Lecture 6

Nonlocal games and Tsirelson's theorem

In this lecture we will discuss *nonlocal games*, which offer a model through which the phenomenon of nonlocality is commonly studied. We will then narrow our focus to *XOR games*, which are a highly restricted form of nonlocal games that can, perhaps surprisingly, be analyzed through semidefinite programming. This is made possible by *Tsirelson's theorem*, which we will prove in this lecture.

6.1 Nonlocal games

We will begin by introducing the *nonlocal games* model. A nonlocal game is a hypothetical game in which two cooperating players, *Alice* and *Bob*, each receive a question from a *referee*, and then respond with an answer. The referee randomly selects the questions according to a known distribution, and, upon receiving answers from Alice and Bob, decides whether they win or lose. The following definition makes this notion precise in mathematical terms.

Definition 6.1. A *nonlocal game* is a 6-tuple $G = (X, Y, A, B, \pi, V)$, where

1. X, Y, A, and B are finite and nonempty sets,
2. $\pi \in \mathcal{P}(\mathcal{X} \times Y)$ is a probability vector, and
3. $V : A \times B \times X \times Y \rightarrow \{0, 1\}$ is a predicate.

In this definition, the sets X and Y are the sets of questions, and A and B are the sets of answers, for Alice and Bob, respectively. The probability vector π determines the probability with which each pair of questions $(x, y) \in X \times Y$ is selected by the referee, and V determines whether or not a pair of answers (a, b) wins or loses for a given pair of questions (x, y). For a given pair of questions

$(x,y) \in X \times Y$ and a pair of answers $(a,b) \in A \times B$, we write the value of the predicate as $V(a,b|x,y)$, because that's the way Ben Toner prefers it to be written—as it helps to stress the idea that (a,b) either wins or loses given that the question pair (x,y) was selected.

Example 6.2 (The CHSH game)**.** The CHSH game (named after Clauser, Horne, Shimony, and Holt) is a nonlocal game in which the questions and answers correspond to binary values, $X = Y = A = B = \{0,1\}$, the probability vector π is uniform,

$$\pi(0,0) = \pi(0,1) = \pi(1,0) = \pi(1,1) = \frac{1}{4}, \tag{6.1}$$

and the predicate V is defined as

$$V(a,b|x,y) = \begin{cases} 1 & \text{if } a \oplus b = x \wedge y \\ 0 & \text{if } a \oplus b \neq x \wedge y, \end{cases} \tag{6.2}$$

where $a \oplus b$ denotes the XOR of a and b, and $x \wedge y$ denotes the AND of x and y.

Intuitively speaking, if the referee selects any of the question pairs $(0,0)$, $(0,1)$, or $(1,0)$, then Alice and Bob must provide a pair of answers (a,b) for which $a = b$ in order to win, while if the referee selects the question pair $(1,1)$, the answer (a,b) wins when $a \neq b$.

Example 6.3 (The FFL game)**.** The FFL game (named after Fortnow, Feige, and Lovász) is a nonlocal game in which the questions and answers correspond to binary values, $X = Y = A = B = \{0,1\}$, the probability vector π is given by

$$\pi(0,0) = \pi(0,1) = \pi(1,0) = \frac{1}{3}, \quad \pi(1,1) = 0, \tag{6.3}$$

and the predicate V is defined as

$$V(a,b|x,y) = \begin{cases} 1 & \text{if } a \vee x \neq b \vee y \\ 0 & \text{if } a \vee x = b \vee y, \end{cases} \tag{6.4}$$

where $a \vee x$ denotes the OR of a and x, and similar for $b \vee y$.

Intuitively speaking, if the referee asks the question pair $(0,0)$, then exactly one of Alice and Bob, but not both, must respond with the answer 1 in order to win. However, if the question pair is either $(0,1)$ or $(1,0)$, then the player who received 0 must answer 0 to win (and it does not matter what the player who received the question 1 answers).

Example 6.4 (Graph coloring games)**.** Suppose that $H = (V, E)$ is an undirected graph and k is a positive integer. Let us also define $n = |V|$ and $m = |E|$, and assume $m \geq 1$. We may form a nonlocal game in the following way. The question sets are both equal to the set of vertices, $X = Y = \{1, \ldots, n\}$, and the answer sets are given by $A = B = \{1, \ldots, k\}$, which we may intuitively think about as colors. The probability vector π is defined as follows:

$$\pi(x, y) = \begin{cases} \frac{1}{2n} & \text{if } x = y \\ \frac{1}{4m} & \text{if } \{x, y\} \in E \\ 0 & \text{otherwise.} \end{cases} \tag{6.5}$$

In words, the referee flips a fair coin, and if the outcome is heads, it randomly selects a vertex and sends it to both players, and if the outcome is tails, it randomly selects an edge and then sends the two incident vertices to the two players (again at random). The predicate is defined as

$$V(a, b|x, y) = \begin{cases} 1 & \text{if } x = y \text{ and } a = b \\ 1 & \text{if } x \neq y \text{ and } a \neq b \\ 0 & \text{otherwise.} \end{cases} \tag{6.6}$$

The idea is that if Alice and Bob receive the same vertex, they should answer with the same color, while if they receive different (adjacent) vertices, they should answer with different colors.

Strategies

The definition of a nonlocal game does not, in itself, specify or restrict the sorts of strategies that Alice and Bob might employ when playing. There are, in fact, different types of strategies that are of interest. Let us start with a short summary of the strategy types that are of interest for this lecture.

1. *Deterministic strategies.* In a deterministic strategy, Alice must deterministically choose her answer a based on her question x alone, and likewise Bob must choose b based on y alone. A deterministic strategy may therefore be described as a pair of functions (f, g), where $f : X \to A$ and $g : Y \to B$.

   Notice that when we consider such a strategy, there is an implicit assumption that Alice cannot see Bob's question (or answer), and likewise Bob cannot see Alice's question (or answer). This sort of implicit assumption is also in place for the other strategy types listed below, and is what makes nonlocal games interesting and motivates their name.

2. *Randomized strategies.* Rather than choosing their answers deterministically, Alice and Bob could choose to make use of randomness when selecting their answers. The randomness could be in the form of local randomness, where Alice and Bob individually generate random numbers to assist in the selection of their answers, or it could be in the form of shared randomness, which one might view as having been generated by Alice and Bob at some point in the past.

   As it turns out, randomized strategies are not helpful to Alice and Bob, assuming their goal is to maximize the probability that they win. This is because randomized strategies can simply be viewed as the random selection of a deterministic strategy, and Alice and Bob might as well just select the optimal deterministic strategy—the average winning probability obviously cannot be larger than the maximum winning probability over all deterministic strategies.

3. *Entangled strategies.* An entangled strategy is one in which Alice and Bob make use of a shared quantum state when playing a nonlocal game. That is, Alice holds a register A and Bob holds a register B, where (A, B) is in a joint state $\rho \in \mathrm{D}(\mathcal{A} \otimes \mathcal{B})$, prior to the referee sending the questions. Upon receiving a question $x \in X$, Alice measures the register A with respect to a measurement described by a collection of measurement operators

   $$\left\{ P_a^x \,:\, a \in A \right\} \subset \mathrm{Pos}(\mathcal{A}), \tag{6.7}$$

   and likewise Bob measures B with respect to a measurement described by measurement operators

   $$\left\{ Q_b^y \,:\, b \in B \right\} \subset \mathrm{Pos}(\mathcal{B}). \tag{6.8}$$

   To be clear, Alice's measurement depends on her question $x \in X$ and Bob's measurement depends on his question $y \in Y$; they each have a measurement for each possible question they might receive.

   Given such a strategy, we see that the probability that Alice and Bob respond to a question pair (x, y) with an answer pair (a, b) is equal to

   $$\left\langle P_a^x \otimes Q_b^y, \rho \right\rangle. \tag{6.9}$$

   Note that ρ is not actually required to be entangled by the definition of an entangled strategy, but also note that if ρ is separable, then the strategy will be equivalent to a classical randomized strategy. So, entanglement is what makes this sort of strategy different from a classical strategy, which perhaps explains the name *entangled strategy*.

There are other types of strategies that are often considered in the study of nonlocal games, including *commuting operator strategies* and *no-signaling strategies*—we will discuss commuting operator strategies in the lecture following the next one. One could also consider *global strategies*, in which there is no implicit assumption that Alice and Bob are separated, so that (a, b) can depend arbitrarily on (x, y), but this class of strategies is not very interesting in a setting in which the nonlocality of Alice and Bob is relevant.

Values of games

When we speak of the *value* of a nonlocal game, we're referring to the supremum probability with which Alice and Bob can win the game, with respect to whatever class of strategies we might wish to consider. For this lecture we will focus on two values: the *classical value* and the *entangled value*.

Definition 6.5 (Classical value of a nonlocal game)**.** The *classical value* of a nonlocal game $G = (X, Y, A, B, \pi, V)$, which is denoted $\omega(G)$, is given by a maximization of the winning probability over all deterministic strategies:

$$\omega(G) = \max_{f,g} \sum_{(x,y)\in X\times Y} \pi(x,y) V\bigl(f(x), g(y)|x,y\bigr), \tag{6.10}$$

where the maximum is over all $f : X \rightarrow A$ and $g : Y \rightarrow B$.

Remark 6.6. As was previously discussed, the deterministic and randomized values of nonlocal games are the same—and so the name *classical value* is justified.

Definition 6.7 (Entangled value of a nonlocal game)**.** The *entangled value* of a nonlocal game $G = (X, Y, A, B, \pi, V)$, which is denoted $\omega^*(G)$, is the supremum of the winning probabilities

$$\sum_{(x,y)\in X\times Y} \pi(x,y) \sum_{(a,b)\in A\times B} V\bigl(a, b|x,y\bigr) \bigl\langle P_a^x \otimes Q_b^y, \rho \bigr\rangle, \tag{6.11}$$

over all choices of complex Euclidean spaces $\mathcal{A}$ and $\mathcal{B}$, states $\rho \in \mathrm{D}(\mathcal{A} \otimes \mathcal{B})$, and sets of measurements

$$\bigl\{P_a^x : a \in A\bigr\}_{x\in X} \subset \mathrm{Pos}(\mathcal{A}) \quad \text{and} \quad \bigl\{Q_b^y : b \in B\bigr\}_{y\in Y} \subset \mathrm{Pos}(\mathcal{B}). \tag{6.12}$$

That is, the entangled value is the supremum winning probability over all entangled strategies.

Remark 6.8. There are nonlocal games for which the winning probability is never achieved, so it is necessary to use the supremum in this definition. The principal issue is that the dimensions of the spaces $\mathcal{A}$ and $\mathcal{B}$ are not bounded as one ranges over all entangled strategies.

Example 6.9 (CHSH game values)**.** Letting G denote the CHSH game, we have that the classical value of this game is $\omega(G) = 3/4$. This may be verified by checking that the winning probability of each of the 16 possible deterministic strategies is at most 3/4, and of course that some of those strategies win with probability 3/4.

The entangled value of the CHSH game is $\omega^*(G) = \cos^2(\pi/8) \approx 0.85$. The fact that this is so will emerge as a simple corollary to Tsirelson's theorem—which is fitting given that the inequality $\omega^*(G) \leq \cos^2(\pi/8)$ is a rephrasing of an inequality known as Tsirelson's bound.

Example 6.10 (FFL game values)**.** If we let G denote the FFL game, then we have that its classical value and quantum value agree: $\omega(G) = \omega^*(G) = 2/3$. The fact that $\omega(G) = 2/3$ is easily established by testing all deterministic classical strategies. I will ask you to prove that $\omega^*(G) = 2/3$ as a homework problem. One way to do this is to prove that even the so-called *no-signaling value*, which upper-bounds the quantum value, of the FFL game is 2/3. The no-signaling value can be computed through linear programming.

Example 6.11 (Graph coloring game values)**.** If G is the graph coloring game determined by a graph H and an integer k, there we see that $\omega(G) = 1$ if and only if the chromatic number of H is at most k. That is, given any *perfect deterministic strategy*, meaning one that wins with certainty, it is possible to recover a k-coloring of H, meaning an assignment of colors $\{1, \ldots, k\}$ to the vertices of H such that no two adjacent vertices share the same color.

There are known examples of graphs H and choices of k for which the associated nonlocal game G satisfies $\omega(G) < 1$ but $\omega^*(G) = 1$.

6.2 XOR games

XOR games are a restricted type of nonlocal game $G = (X, Y, A, B, \pi, V)$ in which both players answer binary values, so that $A = B = \{0, 1\}$, and for which the predicate V takes the form

$$V(a, b|x, y) = \begin{cases} 1 & \text{if } a \oplus b = f(x, y) \\ 0 & \text{if } a \oplus b \neq f(x, y) \end{cases} \tag{6.13}$$

for some choice of a function $f : X \times Y \rightarrow \{0,1\}$. Intuitively speaking, the function f specifies whether a and b should agree or disagree in order to be a winning answer, for each question pair (x, y). Notice that exactly one of the two possibilities, meaning the possibilities that a and b agree or disagree, always wins for each question pair, while the other possibility loses.

As every XOR game is uniquely determined by the sets X and Y, the probability vector $\pi \in \mathcal{P}(X \times Y)$, and the function $f : X \times Y \rightarrow \{0,1\}$, we will identify the corresponding game G with the quadruple (X, Y, π, f) when it is convenient to do that. For example, the CHSH game is an example of an XOR game, corresponding to the quadruple $(\{0,1\}, \{0,1\}, \pi, f)$, for π the uniform probability vector and $f(x, y) = x \wedge y$ being the AND function.

Bias of an XOR game

When analyzing XOR games, it is often convenient to consider the *bias* of games rather than their value. For a given XOR game $G = (X, Y, \pi, f)$, and any strategy for G, we define the bias of that strategy, for that game, to be the probability it wins *minus* the probability it loses—which happens to be the same thing as twice the probability it wins minus 1. The *bias of a game* is defined to be the supremum bias over all strategies under consideration for that game. We will write $\varepsilon(G)$ and $\varepsilon^*(G)$ to denote the classical and quantum biases for G, and so we have

$$\varepsilon(G) = 2\omega(G) - 1 \quad \text{and} \quad \varepsilon^*(G) = 2\omega^*(G) - 1, \tag{6.14}$$

or, alternatively,

$$\omega(G) = \frac{1}{2} + \frac{\varepsilon(G)}{2} \quad \text{and} \quad \omega^*(G) = \frac{1}{2} + \frac{\varepsilon^*(G)}{2}. \tag{6.15}$$

XOR game strategies described by observables

Let $G = (X, Y, \pi, f)$ be an XOR game, and consider any entangled strategy for that game, represented by a state $\rho \in \mathrm{D}(\mathcal{A} \otimes \mathcal{B}\}$ and measurement operators

$$\{P_0^x, P_1^x\}_{x \in X} \subset \mathrm{Pos}(\mathcal{A}) \quad \text{and} \quad \{Q_0^y, Q_1^y\}_{y \in Y} \subset \mathrm{Pos}(\mathcal{B}). \tag{6.16}$$

If we consider the expression

$$\sum_{x,y \in X \times Y} \pi(x, y)(-1)^{f(x,y)} \big\langle \big(P_0^x - P_1^x\big) \otimes \big(Q_0^y - Q_1^y\big), \rho \big\rangle \tag{6.17}$$

for a few moments, we find that it agrees with the bias of the strategy just described. By defining $A_x = P_0^x - P_1^x$ for each $x \in X$ and $B_y = Q_0^y - Q_1^y$ for each

$y \in Y$, we may express this quantity as

$$\sum_{x,y \in X \times Y} \pi(x,y)(-1)^{f(x,y)} \langle A_x \otimes B_y, \rho \rangle. \tag{6.18}$$

The operators A_x and B_y may be viewed as representing *observables* in the parlance of quantum mechanics.

Notice that as one ranges over all binary-valued measurements $\{R_0, R_1\}$, the operator $R_0 - R_1$ ranges over all Hermitian operators H with $\|H\| \leq 1$. Therefore, the bias of a game G is given by the supremum value of the expression (6.18), over all choices of $\{A_x : x \in X\} \subset \mathrm{Herm}(\mathcal{A})$, $\{B_y : y \in Y\} \subset \mathrm{Herm}(\mathcal{B})$, and $\rho \in \mathrm{D}(\mathcal{A} \otimes \mathcal{B})$, subject to the constraints $\|A_x\| \leq 1$ for every $x \in X$ and $\|B_y\| \leq 1$ for every $y \in Y$.

6.3 Tsirelson's theorem

Now we will prove the theorem of Tsirelson mentioned previously. Let us begin with a statement of the theorem.

Theorem 6.12 (Tsirelson's theorem)**.** *For every choice of finite and nonempty sets X and Y and an operator $M \in \mathrm{L}(\mathbb{R}^Y, \mathbb{R}^X)$, the following statements are equivalent.*

1. *There exist complex Euclidean spaces $\mathcal{A}$ and $\mathcal{B}$, a density operator $\rho \in \mathrm{D}(\mathcal{A} \otimes \mathcal{B})$, and two collections $\{A_x : x \in X\} \subset \mathrm{Herm}(\mathcal{A})$ and $\{B_y : y \in Y\} \subset \mathrm{Herm}(\mathcal{B})$ of operators such that $\|A_x\| \leq 1$, $\|B_y\| \leq 1$, and*

$$M(x,y) = \langle A_x \otimes B_y, \rho \rangle \tag{6.19}$$

   for all $x \in X$ and $y \in Y$.

2. *There exist positive semidefinite operators $R \in \mathrm{Pos}(\mathbb{C}^X)$ and $S \in \mathrm{Pos}(\mathbb{C}^Y)$, with $R(x,x) = 1$ and $S(y,y) = 1$ for all $x \in X$ and $y \in Y$, such that*

$$\begin{pmatrix} R & M \\ M^* & S \end{pmatrix} \geq 0. \tag{6.20}$$

Remark 6.13. The second statement in the theorem is equivalent to one in which the requirement that R and S have real number entries is added. In particular, if R_0 and S_0 satisfy the conditions listed in the second statement of the theorem, then so too will

$$R = \frac{R_0 + R_0^{\mathsf{T}}}{2} \quad \text{and} \quad S = \frac{S_0 + S_0^{\mathsf{T}}}{2}, \tag{6.21}$$

by virtue of the fact that M has real-number entries and

$$\begin{pmatrix} R & M \\ M^* & S \end{pmatrix} = \frac{1}{2}\begin{pmatrix} R_0 & M \\ M^* & S_0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} R_0 & M \\ M^* & S_0 \end{pmatrix}^{\mathsf{T}} \tag{6.22}$$

is a positive semidefinite operator whose diagonal entries are all equal to 1.

The first statement of the theorem says that the operator M, which is best viewed as a matrix indexed by pairs $(x, y) \in X \times Y$ in this case, describes exactly the values in the expression (6.18) that depend upon the strategy under consideration. The second statement of the theorem is a suprisingly simple condition on M—and it may come at no surprise to learn that it will be used to define semidefinite programs to calculate XOR game biases. The fact that these two statements are exactly the same thing is a remarkable thing of beauty.

Weyl–Brauer operators

The proof of Tsirelson's theorem will make use of a collection of unitary and Hermitian operators known as *Weyl–Brauer operators*.

Definition 6.14. Let N be a positive integer and let $\mathcal{Z} = \mathbb{C}^2$. The *Weyl–Brauer operators of order N* are the operators $V_1, \ldots, V_{2N+1} \in \mathrm{L}(\mathcal{Z}^{\otimes N})$ defined as

$$\begin{aligned} V_{2k-1} &= \sigma_z^{\otimes(k-1)} \otimes \sigma_x \otimes \mathbb{1}^{\otimes(N-k)}, \\ V_{2k} &= \sigma_z^{\otimes(k-1)} \otimes \sigma_y \otimes \mathbb{1}^{\otimes(N-k)}, \end{aligned} \tag{6.23}$$

for all $k \in \{1, \ldots, N\}$, as well as

$$V_{2N+1} = \sigma_z^{\otimes N}, \tag{6.24}$$

where $\mathbb{1}$, σ_x, σ_y, and σ_z denote the Pauli operators:

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.25}$$

Example 6.15. In the case $N = 3$, the Weyl–Brauer operators $V_1, \ldots, V_7$ are

$$\begin{aligned} V_1 &= \sigma_x \otimes \mathbb{1} \otimes \mathbb{1} \\ V_2 &= \sigma_y \otimes \mathbb{1} \otimes \mathbb{1} \\ V_3 &= \sigma_z \otimes \sigma_x \otimes \mathbb{1} \\ V_4 &= \sigma_z \otimes \sigma_y \otimes \mathbb{1} \\ V_5 &= \sigma_z \otimes \sigma_z \otimes \sigma_x \\ V_6 &= \sigma_z \otimes \sigma_z \otimes \sigma_y \\ V_7 &= \sigma_z \otimes \sigma_z \otimes \sigma_z. \end{aligned} \tag{6.26}$$

A proposition summarizing the properties of the Weyl–Brauer operators that are relevant to the proof of Tsirelson's theorem follows.

Proposition 6.16. *Let N be a positive integer, let $V_1, \ldots, V_{2N+1}$ denote the Weyl–Brauer operators of order N. For every unit vector $u \in \mathbb{R}^{2N+1}$, the operator*

$$\sum_{k=1}^{2N+1} u(k) V_k \tag{6.27}$$

is both unitary and Hermitian, and for any two vectors $u, v \in \mathbb{R}^{2N+1}$, it holds that

$$\frac{1}{2^N}\left\langle \sum_{j=1}^{2N+1} u(j) V_j, \sum_{k=1}^{2N+1} v(k) V_k \right\rangle = \langle u, v \rangle. \tag{6.28}$$

Proof. Each operator V_k is Hermitian, and therefore the operator (6.27) is Hermitian as well.

The Pauli operators anti-commute in pairs:

$$\sigma_x \sigma_y = -\sigma_y \sigma_x, \quad \sigma_x \sigma_z = -\sigma_z \sigma_x, \quad \text{and} \quad \sigma_y \sigma_z = -\sigma_z \sigma_y. \tag{6.29}$$

By an inspection of the definition of the Weyl–Brauer operators, it follows that $V_1, \ldots, V_{2N+1}$ also anti-commute in pairs:

$$V_j V_k = -V_k V_j \tag{6.30}$$

for distinct choices of $j, k \in \{1, \ldots, 2N+1\}$. Moreover, each V_k is unitary (as well as being Hermitian), and therefore $V_k^2 = \mathbb{1}^{\otimes N}$. It follows that

$$\begin{aligned}\left(\sum_{k=1}^{2N+1} u(k) V_k\right)^2 &= \sum_{k=1}^{2N+1} u(k)^2 V_k^2 + \sum_{1 \leq j < k \leq 2N+1} u(j) u(k) \bigl(V_j V_k + V_k V_j\bigr) \\ &= \sum_{k=1}^{2N+1} u(k)^2 \mathbb{1}^{\otimes N} = \mathbb{1}^{\otimes N},\end{aligned} \tag{6.31}$$

and therefore (6.27) is unitary.

Next, observe that

$$\langle V_j, V_k \rangle = \begin{cases} 2^N & \text{if } j = k \\ 0 & \text{if } j \neq k. \end{cases} \tag{6.32}$$

Therefore, one has

$$\begin{aligned}&\frac{1}{2^N}\left\langle \sum_{j=1}^{2N+1} u(j) V_j, \sum_{k=1}^{2N+1} v(k) V_k \right\rangle \\ &\quad = \frac{1}{2^N} \sum_{j=1}^{2N+1} \sum_{k=1}^{2N+1} u(j) v(k) \langle V_j, V_k \rangle = \sum_{k=1}^{2N+1} u(k) v(k) = \langle u, v \rangle,\end{aligned} \tag{6.33}$$

as required. □

Proof of Tsirelson's theorem

Proof of Theorem 6.12. For the sake of simplifying notation, we will make the assumption that $X = \{1, \ldots, n\}$ and $Y = \{1, \ldots, m\}$.

Assume that the first statement is true, and define an operator

$$K = \begin{pmatrix} \operatorname{vec}\big((A_1 \otimes \mathbb{1})\sqrt{\rho}\big)^* \\ \vdots \\ \operatorname{vec}\big((A_n \otimes \mathbb{1})\sqrt{\rho}\big)^* \\ \operatorname{vec}\big((\mathbb{1} \otimes B_1)\sqrt{\rho}\big)^* \\ \vdots \\ \operatorname{vec}\big((\mathbb{1} \otimes B_m)\sqrt{\rho}\big)^* \end{pmatrix} \in \mathrm{L}(\mathcal{A} \otimes \mathcal{B} \otimes \mathcal{A} \otimes \mathcal{B}, \mathbb{C}^n \oplus \mathbb{C}^m). \tag{6.34}$$

The operator $KK^* \in \operatorname{Pos}\big(\mathbb{C}^n \oplus \mathbb{C}^m\big)$ may be written in a block form as

$$KK^* = \begin{pmatrix} P & M \\ M^* & Q \end{pmatrix} \tag{6.35}$$

for $P \in \operatorname{Pos}(\mathbb{C}^n)$ and $Q \in \operatorname{Pos}(\mathbb{C}^m)$; the fact that the off-diagonal blocks are as claimed follows from the calculation

$$\big\langle (A_j \otimes \mathbb{1})\sqrt{\rho}, (\mathbb{1} \otimes B_k)\sqrt{\rho} \big\rangle = \big\langle A_j \otimes B_k, \rho \big\rangle = M(j,k). \tag{6.36}$$

For each $j \in \{1, \ldots, n\}$ one has

$$P(j,j) = \big\langle (A_j \otimes \mathbb{1})\sqrt{\rho}, (A_j \otimes \mathbb{1})\sqrt{\rho} \big\rangle = \big\langle A_j^2 \otimes \mathbb{1}, \rho \big\rangle, \tag{6.37}$$

which is necessarily a nonnegative real number in the interval $[0,1]$; and through a similar calculation, one finds that $Q(k,k)$ is also a nonnegative integer in the interval $[0,1]$ for each $k \in \{1, \ldots, m\}$. A nonnegative real number may be added to each diagonal entry of this operator to yield another positive semidefinite operator, so one has that statement 2 holds.

Next, let us assume statement 2 holds. As was explained in Remark 6.13, we are free to assume that all of the entries of R and S are real numbers.

Now, a matrix with real number entries is positive semidefinite if and only if it is the Gram matrix of a collection of real vectors, and therefore there must exist real vectors $\{u_1, \ldots, u_n, v_1, \ldots, v_m\}$ such that

$$\langle u_j, v_k \rangle = M(j,k) \tag{6.38}$$

for all $j \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, m\}$, as well as

$$\langle u_{j_0}, u_{j_1} \rangle = R(j_0, j_1) \quad \text{and} \quad \langle v_{k_0}, v_{k_1} \rangle = S(k_0, k_1) \tag{6.39}$$

for all $j_0, j_1 \in \{1,\ldots,n\}$ and $k_0, k_1 \in \{1,\ldots,m\}$. There are $n+m$ of these vectors, and therefore they span a real vector space of dimension at most $n+m$, so there is no loss of generality in assuming $u_1,\ldots,u_n,v_1,\ldots,v_m \in \mathbb{R}^{n+m}$. Observe that these vectors are all unit vectors, as the diagonal entries of R and S represent their norm squared.

Choose N so that $2N+1 \geq n+m$ and let $\mathcal{Z} = \mathbb{C}^2$. Define operators $A_1,\ldots,A_n$ and $B_1,\ldots,B_m$, all acting on $\mathrm{L}(\mathcal{Z}^{\otimes N})$, as

$$A_j = \sum_{i=1}^{n+m} u_j(i) V_i \quad \text{and} \quad B_k = \sum_{i=1}^{n+m} v_k(i) V_i^{\mathsf{T}} \tag{6.40}$$

for each $j \in \{1,\ldots,n\}$ and $k \in \{1,\ldots,m\}$, where $V_1,\ldots,V_{n+m}$ are the first $n+m$ Weyl–Brauer operators of order N. By Proposition 6.16, each of these operators is both unitary and Hermitian, and therefore each of these operators has spectral norm equal to 1.

Finally, define

$$\rho = \frac{1}{2^N} \operatorname{vec}\big(\mathbb{1}^{\otimes N}\big) \operatorname{vec}\big(\mathbb{1}^{\otimes N}\big)^* \in \mathrm{D}\big(\mathcal{Z}^{\otimes N} \otimes \mathcal{Z}^{\otimes N}\big). \tag{6.41}$$

Applying Proposition 6.16 again gives

$$\big\langle A_j \otimes B_k, \rho \big\rangle = \frac{1}{2^N} \big\langle A_j, B_k^{\mathsf{T}} \big\rangle = \langle u_j, v_k \rangle = M(j,k), \tag{6.42}$$

for each $j \in \{1,\ldots,n\}$ and $k \in \{1,\ldots,m\}$.

We have proved that statement 2 implies statement 1, for the spaces $\mathcal{A} = \mathcal{Z}^{\otimes N}$ and $\mathcal{B} = \mathcal{Z}^{\otimes N}$, and so the proof is complete. □