varepsilon\sigma_c$, where $\varepsilon = \pm 1$ is the sign of the permutation $(x, y, z) \mapsto (a, b, c)$; in particular, if $a \neq b$, then $\sigma_a\sigma_b = -\sigma_b\sigma_a$.

2.1.3. Facial structure.

Proposition 2.1 (Characterization of faces of D). *There is a one-to-one correspondence between nontrivial subspaces of $\mathbb{C}^d$ and proper faces of $\mathrm{D}(\mathbb{C}^d)$. Given a subspace $\{0\} \subsetneq E \subsetneq \mathbb{C}^d$, the corresponding face $\mathrm{D}(E)$ is the set of states whose range is contained in E:*

$$\mathrm{D}(E) = \{\rho \in \mathrm{D}(\mathbb{C}^d) \ : \ \rho(\mathbb{C}^d) \subset E\}.$$

In particular, pure states (extreme points, i.e., minimal, 0-dimensional faces) correspond to the case $\dim E = 1$. In the direction opposed to a pure state $|x\rangle\langle x|$ lies a face which corresponds to all states with a range orthogonal to x; these are maximal proper faces.

Remark 2.2. All faces of $\mathrm{D}(\mathbb{C}^d)$ are exposed (as defined in Exercise 1.5) since $\mathrm{D}(E)$ is the intersection of $\mathrm{D}(\mathbb{C}^d)$ with the hyperplane $\{X \ : \ \mathrm{Tr}(XP_E) = 1\}$.

Proof of Proposition 2.1. Denote by $\mathrm{range}(\rho) = \rho(\mathbb{C}^d)$ the range of a state $\rho \in \mathrm{D}(\mathbb{C}^d)$. We use the following observation: if $\rho, \sigma \in \mathrm{D}(\mathbb{C}^d)$ and $\lambda \in (0, 1)$, then

$$\mathrm{range}(\lambda\rho + (1-\lambda)\sigma) = \mathrm{range}(\rho) + \mathrm{range}(\sigma). \tag{2.8}$$

We first check that, for any nontrivial subspace $E \subset \mathbb{C}^d$, $\mathrm{D}(E)$ is a face of $\mathrm{D}(\mathbb{C}^d)$. For indeed, if $\rho \in \mathrm{D}(E)$ can be written as $\lambda\rho_1 + (1-\lambda)\rho_2$ for $\rho_1, \rho_2 \in \mathrm{D}(\mathbb{C}^d)$ and $\lambda \in (0, 1)$, then (2.8) implies that $\mathrm{range}(\rho_1) \subset E$ and $\mathrm{range}(\rho_2) \subset E$.

Conversely, let $F \subset \mathrm{D}(\mathbb{C}^d)$ be a proper face. Define $E = \bigcup\{\mathrm{range}(\rho) \ : \ \rho \in F\}$. It follows—from (2.8) and from the fact that F is convex—that E is actually a

CHAPTER 2

The mathematics of quantum information theory

This chapter puts into mathematical perspective some basic concepts of quantum information theory. (For a physically motivated approach, see Chapter 3.) We discuss the geometry of the set of quantum states, the entanglement vs. separability dichotomy, and introduce completely positive maps and quantum channels. All these concepts will be extensively used in Chapters 8–12.

2.1. On the geometry of the set of quantum states

2.1.1. Pure and mixed states. In this section we take a closer look at the set $\mathrm{D}(\mathcal{H})$ (or simply D) of quantum states on a finite-dimensional complex Hilbert space $\mathcal{H}$. By definition (see Section 0.10), we have

$$\mathrm{D}(\mathcal{H}) = \{\rho \in B_{\mathrm{sa}}(\mathcal{H}) \ : \ \rho \geqslant 0, \operatorname{Tr} \rho = 1\}. \tag{2.1}$$

If $\mathcal{H} = \mathbb{C}^d$, the definition (2.1) simply says that $\mathrm{D}(\mathbb{C}^d)$ is the base of the positive semi-definite cone $\mathcal{PSD}(\mathbb{C}^d)$ defined by the hyperplane $H_1 \subset \mathsf{M}_d^{\mathrm{sa}}$ of trace one Hermitian matrices (cf. (1.22)). The (real) dimension of the set $\mathrm{D}(\mathbb{C}^d)$ equals $d^2 - 1$: it has non-empty interior inside H_1. (This follows from $\mathcal{PSD}(\mathbb{C}^d)$ being a full cone.)

A state $\rho \in \mathrm{D}(\mathcal{H})$ is called *pure* if it has rank 1, i.e., if there is a unit vector $\psi \in \mathcal{H}$ such that

$$\rho = |\psi\rangle\langle\psi|.$$

Note that $|\psi\rangle\langle\psi|$ is the orthogonal projection onto the (complex) line spanned by ψ. We sometimes use the terminology "consider a pure state ψ" (such language is prevalent in physics literature). What we mean is that ψ is a unit vector and we consider the corresponding pure state $|\psi\rangle\langle\psi|$. We use the terminology of *mixed* states when we want to emphasize that we consider the set of all states, not necessarily pure.

Let ψ, χ be unit vectors in $\mathcal{H}$. Then the pure states $|\psi\rangle\langle\psi|$ and $|\chi\rangle\langle\chi|$ coincide if and only if there is a complex number λ with $|\lambda| = 1$ such that $\chi = \lambda\psi$. Therefore the set of pure states identifies with $\mathsf{P}(\mathcal{H})$, the projective space on $\mathcal{H}$. (See Appendix B.2; note that the space $\mathsf{P}(\mathbb{C}^d)$ is more commonly denoted by $\mathbb{CP}^{d-1}$.)

The set $\mathrm{D}(\mathcal{H})$ is a compact convex set, and it is easily checked that the extreme points of $\mathrm{D}(\mathcal{H})$ are exactly the pure states (cf. Proposition 1.9 and Corollary 1.10). It follows from general convexity theory (Krein–Milman and Carathéodory's theorems) that any state is a convex combination of at most $(\dim \mathcal{H})^2$ pure states. However, using the spectral theorem instead tells us more: any state is a convex combination of at most $\dim \mathcal{H}$ pure states $|\psi_i\rangle\langle\psi_i|$, where (ψ_i) are pairwise orthogonal unit vectors (cf. Exercise 1.45). A fundamental consequence is that whenever we want to maximize a convex function (or minimize a concave function) over the

(iii) *There is an $n \times n$ bistochastic matrix B such that $y = Bx$ (a matrix is bistochastic if its entries are non-negative, and add up to 1 in each row and each column).*

(iv) *Whenever ϕ is a permutationally invariant convex function on $\mathbb{R}^n$, then $\phi(x) \leqslant \phi(y)$.*

(v) *For every $t \in \mathbb{R}$, we have $\sum_{i=1}^n |x_i - t| \leqslant \sum_{i=1}^n |y_i - t|$.*

(vi) *For every $t \in \mathbb{R}$, we have $\sum_{i=1}^n (x_i - t)^+ \leqslant \sum_{i=1}^n (y_i - t)^+$, where $x^+ = \max(x, 0)$.*

SKETCH OF THE PROOF. Fix $y \in \mathbb{R}^n$, and consider the non-empty convex compact set

$$K_y = \{x \in \mathbb{R}^n \ : \ x \prec y\}.$$

It is easily checked that x is an extreme point of K_y if and only if $x^\downarrow = y^\downarrow$, and it follows from the Krein–Milman theorem that (i) is equivalent to (ii). Similarly, the classical Birkhoff theorem, which asserts that extreme points of the set of bistochastic matrices are exactly permutation matrices, gives the equivalence of (ii) and (iii). The implications (ii) $\Rightarrow$ (iv) $\Rightarrow$ (v) are obvious. We check that (v) and (vi) are equivalent since $|x| = 2x^+ - x$ (using the fact that $\sum x_i = \sum y_i$). Finally, for $t = y_k^\downarrow$, we compute

$$\sum_{i=1}^n (y_i - t)^+ = \sum_{i=1}^k (y_i^\downarrow - t) = \sum_{i=1}^k y_i^\downarrow - kt$$

$$\sum_{i=1}^n (x_i - t)^+ = \sum_{i=1}^n (x_i^\downarrow - t)^+ \geqslant \sum_{i=1}^k (x_i^\downarrow - t)^+ \geqslant \sum_{i=1}^k (x_i^\downarrow - t) = \sum_{i=1}^k x_i^\downarrow - kt.$$

Therefore, the inequality from (vi) implies that $\sum_{i=1}^k x_i^\downarrow \leqslant \sum_{i=1}^k y_i^\downarrow$, hence $x \prec y$. □

EXERCISE 1.37. Show that, in the statement of Proposition 1.12, we have to assume the hypothesis $\sum x_i = \sum y_i$ only in (vi); in (ii)–(v) this property follows formally.

EXERCISE 1.38 (Submajorization). Given $x, y \in \mathbb{R}^n$, we say that x is *submajorized* by y and write $x \prec_w y$ if (1.28) holds (the difference with majorization is that we do not assume $\sum x_i = \sum y_i$). Show that $x \prec_w y$ if and only if there exists $u \in \mathbb{R}^n$ such that $u \prec y$ and $x_k \leqslant u_k$ for every $1 \leqslant k \leqslant n$.

1.3.2. Schatten norms. Recall that the space $\mathsf{M}_{m,n}$ of (real or complex) $m \times n$ matrices carries a Euclidean structure given by the Hilbert–Schmidt inner product (see Section 0.6). The Hilbert–Schmidt norm is a special case of the *Schatten p-norms*, which are the noncommutative analogues of the ℓ_p-norms. If $M \in \mathsf{M}_{m,n}$, define $|M| := (M^\dagger M)^{1/2}$, and for $1 \leqslant p \leqslant \infty$,

$$\|M\|_p := (\operatorname{Tr}|M|^p)^{1/p}.$$

Note that $\|\cdot\|_{\mathrm{HS}} = \|\cdot\|_2$. The case $p = \infty$ should be interpreted as the limit $p \to \infty$ of the above, and corresponds to the usual operator norm

$$\|M\|_\infty = \|M\|_{\mathrm{op}} := \sup_{|x| \leqslant 1} |Mx|.$$

The quantity $\|M\|_1 = \operatorname{Tr}|M|$ is called the *trace norm* of M. Occasionally we will loosely refer to various matrix spaces endowed with Schatten norms as *Schatten spaces* or p-Schatten spaces.

There is ambiguity in the notation $\|\cdot\|_p$ in that it has two possible meanings: the Schatten p-norm on $\mathsf{M}_{m,n}$ (matrices) and the usual ℓ_p-norm on $\mathbb{R}^n$ or $\mathbb{C}^n$ (sequences). However, it will be always clear from the context which of the two is the intended one.

If $M \in \mathsf{M}_{m,n}$, and if we denote by $s(M) = (s_1(M), \ldots, s_n(M))$ the singular values of M (i.e., the eigenvalues of $|M|$) arranged in the non-increasing order, then for any p,

$$\|M\|_p = \|s(M)\|_p. \tag{1.29}$$

The following lemma allows us to reduce the study of Schatten norms to the case of self-adjoint matrices.

LEMMA 1.13. *Let $M \in \mathsf{M}_{m,n}$, and $\tilde{M} \in \mathsf{M}_{m+n}$ be the self-adjoint matrix defined by*

$$\tilde{M} = \begin{bmatrix} 0 & M \\ M^\dagger & 0 \end{bmatrix}.$$

Then we have $\|\tilde{M}\|_p = 2^{1/p}\|M\|_p$ for $1 \leqslant p \leqslant \infty$. Similarly, if $M, N \in \mathsf{M}_{m,n}$, then $\operatorname{Tr} \tilde{M}\tilde{N} = 2 \operatorname{Re} \operatorname{Tr} M^\dagger N$.

PROOF. For the first assertion, it suffices to notice that the eigenvalues of $\tilde{M}$ are equal to $\pm s_i(M)$. The second assertion is verified by direct calculation. □

The next lemma shows how the concept of majorization relates to eigenvalues/singular values of a matrix.

LEMMA 1.14 (Spectrum majorizes the diagonal). *Let $M \in \mathsf{M}_n$ be a self-adjoint matrix, let $d(M) = (m_{ii}) \in \mathbb{R}^m$ be the vector of diagonal entries of M, and let $\operatorname{spec}(M) = (\lambda_i) \in \mathbb{R}^m$ be the vector of eigenvalues of M, arranged in non-increasing order. Then $d(M) \prec \operatorname{spec}(M)$.*

PROOF. First, it is known from linear algebra that $\sum_i m_{ii} = \sum_i \lambda_i$, so majorization is in principle possible. Write M as $M = U\Lambda U^\dagger$, where Λ is a diagonal matrix whose entries are the eigenvalues of M, and U is a unitary matrix. We then have

$$m_{ii} = \sum_j u_{ij}\lambda_j \overline{u_{ji}} = \sum_j |u_{ij}|^2 \lambda_j.$$

Since the matrix with entries $|u_{ij}|^2$ is bistochastic, the assertion follows from Proposition 1.12 (iii). □

We now state the Davis convexity theorem, which gives a characterization of all convex functions f on M_m^{sa} that are unitarily invariant.

PROPOSITION 1.15 (Davis convexity theorem). *Let $f : \mathsf{M}_m^{sa} \to \mathbb{R}$ a function which is unitarily invariant, i.e., such that $f(UAU^\dagger) = f(A)$ for any self-adjoint matrix A and any unitary matrix U. Then f is convex if and only if the restriction of f to the subspace of diagonal matrices is convex.*

PROOF. Assume that the restriction of f to diagonal matrices is convex (the converse implication being obvious). This restriction, when considered as a function on $\mathbb{R}^m$, is permutationally invariant, as can be checked by choosing for U a permutation matrix. Given $0 < \lambda < 1$ and $A, B \in \mathsf{M}_m^{sa}$, we need to show that

$$f(\lambda A + (1-\lambda)B) \leqslant \lambda f(A) + (1-\lambda)f(B). \tag{1.30}$$

Since f is unitarily invariant, we may assume that the matrix $\lambda A + (1-\lambda)B$ is diagonal. Denoting by $\operatorname{diag} A$ the matrix obtained from a matrix A by changing all its off-diagonal elements to 0, the hypothesis on f implies

$$f(\lambda A + (1-\lambda)B) \leqslant \lambda f(\operatorname{diag} A) + (1-\lambda)f(\operatorname{diag} B).$$

Using Lemma 1.14 and Proposition 1.12(iv), it follows that $f(\operatorname{diag} A) \leqslant f(A)$ and $f(\operatorname{diag} B) \leqslant f(B)$, showing (1.30). □

An immediate consequence of the Davis convexity theorem is that the Schatten p-norms satisfy the triangle inequality.

PROPOSITION 1.16. *For $1 \leqslant p \leqslant \infty$, if $M, N \in \mathsf{M}_{m,n}$, we have*

$$\|M + N\|_p \leqslant \|M\|_p + \|N\|_p.$$

PROOF. By the first assertion of Lemma 1.13, it is enough to consider the case of $m = n$ and self-adjoint M, N. We now use Proposition 1.15 for the unitarily invariant function $f(\cdot) = \|\cdot\|_p$. The restriction of $\|\cdot\|_p$ to the subspace of diagonal matrices identifies with the usual (commutative) ℓ_p-norm on $\mathbb{R}^n$, and hence, by Proposition 1.15, the function $\|\cdot\|_p$ is convex on M_m^{sa}. Since it is also positively homogeneous, the triangle inequality follows. □

Obviously, the Schatten p-norms of a given matrix satisfy the same inequalities as the ℓ_p-norms: if $1 \leqslant p \leqslant q \leqslant \infty$, and M is an $m \times n$ matrix (with $m \leqslant n$; what is important is that the rank of M is at most m), then

$$\|M\|_q \leqslant \|M\|_p \leqslant m^{1/p-1/q}\|M\|_p. \tag{1.31}$$

Duality between Schatten p-norms holds as in the commutative case.

PROPOSITION 1.17 (The non-commutative Hölder inequality). *Let $1 \leqslant p, q \leqslant \infty$ such that $1/p + 1/q = 1$, and $M \in \mathsf{M}_{m,n}$, $N \in \mathsf{M}_{n,m}$. We have*

$$|\operatorname{Tr} MN| \leqslant \|M\|_p\|N\|_q. \tag{1.32}$$

As a consequence, the Schatten p-norm and q-norm are dual to each other. This holds in all settings: for rectangular matrices (real or complex), for Hermitian matrices, and for real symmetric matrices.

As in the case of ℓ_p^n-spaces, the above duality relation can be equivalently expressed in terms of polars. Denote by $S_p^{m,n}$ the unit ball associated to the Schatten norm $\|\cdot\|_p$ on $\mathsf{M}_{m,n}$ and $S_p^{m,\mathrm{sa}} := S_p^{m,m} \cap \mathsf{M}_m^{\mathrm{sa}}$. (Again, there are two settings, real and complex, and some care needs to be exercised as minor subtleties occasionally arise.) We then have

COROLLARY 1.18. *If $1 \leqslant p, q \leqslant \infty$ with $1/p + 1/q = 1$, then*

$$S_q^{m,n} = \{A \in \mathsf{M}_{m,n} : |\langle X, A\rangle| \leqslant 1 \text{ for all } X \in S_p^{m,n}\} \tag{1.33}$$

$$= \{A \in \mathsf{M}_{m,n} : \mathrm{Re}\langle X, A\rangle \leqslant 1 \text{ for all } X \in S_p^{m,n}\} \tag{1.34}$$

$$S_q^{m,\mathrm{sa}} = \left(S_p^{m,\mathrm{sa}}\right)^\circ, \tag{1.35}$$

where $\langle\cdot,\cdot\rangle$ and $^\circ$ are meant in the sense of trace duality (0.4).

While (1.33) and (1.35) are simply straightforward reformulations of duality relations from Proposition 1.17, the equality in (1.34) needs to be justified (only the inclusion "$\subset$" is immediate). Given $A \in \mathsf{M}_{m,n}$ and $X \in S_p^{m,n}$ such that $|\langle X, A\rangle| > 1$, let $\xi = \frac{\langle X, A\rangle}{|\langle X, A\rangle|}$. Then, setting $X' = \bar{\xi} X$, we see that $X' \in S_p^{m,n}$, while $\mathrm{Re}\langle X', A\rangle = |\langle X, A\rangle| > 1$, which yields the other inclusion "$\supset$" in (1.34). The expression in (1.34) can be thought of as a definition of the polar $\left(S_p^{m,n}\right)^\circ$ by "dropping the complex structure"; see Exercise 1.48 for the general principle. Another potential complication is that, in the complex setting, the identification with the dual space is anti-linear, see Section 0.2. Note that no issues of such nature arise in defining the polar of $S_p^{m,\mathrm{sa}}$, as that set "lives" in a real inner product space irrespectively of the setting.

PROOF OF PROPOSITION 1.17. Consider first the Hermitian case. By unitary invariance, we may assume that M is diagonal. We then have

$$|\operatorname{Tr}(MN)| = \Big|\sum_i m_{ii} n_{ii}\Big| \leqslant \|(m_{ii})\|_p \|(n_{ii})\|_q \leqslant \|M\|_p \|N\|_q,$$

where we used the commutative Hölder inequality, Lemma 1.14, and Proposition 1.12 (iv).

In the general case, Lemma 1.13 and the Hermitian case of (1.32) shown above imply that, for all $M, N \in \mathsf{M}_{n,m}$,

$$\mathrm{Re}\operatorname{Tr} M^\dagger N \leqslant \|M\|_p \|N\|_q,$$

and the same bound for $|\operatorname{Tr}(MN)|$ (or $|\operatorname{Tr}(M^\dagger N)|$) follows by the same trick as the one used to establish equality in (1.34) (see the paragraph following Corollary 1.18).

As in the commutative case, Hölder's inequality constitutes "the hard part" of the duality assertion, such as the inclusion $S_q^{m,\mathrm{sa}} \subset \left(S_p^{m,\mathrm{sa}}\right)^\circ$ in (1.35). "The easy part" involves establishing that for every M, there is $N \neq 0$ such that we have equality in (1.32). In the Hermitian case, this follows readily by restricting attention to matrices that diagonalize in the same orthonormal basis as M and by appealing to the analogous statement for the usual ℓ_p-norm. In the general case one considers similarly the singular value decomposition (SVD) of M. □

EXERCISE 1.39 (Davis convexity theorem, the real case). State and prove a real version of Proposition 1.15, i.e., for functions defined on the set of real symmetric matrices.

$2^{N-2k_N-2} \leqslant \frac{1}{2}s_0(2^{k_N})$, so that $\pi_{2^{k_N},2^{N-2k_N-2}} \leqslant \delta$. By Exercise 10.7(i), this implies that $p_{k_N+1} \leqslant \delta$ and the Corollary follows.

Exercise 10.9. (i) We have $\mathrm{Tr}(\rho^2) = \frac{1}{d^2} + \mathrm{Tr}(W\rho)$. The value of $\mathbf{E}\,\mathrm{Tr}(\rho^2)$ was computed in Exercise 6.45. To obtain concentration, use the fact that $\mathrm{Tr}\,\rho^2$ is related to the Schatten 4-norm of M when $\rho = MM^\dagger$, with M distributed on the Hilbert–Schmidt sphere in $\mathsf{M}_{d^2,s}$. (ii) Let Π be the orthogonal projection onto the subspace $\mathbb{C}x \otimes \mathbb{C}^s$. The function $|\Pi\psi| = \sqrt{\langle x|\rho|x\rangle}$ is 1-Lipschitz as a function of ψ and satisfies $\mathbf{E}\,|\Pi\psi|^2 = 1/d^2$; use Exercise 5.46. (iii) Use Lemma 9.4 and the union bound.

Exercise 10.10. It follows from (the proof of) Carathéodory's theorem (see Exercise 1.1) that the infimum in (10.15) can be restricted to convex combinations of length at most d^4. Then use a compactness argument.

Exercise 10.11. The inradius of PPT is the same as that of Sep (see Table 9.1), so the argument that led to (10.14) carries over to the present setting. For the bound in (i), the relevant range of s is $\Theta(d^2)$.

Chapter 11

Exercise 11.1. If $n \geqslant 3$ is odd, argue as in the comment following Lemma 11.1. If $n = 2k$, identify $\mathbb{R}^n$ with $\mathbb{R}^2 \otimes \mathbb{R}^k$ and consider $E = F \otimes \mathrm{I}_k$, where $F \subset \mathsf{M}_2(\mathbb{R})$ is the subspace spanned by the two real Pauli matrices.

Exercise 11.2. This can be seen directly from the definition. Alternatively, we may use the description from Proposition 11.8. Let $\lambda \in (0,1)$ and $(a_{ij}), (a'_{ij}) \in \mathsf{QC}_{m,n}$. We have $a_{ij} = \langle x_i, y_j\rangle$ and $a'_{ij} = \langle x'_i, y'_j\rangle$. Defining $\tilde{x}_i = \sqrt{\lambda}\, x_i \oplus \sqrt{1-\lambda}\, x'_i$ and $\tilde{y}_j = \sqrt{\lambda}\, y_j \oplus \sqrt{1-\lambda}\, y'_j$ leads to $\lambda a_{ij} + (1-\lambda)a'_{ij} = \langle \tilde{x}_i, \tilde{y}_j\rangle$. We then argue as in the end of the proof of Proposition 11.8 to ensure that vectors live in $\mathbb{R}^{\min(m,n)}$.

Exercise 11.3. For vectors x_i, y_j of norm at most 1, the unit vectors $x'_i = x_i + \sqrt{1-|x_i|^2}\, u$ and $y'_j = y_j + \sqrt{1-|y_j|^2}\, v$ satisfy $\langle x_i, y_j\rangle = \langle x'_i, y'_j\rangle$ provided u, v are unit vectors in $\{y_j \;:\; 1 \leqslant j \leqslant n\}^\perp \cap \{x_i \;:\; 1 \leqslant i \leqslant m\}^\perp$ such that $u \perp v$.

Exercise 11.4. When considered as elements of $\mathbb{R}^4$, the 8 distinct matrices $A^{\xi,\eta} = (\xi_i\eta_j)_{i,j=1}^2$ are either opposite or orthogonal. A less explicit argument goes as follows: use Proposition 11.7, the fact that B^2_∞ is congruent to $\sqrt{2}\, B^2_1$, and that $B^m_1 \hat{\otimes} B^n_1$ identifies with B^{mn}_1 (cf. Exercise 11.8).

Exercise 11.5. Given $\xi \in \{-1,1\}^m$ and $\eta \in \{-1,1\}^n$, let $I = \{i \;:\; \xi_i = 1\}$ and $J = \{j \;:\; \eta_j = 1\}$ and split the overall sum $\sum b_{ij}\xi_i\eta_j$ into 4 sums according to whether $i \in I$ or not, $j \in J$ or not; then use the triangle inequality.

Exercise 11.6. For the first statement, note that $\{-1,1\}^k$ is exactly the set of extreme points of $B^k_\infty = (B^k_1)^\circ$. The second statement is even more straightforward from Proposition 11.8: $\{(x_i)_{i=1}^k \;:\; x_i \in \mathcal{H}, |x_i| \leqslant 1\}$ is exactly the unit ball of $\ell^k_\infty(\mathcal{H}) = (\ell^k_1(\mathcal{H}))^*$.

Exercise 11.7. Choose $\sigma = \tau = \sigma_z$, $\tilde{X}_i = X_i \otimes |0\rangle\langle 0|$, and $\tilde{Y}_j = Y_j \otimes |0\rangle\langle 0|$.

Exercise 11.8. First observe that Proposition 11.7 generalizes to the present context, with the same proof: $\mathsf{LC}_{2,\dots,2}$ identifies with $(B^2_\infty)^{\hat{\otimes} k}$. Next use the facts that B^2_∞ is congruent to $\sqrt{2}B^2_1$, and that $(B^2_1)^{\hat{\otimes} k}$ identifies with $B_1^{2^k}$. It follows that $\mathsf{LC}_{2,\dots,2}$ is congruent to $2^{k/2}B_1^{2^k}$, a polytope with 2^{k+1} vertices and 2^{2^k} facets.

Exercise 11.9. The answers are most conveniently deduced from the characterizations given by Propositions 11.7 and 11.8. The outradius is in both cases easily seen to be $\sqrt{mn}$. It is a little more delicate to establish that the inradii are 1. For the lower bound on the inradius of $\mathsf{LC}_{m,n} = B_\infty^m \hat{\otimes} B_\infty^n$, note that it is in Löwner position by Lemma 4.9 and then appeal to Exercise 4.20. For the remaining conclusions, use $\mathsf{LC}_{m,n} \subset \mathsf{QC}_{m,n} \subset B_\infty^{mn}$.

Exercise 11.10. Since $\mathsf{LC}_{m,n} = B_\infty^m \hat{\otimes} B_\infty^n$, this follows from Exercise 4.27 and the fact that a cube has enough symmetries (Exercise 4.25). More concretely, symmetries of LC are generated by permutations and sign flips of rows and columns. Since these operations are also symmetries for QC, it follows that QC has likewise enough symmetries.

Exercise 11.11. Taking into account Remark 11.9, it is enough to check that for every self-adjoint operator X_1, X_2, Y_1, Y_2 with $X_1^2 = X_2^2 = \mathrm{I}$ and $Y_1^2 = Y_2^2 = \mathrm{I}$, we have $\operatorname{Tr} \rho B \leqslant 2\sqrt{2}$, where $B = X_1 \otimes Y_1 + X_1 \otimes Y_2 + X_2 \otimes Y_1 - X_2 \otimes Y_2$. To that end, show that $B^2 = 4\mathrm{I} - (X_1X_2 - X_2X_1) \otimes (Y_1Y_2 - Y_2Y_1)$ and conclude that $\|B^2\|_{\mathrm{op}} \leqslant 8$. For an example giving violation $\sqrt{2}$, appeal to Proposition 11.8 and consider the case where $x_1, y_1, x_2, y_2 \in \mathbb{R}^2$ are unit vectors separated by successive 45° angles.

Here is an alternative argument which allows us to arrive at an example without guessing. First, observe that

$$\sup\{\varphi_{CHSH}(A) \;:\; A \in \mathsf{QC}_{2,2}\} = \frac{1}{2}\sup\{|y_1 + y_2| + |y_1 - y_2| \;:\; y_j \in \mathcal{H},\ |y_j| \leqslant 1\}$$

(cf. Exercise 11.6). Next, note that for such y_1, y_2,

$$|y_1 + y_2| + |y_1 - y_2| \leqslant \sqrt{2}\big(|y_1 + y_2|^2 + |y_1 - y_2|^2\big)^{1/2} = 2\big(|y_1|^2 + |y_2|^2\big)^{1/2} \leqslant 2\sqrt{2}$$

and verify when equalities occur.

Exercise 11.12. By Exercise 11.4 and its hint, every normal to a facet is proportional to the sum of four vertices of that facet, which in turn are of the form $A^{\xi,\eta}$. All such sums can then be listed and classified: there are 8 that exhibit the CHSH pattern and another 8 with only one non-zero entry. Alternatively, one may notice that every such sum is a matrix of Hilbert-Schmidt norm 4, whose entries are even integers that sum up to ± 4. Finally, the functionals corresponding to matrices with only one non-zero entry cannot distinguish between classical and quantum correlations.

Exercise 11.13. If $m > n$, the set $\mathsf{LC}_{n,n}$ can be seen in a canonical way as a section of $\mathsf{LC}_{m,n}$, which in turn is a section of $\mathsf{LC}_{m,m}$, and similarly for $\mathsf{QC}_{n,n}$, $\mathsf{QC}_{m,n}$ and $\mathsf{QC}_{m,m}$. The fact that $\mathbf{K}_G^{(2)} \geqslant \sqrt{2}$ follows from Exercise 11.11, and the opposite inequality by combining Exercises 11.11 and 11.12.

Exercise 11.14. We have $\mathsf{LC}_{2,n} = B_\infty^2 \hat{\otimes} B_\infty^n$. Since B_∞^2 is congruent to $\sqrt{2}B_1^2$, it follows that $\frac{1}{\sqrt{2}}\mathsf{LC}_{2,n}$ is congruent to $B_1^2 \hat{\otimes} B_\infty^n$, which identifies with $B_\infty^n \oplus_1 B_\infty^n :=$ $\operatorname{conv}\big(\{(x,0) \;:\; x \in B_\infty^n\} \cup \{(0,x) \;:\; x \in B_\infty^n\}\big)$. The facets of $B_\infty^n \oplus_1 B_\infty^n$ are of the form $\operatorname{conv}(F \times \{0\}, \{0\} \times G)$, where F, G are facets of B_∞^n. (This can be easily seen by identifying $(B_\infty^n \oplus_1 B_\infty^n)^\circ$ with $B_1^n \times B_1^n$.) It follows that $\mathsf{LC}_{2,n}$ has $(2n)^2$ facets: $4n$ facets express the fact that each entry of a correlation matrix belongs to $[-1,1]$, and $8\binom{n}{2} = 4n^2 - 4n$ are equivalent to the CHSH inequality.

Exercise 11.15. Fix $1 \leqslant i, j \leqslant 3$ and denote $E \subset \mathbb{R}^3 \times \mathbb{R}^3$ the subspace of matrices for which the ith row and the jth column are zero. It is clear from the definition that $P_E \mathsf{LC}_{3,3} = \mathsf{LC}_{2,2}$, where E is identified with $\mathbb{R}^2 \times \mathbb{R}^2$. It follows that whenever $\{\phi(\cdot) \leqslant 1\}$ is a facet-defining inequality for $\mathsf{LC}_{2,2}$, then $\{\phi(P_E(\cdot)) \leqslant 1\}$ is a facet-defining inequality for $\mathsf{LC}_{3,3}$. A careful counting (cf. Exercise 11.12) shows that this construction produces 18 facets of the kind $\pm a_{ij} \leqslant 1$ and $9 \times 8 = 72$ facets defined by inequalities equivalent to CHSH up to symmetries. The information that $\mathsf{LC}_{3,3}$ has 90 facets implies that $\mathsf{LC}_{3,3}$ is the intersection of the half-spaces associated to these $18 + 72 = 90$ facets. Since $P_E \mathsf{QC}_{3,3} = \mathsf{QC}_{2,2} \subset \sqrt{2}\mathsf{LC}_{2,2}$, it follows that $\mathsf{QC}_{3,3} \subset \sqrt{2}\mathsf{LC}_{3,3}$.

Exercise 11.16. If $M = (m_{ij})$, then

$$\|M : \ell_\infty^2(\mathbb{C}) \to \ell_1^2(\mathbb{C})\| = \max\Big\{ \sum_{i=1}^{m} \Big| \sum_{j=1}^{n} m_{ij} z_j \Big| \ : \ z_j \in \mathbb{C},\ |z_j| \leqslant 1,\ j = 1, \ldots, m \Big\}.$$

Since, as a real normed space, $(\mathbb{C}, |\cdot|)$ coincides with $(\mathbb{R}^2, |\cdot|)$, it remains to appeal to Exercise 11.6. (Note that we are concerned here with the case $m = n = 2$, but a similar argument works if $\min\{m, n\} = 2$.)

Exercise 11.17. Let $a, b, c, d \in \mathbb{C}$ and let $\phi : \mathbb{C} \to \mathbb{R}_+$ be defined by $\phi(z) = |az + b| + |cz + d|$. Then ϕ is convex and, in particular, its maximal value over the (closed) unit disk is attained on its boundary $\mathbb{T}$. Next, note that for $\eta_1, \eta_2 \in \mathbb{T}$ we have

$$|a\eta_1 + b\eta_2| + |c\eta_1 + d\eta_2| = \phi(\eta_1 \overline{\eta_2})$$

and, similarly, for $y_1, y_2 \in \mathbb{C}^2$ with $|y_1| = |y_2| = 1$,

$$|ay_1 + by_2| + |cy_1 + dy_2| = \phi(\langle y_1 | y_2 \rangle).$$

By the first observation, the maxima of these two expressions (over $\eta_1, \eta_2 \in \mathbb{T}$ and over unit vectors $y_1, y_2 \in \mathbb{C}^2$ respectively) coincide and it remains to notice that these maxima represent the expressions on the two sides of the inequality (11.37) for $[m_{ij}] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Exercise 11.18. The polytope $\mathsf{LC}_{n,n}$ is a symmetric polytope with 2^{2n-1} vertices and dimension n^2 (see Proposition 11.7), so the result follows. For the "moreover" part, combine Exercise 7.15 and, if needed, Theorem 11.12. Note that, from general principles (see Exercise 4.20), $a(\mathsf{LC}_{n,n}) \leqslant n$ and $a(\mathsf{QC}_{n,n}) \leqslant n$ (in fact we have equality by Exercise 11.9).

Exercise 11.19. Via Santaló inequality and its reverse, Proposition 11.15 implies that $\operatorname{vrad}(\mathsf{LC}_{n,n}^\circ) = \Theta(1/\sqrt{n})$. Since $\operatorname{outrad}(\mathsf{LC}_{n,n}^\circ) = \operatorname{inrad}(\mathsf{LC}_{n,n}) = 1$ (see Exercise 11.9). Proposition 6.3 implies that $\mathsf{LC}_{n,n}^\circ$ has $\exp(\Omega(n))$ vertices, or equivalently that $\mathsf{LC}_{n,n}$ has $\exp(\Omega(n))$ facets.

Exercise 11.20. (a) The value of the game is $\sum_{i,j} \pi(i,j) m_{ij} \xi_i \eta_j$, where $(\pi(i,j))$ is the distribution on the set of inputs. If $\pi(i_0, j_0) < \frac{1}{4}$, choose ξ, η so that $(\xi_i \eta_j)$ agrees with (m_{ij}) except for the (i_0, j_0)th entry. (b) First, replacing (ξ, η) by $(-\xi, -\eta)$ does not change the outcome, so for each such pair of strategies only the sum of their probabilities matters. Next, there are four pairs of that kind that saturate (11.4) and (11.12), with each pair leading to a mismatch in exactly one of the four entries of the 2×2 matrices $(\xi_i \eta_j)$ and (m_{ij}). If one of these four pairs entered into the random strategy with a weight strictly larger than $\frac{1}{4}$, the referee could use as

the setting (i, j) the index of the corresponding mismatched entry.
The combination of (a) and (b) describes the von Neumann–Nash-type equilibrium for the CHSH game.

Exercise 11.21. Alice and Bob have a quantum strategy which gives the value of at least $\frac{\sqrt{2}}{2}$ independently of the distribution $(\pi(i, j))$ on the set of inputs; moreover, if that distribution is not uniform, they have a quantum strategy yielding a value strictly larger than $\frac{\sqrt{2}}{2}$. For the universal strategy, use the same x_i, y_j as those implicit in the hint to Exercise 11.11; it follows from the argument there that, when expressed in terms of x_i, y_j, such strategy is unique up to isometries of the Hilbert space in question. If $(\pi(i, j))$ is not uniform, then either $|\pi(1,1)y_1 + \pi(1,2)y_2| + |\pi(2,1)y_1 - \pi(2,2)y_2|$ or $|\pi(1,1)x_1 + \pi(2,1)x_2| + |\pi(1,2)x_1 - \pi(2,2)x_2|$ is strictly larger than $2\sqrt{2}$.

Exercise 11.22. Extreme points of the set $\mathrm{K}_{k,m}$ defined in (11.23) are deterministic distributions that are of the form $p(\xi|i) = \delta_{\xi,f(i)}$ for some function f. It follows from the Krein–Milman theorem that any conditional probability distribution is a convex combination of deterministic distributions. Since $\mathsf{LB} = \mathrm{K}_{k,m} \hat{\otimes} \mathrm{K}_{l,n}$, the result follows.

Exercise 11.23. Consider $\lambda \in (0,1)$ and two boxes $P, \bar{P} \in \mathsf{QB}$. Represent $P = \{p(\xi, \eta|i, j)\}$ as $\{\operatorname{Tr} \rho(E_i^\xi \otimes F_j^\eta)\}$ and $\bar{P} = \{\bar{p}(\xi, \eta|i, j)\}$ as $\{\operatorname{Tr} \bar{\rho}(\bar{E}_i^\xi \otimes \bar{F}_j^\eta)\}$, where the operators $E_i^\xi, F_j^\eta, \bar{E}_i^\xi, \bar{F}_j^\eta$ act respectively on the Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B, \bar{\mathcal{H}}_A, \bar{\mathcal{H}}_B$. Verify that

$$\lambda p(\xi, \eta|i, j) + (1-\lambda)\bar{p}(\xi, \eta|i, j) = \operatorname{Tr}\Big(\sigma\Big((E_i^\xi \oplus \bar{E}_i^\xi) \otimes (F_j^\eta \oplus \bar{F}_j^\eta)\Big)\Big),$$

where $\sigma = \lambda\rho \oplus (1-\lambda)\bar{\rho}$ is a state acting on the diagonal subspace $\mathcal{H}_A \otimes \mathcal{H}_B \oplus \bar{\mathcal{H}}_A \otimes \bar{\mathcal{H}}_B \subset (\mathcal{H}_A \oplus \bar{\mathcal{H}}_A) \otimes (\mathcal{H}_B \oplus \bar{\mathcal{H}}_B)$.

Exercise 11.24. Replace ρ by its appropriate purification (see Section 3.4), i.e., represent $\rho \in \mathcal{H}_A \otimes \mathcal{H}_B$ as $\rho = \operatorname{Tr}_{\mathcal{H}_C} |\psi\rangle\langle\psi|$ for some $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. Then write $\operatorname{Tr} \rho(E_i^\xi \otimes F_j^\eta) = \langle\psi|E_i^\xi \otimes \overline{F}_j^\eta|\psi\rangle$, where $\overline{F}_j^\eta = F_j^\eta \otimes \mathrm{I}_{\mathcal{H}_C}$.

Exercise 11.25. (i) By Exercise 11.23, it is enough to show that $\mathsf{RB} \subset \mathsf{QB}$, which is easy. Note that a product box $P \in \mathsf{RB}$ can be represented in a trivial way: take $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}$, $\rho = \mathrm{I}_{\mathbb{C}\otimes\mathbb{C}}$ and $E_i^\xi = p(\xi|i)\,\mathrm{I}_{\mathbb{C}}$, $F_j^\eta = p(\eta|j)\,\mathrm{I}_{\mathbb{C}}$. (ii) Consider a local box of the form (11.20). By Carathéodory's theorem, we may assume that the index set Λ is finite. To obtain a representation as a quantum box with a separable state, consider $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^\Lambda$ and let $(|\lambda\rangle)_{\lambda\in\Lambda}$ be the canonical basis in $\mathbb{C}^\Lambda$. Define $\rho = \sum_\lambda \mu(\lambda)|\lambda\rangle\langle\lambda| \otimes |\lambda\rangle\langle\lambda|$, $E_i^\xi = \sum_\lambda p(\xi|i,\lambda)|\lambda\rangle\langle\lambda|$ and $F_j^\eta = \sum_\lambda p(\eta|j,\lambda)|\lambda\rangle\langle\lambda|$. One checks then that the representation (11.21) holds. Note that this construction is essentially the argument used in Exercise 11.23 to prove convexity of QB, specified to the present (simpler) setting.

Exercise 11.26. Since LB is convex, it suffices to prove the result when ρ is a product state, in which case it is almost immediate.

Exercise 11.27. Use (11.24) in combination with Exercises 4.13 and 4.15. Note that the affine space $V_{k,m}$ generated by $\mathrm{K}_{k,m}$ does not contain 0 and similarly for $\mathrm{K}_{l,n}$.

Exercise 11.28. If $p_A(\cdot|i) \in \mathrm{K}_{k,m}$ and $p_B(\cdot|j) \in \mathrm{K}_{l,n}$, the dimension of the set of boxes $P = \{p(\xi, \eta|i, j)\}$ verifying (11.25) for inputs i, j and for that particular choice of p_A, p_B is $(k-1)(l-1)$. Consequently, $\dim \mathsf{NSB} \leqslant mn(k-1)(l-1) + \dim \mathrm{K}_{k,m} +$

$\dim K_{l,n}$, which coincides with the value of $\dim \mathsf{LB}$ calculated in Exercise 11.27. Since $\mathsf{LB} \subset \mathsf{QB} \subset \mathsf{NSB}$, all dimensions must be the same. They are all convex sets with nonempty interior in the affine space $V_{k,m} \hat{\otimes} V_{l,n}$ analyzed in Exercise 11.27.

Exercise 11.29. Let $P = \{\operatorname{Tr} \rho(E_i^\xi \otimes F_j^\eta)\} \in \mathsf{QB}$ and $\bar{P} = \{\operatorname{Tr} \rho_*(E_i^\xi \otimes F_j^\eta)\}$, where ρ_* is the maximally mixed state. Since ρ_* is an interior point of Sep, it follows from Exercise 11.26 that the intersection of the segment $[\bar{P}, P]$ with LB is a segment of nonzero length, in particular P belongs to the affine subspace generated by LB. Since $P \in \mathsf{QB}$ was arbitrary, we conclude that QB is contained in that subspace and, in particular, $\dim \mathsf{QB} \leqslant \dim \mathsf{LB}$. (The converse inequality is trivial.)

Exercise 11.30. If $H \subset \mathbb{R}^N$ is an affine subspace not containing 0 and if V is an affine functional on $\mathbb{R}^N$, then there exists $v \in \mathbb{R}^N$ such that $\langle v, x \rangle = V(x)$ for $x \in H$.

Exercise 11.31. The first part is straightforward from the definitions. For the second part, note that we cannot have $\mathsf{LB}_\wp \subset b\mathsf{QB}_\wp$ if $|b| < 1$, and then appeal to the first part.

Exercise 11.32. (i) By Exercise 11.30, we can use affine functionals to exhibit violations. Given such functional V, the largest violation among functionals of the form $V_s = s + V$ (where $s \in \mathbb{R}$) occurs when $V_s(\mathsf{LB})$ is an interval of the form $[-a, a]$. Hence if V yields the maximal quantum violation, then

$$[-a, a] = V(\mathsf{LB}) \subset V(\mathsf{QB}) \subset [-a\omega_{\mathrm{Q}}(V), a\omega_{\mathrm{Q}}(V)]$$

and the last two intervals share (at least) one of the endpoints. In particular, the ratio of the lengths of the intervals $V(\mathsf{QB})$ and $V(\mathsf{LB})$ is between $(1 + \omega_{\mathrm{Q}}(V))/2$ and $\omega_{\mathrm{Q}}(V)$. (ii) Replace everywhere QB by NSB.

Exercise 11.33. First, the PR-box yields value 4 (in the normalization given by (11.29)). In the opposite direction, use the fact that, for each i, j, $p(\xi, \eta | i, j)$ is a joint density to deduce that $\left|\sum_{\xi, \eta} p(\xi, \eta | i, j)\right| \leqslant 1$. The second statement follows then from Exercise 11.15 and the proof of Proposition 11.19.

Exercise 11.34. Reverse engineer the proof of Proposition 11.8 starting from the configuration $x_1, y_1, x_2, y_2 \in \mathbb{R}^2$ from the hint to Exercise 11.11. This leads (for example) to ρ being the maximally entangled state on $\mathbb{C}^2 \otimes \mathbb{C}^2$, the isometries $X_1 = \sigma_x$, $X_2 = \sigma_z$ (the Pauli matrices), $Y_1 = 2^{-1/2}(\sigma_x + \sigma_z)$, $Y_2 = 2^{-1/2}(\sigma_x - \sigma_z)$ and, finally, to the POVMs consisting of spectral projections of X_i's and Y_j's (as in the formulas following (11.13)). The last step is somewhat tedious, but instructive.

Exercise 11.35. (i) The composition rules for Pauli matrices are in Exercise 2.4. (ii) Multiply all the numbers in the matrix. (iii)(a) Use part (ii); it follows that the probability of winning under any classical strategy is at most 8/9. (b) First, the product of the elements of Alice's output string must be an eigenvalue of the composition of the corresponding operators, and similarly for Bob, and therefore by (i)(b) their answers are valid. Next, we can compute (as in Section 3.8) the joint probability distribution of outcomes when Alice and Bob measure a single shared ϕ^+ in the eigenbasis of a Pauli matrix: for σ_x and σ_z both outcomes are always equal, and for σ_y both outcomes are always different. It follows that for each of the entries in Table 11.1, the outcomes of Alice's and Bob's measurements on $\phi_+ \otimes \phi_+$ always coincide.