QIC 890 Entanglement and Nonlocal Effects
Lecture Notes

R. Cleve

April 22, 2019

Abstract

These are somewhat rough notes to accompany the course lectures. They include corrections and improvements that have been suggested by colleagues and students in the class (an acknowledgements section provides more detail).

Contents

1 Review of some basic definitions

1.1 Quantum states

A (pure) quantum state is a vector ψ in a Hilbert space of norm 1 (i.e., $\|\psi\| = 1$). For starters, we will only consider finite-dimensional Hilbert spaces, which are without loss of generality, characterized by their dimension. Thus, we can identify a d-dimensional Hilbert space with $\mathbb{C}^d$. The *computational basis* is referred to as $|0\rangle, \ldots, |d-1\rangle$.

When two separate quantum systems are considered as one (say, the state in Alice's lab combined with the state in Bob's lab), the joint Hilbert space is given by the tensor product of Alice's Hilbert space with that of Bob. If they are both d-dimensional then the computational basis of the tensor product consists of states of the form $|j\rangle \otimes |k\rangle$, where $j, k \in \{0, 1, \ldots, d-1\}$. When there is no ambiguity, it is common to leave out the $\otimes$ and just write $|j\rangle|k\rangle$. Other alternative notations are $|j, k\rangle$ and $|jk\rangle$. In the last one, "jk" is the concatination of the digits j and k (not their product!), and this can be read as a two-digit number in base d representing an element of $\{0, 1, \ldots, d^2-1\}$ written in base d. Thus, $\mathbb{C}^d \otimes \mathbb{C}^d$ is the same as $\mathbb{C}^{d^2}$.

1.2 Measurements

A projective measurement is defined by a set of complete orthogonal projectors $\Pi_1, \ldots, \Pi_m$. (Being *projectors* means $\Pi_k^2 = \Pi_k$; being *orthogonal* means $\Pi_j \Pi_k = 0$ for all $j \neq k$; and being *complete* means $\sum_{k=1}^m \Pi_k = I$). When the state is $\psi \in \mathcal{H}$ (where $\mathcal{H}$ is the space on which the projectors act), the corresponding measurement operation produces outcome k with probability $\langle\psi|\Pi_k|\psi\rangle$ (and the state is destroyed).

A POVM measurement[1] is defined by a set of positive[2] operators $A_1, ..., A_m$ (i.e., $A_k \geq 0$ for each k) such that $\sum_k A_k = I$. When state is ψ, the correspinding measurement outcome k occurs with probability $\langle\psi|A_k|\psi\rangle$.

Note about notation

We sometimes use the so-called ket notation that is prevalent in physics, where $|\psi\rangle = \psi$ and $\langle\psi| = \psi^*$. Also, $\langle\psi|\Pi_k|\psi\rangle = \langle\psi, \Pi_k\psi\rangle$. Moreover, A^*, which denotes the adjoint (conjugate-transpose) of A is sometimes denoted as $A^\dagger$.

1.3 Nonlocal games

Following [4], a *nonlocal game* is defined as $G = (S, T, A, B, \pi, V)$, where S, T, A, B are finite sets, π is a probability distribution on $S \times T$ and $V : A \times B \times S \times T \to \mathbb{R}$. V is the *payoff function* and in many nonlocal games of interest it is $\{0, 1\}$-valued (where 1 means "win" and 0 means "lose").

We think of such a game as a process, where there are two cooperating parties, usually called Alice and Bob, who are restricted so as not being able to communicate with each other. They are each provided a question as input ($s \in S$ for Alice and $t \in T$ for Bob). The question pair $(s, t) \in S \times T$ is generated according to the probability distribution π. After receiving their questions, they are required to produce answers as output ($a \in A$ for Alice and $b \in B$ for Bob). The value that they attain is given by the payoff function evaluated at their inputs/outputs, i.e., $V(a, b, s, t)$.

The prohibition on communication means Alice doesn't know Bob's question and Bob doesn't know Alice's question. This restriction is what makes these games interesting, as we'll see in the examples coming up.

Classical strategies

We define deterministic classical strategies as functions $a : S \to A$ and $b : T \to B$. On inputs $(s, t) \in S \times T$, Alice outputs $a(s)$ and Bob outputs $b(t)$.

(Note that we could also consider probabilistic strategies, but, using a convexity argument, it can be shown that these do not increase the value; there is always an optimal strategy that is deterministic.)

Classical value

The classical value (sometimes denoted as $\omega(G)$ or as $\omega_c(G)$) is

$$\omega_c(G) = \sum_{(s,t)\in S\times T} \pi(s, t)V(a(s), b(t), s, t). \tag{1}$$

Entangled strategies (a.k.a, quantum strategies)

Here, we assume that Alice and Bob are using entanglement from some tensor product of two finite-dimensional Hilbert spaces, $\mathcal{H}_A$ and $\mathcal{H}_B$ (say $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^d$). (Other, more exotic, notions of entanglement may be discussed later on.)

[1] POVM stands for "positive operator-valued measure", but this is not important for us here.

[2] Operator A is positive (a.k.a., positive semidefinite) if there exists an operator B such that $A = B^*B$.

We assume that Alice and Bob start with some entangled state $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$. For each input $s \in S$ to Alice, there is a POVM measurement $(A_s^a)_{a \in A}$, which Alice performs when input $s \in S$ is received. This generates an output $a \in A$ for Alice's output. Similarly, for each $t \in T$ to Bob, there is a POVM measurement $(B_t^b)_{b \in B}$, which Bob performs when input $t \in T$ is received. This generates an output $b \in B$ for Bob's output.

The value attained by any particular such strategy is

$$\sum_{(s,t) \in S \times T} \pi(s,t) \sum_{(a,b) \in A \times B} \langle \psi | A_s^a \otimes B_t^b | \psi \rangle \, V(a,b,s,t). \tag{2}$$

Entangled value (a.k.a., quantum value)

The engangled value of a game G (sometimes denoted as $\omega^*(G)$ or as $\omega_q(G)$) is the supremum of Eq. (2), over all possible entangled strategies.

1.4 Observables

An observable is a Hermitian operator, and it is useful to associate it with the following measurement process. Every Hermitian operator A can be expressed as

$$A = \sum_{k=1}^{m} \lambda_k \Pi_k, \tag{3}$$

where $\Pi_1, ..., \Pi_m$ are projectors and $\lambda_1, ..., \lambda_m \in \mathbb{R}$ are the unique eigenvalues. The measurement process is to perform the projective measurement with projectors $\Pi_1, ..., \Pi_m$ where, on outcome k, the information that is output is λ_k. The expected value of the outcome is $\langle \psi | A | \psi \rangle$.

Tensor products of observables

If A is an observable on the Hilbert space $\mathcal{H}_A$ and B is an observable on the Hilbert space $\mathcal{H}_B$ then $A \otimes B$ is an observable on the space $\mathcal{H}_A \otimes \mathcal{H}_B$. The expected value of this observable is the same as that of the following operational process. Separately measure the two systems, yielding an eigenvalue of each one ($\lambda_A \in \mathbb{R}$ for the first system and $\lambda_B \in \mathbb{R}$ for the second system), and then multiply the two outcomes to obtain $\lambda_A \lambda_B \in \mathbb{R}$.

If ψ is a state on the joint system then the expected value of the above process is given by $\langle \psi | A \otimes B | \psi \rangle$.

Binary observables and the bias of a binary observable

A *binary observable* is one whose eigenvalues are in $\{+1, -1\}$. Examples of observables are the Pauli matrices

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4}$$

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{5}$$

$$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{6}$$

Note that: for Z, the measurement is with respect to the computational basis $|0\rangle$ and $|1\rangle$; for X, the measurement is with respect to the Hadamard basis, often denoted as $|+\rangle$, $|-\rangle$, where

$$|+\rangle := \tfrac{1}{\sqrt{2}}|0\rangle + \tfrac{1}{\sqrt{2}}|1\rangle \tag{7}$$

$$|-\rangle := \tfrac{1}{\sqrt{2}}|0\rangle - \tfrac{1}{\sqrt{2}}|1\rangle; \tag{8}$$

and, for Y, the measurement is with respect to the basis

$$|+i\rangle := \tfrac{1}{\sqrt{2}}|0\rangle + \tfrac{i}{\sqrt{2}}|1\rangle \tag{9}$$

$$|-i\rangle := \tfrac{1}{\sqrt{2}}|0\rangle - \tfrac{i}{\sqrt{2}}|1\rangle. \tag{10}$$

Note that

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{11}$$

$$H = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \tag{12}$$

are also binary observables.

For a binary observable A, it is straightforward to calculate that the probability of outcome $+1$ is $p_+ = \frac{1+\langle\psi|A|\psi\rangle}{2}$ and the probability of outcome -1 is $p_- = \frac{1-\langle\psi|A|\psi\rangle}{2}$. The *bias towards* $+1$ is defined as $p_+ - p_- = \langle\psi|A|\psi\rangle$ (this is a quantity in $[-1, +1]$).

1.5 Frobenius inner product

For $d \times d$ matrices A and B, define their *Frobenius inner product* as $\mathrm{Tr}(A^*B)$. Note that

$$\mathrm{Tr}(A^*B) = \sum_{j=1}^{d}\sum_{k=1}^{d} \overline{A}_{jk} B_{jk}. \tag{13}$$

Therefore, $\mathrm{Tr}(A^*B)$ is the "dot product" of A and B regarded as d^2-dimensional vectors.

1.6 Maximally entangled states

In this section, $|\psi\rangle$ refers to the following maximally entangled state

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle|k\rangle. \tag{14}$$

Also, let A^T denote the *transpose* of A (with respect to the computational basis) and $\overline{A}$ refers to the *element-wise conjugate* of A (also with respect to the computational basis).

The following are straightforward to verify by calculation.

Lemma 1.1. *For any $A \in \mathbb{C}^{d\times d}$, $(A \otimes I)|\psi\rangle = (I \otimes A^T)|\psi\rangle$.*

Lemma 1.2. *For any $A \in \mathbb{C}^{d\times d}$, $(A \otimes I)|\psi\rangle = \frac{1}{d}\mathrm{Tr}(A)$.*

Corollary 1.3. *For any $A, B \in \mathbb{C}^{d\times d}$, $\langle\psi|(\overline{A} \otimes B)|\psi\rangle = \frac{1}{d}\mathrm{Tr}(A^*B)$.*

Analysis of what happens when $\psi = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$

For the Pauli operators X, Y, Z, we have

$$\langle\psi|Z \otimes Z|\psi\rangle = +1 \tag{15}$$
$$\langle\psi|X \otimes X|\psi\rangle = +1 \tag{16}$$
$$\langle\psi|Y \otimes Y|\psi\rangle = -1, \tag{17}$$

where the last equation is because $\overline{Y} = -Y = Y^T$. This can be verified by calculating that

$$Z \otimes Z|\psi\rangle = |\psi\rangle \tag{18}$$
$$X \otimes X|\psi\rangle = |\psi\rangle \tag{19}$$
$$Y \otimes Y|\psi\rangle = -|\psi\rangle. \tag{20}$$

To get rid of the asymmetry arising for the case of Y, we can get $+1$ in all three equations, by taking the transpose of the second Pauli in each equation, so that

$$\langle\psi|Z \otimes Z^{\mathrm{T}}|\psi\rangle = \langle\psi|X \otimes X^{\mathrm{T}}|\psi\rangle = \langle\psi|Y \otimes Y^{\mathrm{T}}|\psi\rangle = +1. \tag{21}$$

2 The CHSH game

In this game (named after the authors of [3]), the input and output alphabets are $\{0,1\}$ (i.e., $S = T = A = B = \{0,1\}$), π is the uniform distribution, and

$$V(a,b,s,t) = \begin{cases} 1 & \text{if } a \oplus b = s \wedge t \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

The classical value of CHSH is $3/4$. To see why this is so, note that every (deterministic) classical strategy is characterized by four bits, a_0, a_1, b_0, b_1 (Alice's output on input s is a_s and Bob's output on input t is b_t).

The winning conditions for the four possible questions are

$$a_0 \oplus b_0 = 0 \tag{23}$$
$$a_0 \oplus b_1 = 0 \tag{24}$$
$$a_1 \oplus b_0 = 0 \tag{25}$$
$$a_1 \oplus b_1 = 1. \tag{26}$$

By summing the left and right sides of these equations in mod 2 arithmetic, we get $0 = 1$, which implies that it is impossible to satisfy all four equations simultaneously. However, it is possible to satisfy any three of the four equations, which leads to a strategy that succeeds with probability $3/4$. Therefore, $\omega_c(\mathrm{CHSH}) = 3/4$.

Sometimes it is convenient to refer to the *bias* (towards 1) of a strategy rather than its winning probability, where the bias is defined as the winning probability minus the losing probability. This is denoted as $\beta(G)$ and we have $\beta(\mathrm{CHSH}) = 3/4 - 1/4 = 1/2$. The bias arises by considering the CHSH game in its multiplicative form where $S = T = \{0,1\}$, $A = B = \{+1,-1\}$ and

$$V(a,b,s,t) = \begin{cases} ab & \text{if } s \wedge t = 0 \\ -ab & \text{otherwise.} \end{cases} \tag{27}$$

Clearly, the value of this version of CHSH is $1/2$.

2.1 Entangled strategy for CHSH

Now we consider the entangled value of CHSH and show that $\omega_q(\text{CHSH}) \geq \frac{1}{2} + \frac{\sqrt{2}}{4}$. Note that this is equivalent to the bias being $\frac{1}{\sqrt{2}}$ and we will analyze the CHSH game in this form (that is, the value of CHSH in the multiplicative form, where the outputs are in $\{+1, -1\}$). Let $\psi \in \mathbb{C}^d \otimes \mathbb{C}^d$ be the entangled state used. Then any strategy corresponds to four d-dimensional binary observables A_0, A_1, B_0, B_1 (Alice's observable on input s is A_s and Bob's observable on input t is B_t). Note that, for input $(s,t) \in S \times T$, the expected payoff is

$$(-1)^{s \wedge t} \langle \psi | A_s \otimes B_t | \psi \rangle. \tag{28}$$

Therefore, the classical bias is

$$\beta_c(\text{CHSH}) = \frac{1}{4} \langle \psi | A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1 | \psi \rangle. \tag{29}$$

Consider the entangled strategy that uses the entangled state $\psi = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$ and observables

$$A_0 = Z \tag{30}$$
$$A_1 = X \tag{31}$$
$$B_0 = H \tag{32}$$
$$B_1 = -XHX = ZHZ. \tag{33}$$

We can calculate that

$$\langle \psi | A_0 \otimes B_0 | \psi \rangle = \tfrac{1}{2}\text{Tr}(ZH) = \tfrac{1}{\sqrt{2}} \tag{34}$$
$$\langle \psi | A_0 \otimes B_1 | \psi \rangle = \tfrac{1}{2}\text{Tr}(Z\,ZHZ) = \tfrac{1}{\sqrt{2}} \tag{35}$$
$$\langle \psi | A_1 \otimes B_0 | \psi \rangle = \tfrac{1}{2}\text{Tr}(XH) = \tfrac{1}{\sqrt{2}} \tag{36}$$
$$\langle \psi | A_1 \otimes B_1 | \psi \rangle = \tfrac{1}{2}\text{Tr}(X(-XHX)) = -\tfrac{1}{\sqrt{2}}, \tag{37}$$

which implies that the quantum bias is $\beta_q(\text{CHSH}) \geq \frac{1}{\sqrt{2}}$. Translating this into the original formulation of CHSH with output alphabets $\{0,1\}$, we obtain $\omega_q(\text{CHSH}) \geq \left(1 + \frac{1}{\sqrt{2}}\right)/2 = \frac{1}{2} + \frac{\sqrt{2}}{4}$.

At this point, we have an entangled strategy that succeeds with probability $\frac{1}{2} + \frac{\sqrt{2}}{4} = 0.853\ldots$, but we haven't ruled out the possibility of attaining a higher success probability using some other entangled strategy. We do this next, in section 2.2.

2.2 Optimality proof for CHSH (Tsirelson bound)

Theorem 2.1. *The maximum value attained by any entangled strategy for the CHSH game is* $\frac{1}{2} + \frac{\sqrt{2}}{4}$.

Proof. Let the entangled state be ψ, Alice's observables be A_0 and A_1, and Bob's observables be B_0 and B_1. (Recall that these are binary observables.) It suffices to show that

$$\frac{1}{4} \langle \psi | A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1 | \psi \rangle \leq \frac{1}{\sqrt{2}}. \tag{38}$$

Define the vectors $v_0 = A_0 \otimes I|\psi\rangle$, $v_1 = A_1 \otimes I|\psi\rangle$, $w_0 = I \otimes B_0|\psi\rangle$, and $w_1 = I \otimes B_1|\psi\rangle$. Note that v_0, v_1, w_0, w_1 are unit vectors and, for each $j, k \in \{0,1\}$, $\langle\psi|A_j \otimes B_k|\psi\rangle = v_j \cdot w_k$ (where "$\cdot$" denotes the inner product).

Therefore, the LHS of Eq. (38) is

$$\frac{1}{4}\big(v_0 \cdot w_0 + v_0 \cdot w_1 + v_1 \cdot w_0 - v_1 \cdot w_1\big) = \frac{1}{4}\big(v_0 \cdot (w_0 + w_1) + v_1 \cdot (w_0 - w_1)\big). \tag{39}$$

Now, using the fact that v_0, v_1, w_0, w_1 are unit vectors,

$$\begin{aligned} v_0 \cdot (w_0 + w_1) + v_1 \cdot (w_0 - w_1) &\leq |v_0 \cdot (w_0 + w_1)| + |v_1 \cdot (w_0 - w_1)| && (40)\\ &\leq \|w_0 + w_1\| + \|w_0 - w_1\| && (41)\\ &= (1,1) \cdot (\|w_0 + w_1\|, \|w_0 - w_1\|) && (42)\\ &\leq \sqrt{2}\sqrt{\|w_0 + w_1\|^2 + \|w_0 - w_1\|^2} && (43) \end{aligned}$$

(where we have used the Cauchy-Schwarz inequality). Moreover,

$$\begin{aligned} \|w_0 + w_1\|^2 + \|w_0 - w_1\|^2 &= (w_0 + w_1) \cdot (w_0 + w_1) + (w_0 - w_1) \cdot (w_0 - w_1) && (44)\\ &= w_0 \cdot w_0 + w_1 \cdot w_1 + w_0 \cdot w_0 + w_1 \cdot w_1 && (45)\\ &= 2\|w_0\|^2 + 2\|w_1\|^2 && (46)\\ &= 4. && (47) \end{aligned}$$

Combining these equations, we get

$$\frac{1}{4}\langle\psi|A_0B_0 + A_0B_1 + A_1B_0 - A_1B_1|\psi\rangle \leq \frac{1}{4}\sqrt{2}\sqrt{4} = \frac{1}{\sqrt{2}}, \tag{48}$$

as required. □

2.3 XOR games

The CHSH game is from a class of games that are, commonly referred to as *XOR games*, which are defined are nonlocal games where: the output alphabets are $\{0,1\}$, and $V(a,b,s,t) = f(a \oplus b, s, t)$ for some function $f : \{0,1\} \times S \times T \to \{0,1\}$. Thus, for any questions, the payoff is the same for the output $(0,0)$ as it is for the output $(1,1)$, and the payoff is the same for the output $(0,1)$ as it is for the output $(1,0)$.

It turns out that the XOR games have nice structural properties and are amenable to an analysis similar to that of CHSH.

These games will be discussed in further detail later on.

3 The Magic Square game

Considering the system of six equations in nine variables $v_1, \ldots, v_9$:

$$v_1 \oplus v_2 \oplus v_3 = 0 \tag{49}$$
$$v_4 \oplus v_5 \oplus v_6 = 0 \tag{50}$$
$$v_7 \oplus v_8 \oplus v_9 = 0 \tag{51}$$
$$v_1 \oplus v_4 \oplus v_7 = 0 \tag{52}$$
$$v_2 \oplus v_5 \oplus v_8 = 0 \tag{53}$$
$$v_3 \oplus v_6 \oplus v_9 = 1. \tag{54}$$

This is most easily conceptualized as the parities of the rows and columns of the following 3×3 table of variables

v_1	v_2	v_3
v_4	v_5	v_6
v_7	v_8	v_9

There is no simultaneous solution to these equations (in other words, there is no way to fill in the entries of the above table with bits so that the parity along all rows and the first two columns is even, whereas the parity along the third column is odd).

The magic square game is one in which: one of the six equations is specified to Alice as input and she is required to give an assignment of bits to the variables in that equation as output (so we can take $S = \{1, 2, 3, 4, 5, 6\}$ and $A = \{0, 1\}^3$); one of the nine variables is specified Bob as input and he is required to give an assignment to that variable that is consistent with Alice's (so we can take $T = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and $B = \{0, 1\}$). There are 18 possible questions (the questions arise from a proper subset of $S \times T$ because, for example, if s specifies the first equation, then $t \in \{1, 2, 3\}$).

There is no classical strategy that attains success probability 1. The best winning probabiiity that can be attained is $\frac{17}{18}$ (because 17 out of the 18 instances of (s, t) can be satisfied).

It is remarkable that there is as an entangled strategy that attains success probability 1. The six equations can be equivalently written in multiplicative form, where the variables take values in $\{+1, -1\}$ and their products are $+1$ for the first eight equations and -1 for the ninth equation.

The entangled strategy is based on an *operator solution* to the equations, such as

$I \otimes Z$	$Z \otimes I$	$Z \otimes Z$
$X \otimes I$	$I \otimes X$	$X \otimes X$
$X \otimes Z$	$Z \otimes X$	$Y \otimes Y$

where the variables are assigned 4×4 matrices, the matrices in each row/column commute, and the products are $+I$ and $-I$.

The entangled state[3] $\psi = (\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle) \otimes (\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle)$ (which is a maximally entangled state).

[3] To clarify who has which qubit, we can write the entangled state as
$\psi = (\frac{1}{\sqrt{2}}|00\rangle_{A_1B_1} + \frac{1}{\sqrt{2}}|11\rangle_{A_1B_1}) \otimes (\frac{1}{\sqrt{2}}|00\rangle_{A_2B_2} + \frac{1}{\sqrt{2}}|11\rangle_{A_2B_2}) = \frac{1}{\sqrt{4}} \sum_{k \in \{00,01,10,11\}} |k\rangle_A \otimes |k\rangle_B$.

Alice's strategy is to measure the observables in the row or column she is assigned and output their value. The fact that the observables in each row/column commute imply that the outcomes of these measurements are well-defined. The products along the rows/columns imply that the parity of the bits Alice outputs are correct (in fact, this would hold for any two-qubit state that Alice is applies the measurements on). Bob measures the transpose of the observable corresponding to the variable he is assigned.

For example, if Alice is asked the first equation ($v_1 \oplus v_2 \oplus v_3 = 0$) and Bob is asked the second variable (v_2) then Alice measures with respect to the observables $I \otimes Z$, $Z \otimes I$, $Z \otimes Z$ and outputs the three bits and Bob measures with respect to the observable $(Z \otimes I)^{\mathrm{T}} = Z \otimes I$ and outputs the bit. The two ± 1-valued bits are consistent because their product is $+1$ with probability

$$\begin{aligned} \langle\psi|(Z \otimes I) \otimes (Z \otimes I)^{\mathrm{T}}|\psi\rangle &= \tfrac{1}{4}\mathrm{Tr}\big((Z \otimes I)(Z \otimes I)\big) && (55)\\ &= \tfrac{1}{4}\mathrm{Tr}(I \otimes I) && (56)\\ &= 1. && (57) \end{aligned}$$

4 Odd Cycle game

Consider a cyclic graph with an odd number of vertices. Note that this graph is not 2-colorable. The Odd Cycle game is related to this. For some odd $n \geq 3$, let $S = T = \mathbb{Z}_n$ and $A = B = \{0, 1\}$. The distribution π on the questions is the uniform distribution on the following subset of $S \times T$:

$$\{(s, t) \in S \times T : t = s \text{ or } t = s + 1 \bmod n\}. \tag{58}$$

The payoff is defined as

$$V(a, b|s, t) = \begin{cases} 1 & \text{if } t = s \text{ and } a \oplus b = 0 \\ 1 & \text{if } t = s + 1 \text{ and } a \oplus b = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{59}$$

Intuitively, Alice and Bob are each asked for colors of vertices of the n-cycle graph. If they are asked the same vertex they must return the same color to win. If they are asked adjacent vertices then they must return different colors. And they are always asked for either the same vertex or adjacent vertices.

This payoff function can be expressed as

$$V(a, b|s, t) = \begin{cases} 1 & \text{if } a \oplus b = f(s, t) \\ 0 & \text{otherwise,} \end{cases} \tag{60}$$

where

$$f(s, t) = \begin{cases} 0 & \text{if } t = s \\ 1 & \text{if } t = s + 1. \end{cases} \tag{61}$$

The classical value of this game is $1 - \frac{1}{2n}$. The strategy that attains this value is to take an almost-2-coloring of the vertices—one that violates all but one edge—and for Alice and Bob to each return bits from that. This strategy wins for all but one of the of the $2n$ possible questions

$(s,t) \in S \times T$. To check the optimality of this strategy, not that the only better strategy would be one that wins on all possible questions. It is straightforward to check that if such a strategy exists then the n-cycle graph is 2-colorable, which is a contradiction.

We next describe a quantum strategy that performs better than the above classical strategy. It is based on the entangled state

$$|\psi\rangle = \tfrac{1}{\sqrt{2}}|00\rangle - \tfrac{1}{\sqrt{2}}|11\rangle. \tag{62}$$

This state has a nice form if local rotations are applied to it. That is, for

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{63}$$

$$\big(R(\theta_A) \otimes R(\theta_B)\big)\Big(\tfrac{1}{\sqrt{2}}|00\rangle - \tfrac{1}{\sqrt{2}}|11\rangle\Big) \tag{64}$$

$$= \cos(\theta_A + \theta_B)\Big(\tfrac{1}{\sqrt{2}}|00\rangle - \tfrac{1}{\sqrt{2}}|11\rangle\Big) + \sin(\theta_A + \theta_B)\Big(\tfrac{1}{\sqrt{2}}|01\rangle + \tfrac{1}{\sqrt{2}}|10\rangle\Big). \tag{65}$$

The protocol for Alice is: on input $s \in \mathbb{Z}_n$, apply the rotation $R\big(s(\frac{\pi}{2} - \frac{\pi}{2n})\big)$, measure in the computational basis and output the resulting bit. The protocol for Bob is: on input $t \in \mathbb{Z}_n$, apply the rotation $R\big(\frac{\pi}{4n} - t(\frac{\pi}{2} - \frac{\pi}{2n})\big)$, measure in the computational basis and output the resulting bit.

If $t = s$ then the resulting state is

$$\cos(\tfrac{\pi}{4n})\Big(\tfrac{1}{\sqrt{2}}|00\rangle - \tfrac{1}{\sqrt{2}}|11\rangle\Big) + \sin(\tfrac{\pi}{4n})\Big(\tfrac{1}{\sqrt{2}}|01\rangle + \tfrac{1}{\sqrt{2}}|10\rangle\Big) \tag{66}$$

and the probability that $a = b$ is $\cos^2(\frac{\pi}{4n})$ If $t = s + 1$ then the resulting state is

$$\cos(\tfrac{\pi}{2} - \tfrac{\pi}{4n})\Big(\tfrac{1}{\sqrt{2}}|00\rangle - \tfrac{1}{\sqrt{2}}|11\rangle\Big) + \sin(\tfrac{\pi}{2} - \tfrac{\pi}{4n})\Big(\tfrac{1}{\sqrt{2}}|01\rangle + \tfrac{1}{\sqrt{2}}|10\rangle\Big) \tag{67}$$

and the probability that $a = b$ is $\cos^2(\frac{\pi}{2} - \frac{\pi}{4n}) = \sin^2(\frac{\pi}{4n}) = 1 - \cos^2(\frac{\pi}{4n})$. Therefore, the success probability of this strategy is $\cos^2(\frac{\pi}{4n}) = 1 - (\frac{\pi}{4})^2 \frac{1}{n^2} + O(\frac{1}{n^4})$. As a specific example, for $n = 5$, we have $\omega_c(G) = 0.9$ and $\omega_q(G) \geq 0.9755\ldots$.

Is the above entangled strategy optimal? We will revisit this question in section 6.2.

5 Correlations for XOR games and Tsirelson's correspondence

Recall that an entangled strategy for an XOR game consists of an entangled state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ and sets of binary observables $\{A_s \in B(\mathcal{H}_A) : s \in S\}$ and $\{B_t \in B(\mathcal{H}_B) : t \in T\}$.

Associated with each strategy for an XOR game, is a $|S| \times |T|$ *correlation matrix* Q defined as

$$Q_{s,t} = \langle\psi|A_s \otimes B_t|\psi\rangle, \tag{68}$$

for each question pair $(s,t) \in S \times T$. $Q_{s,t} \in [-1,+1]$ and represents the bias of $a \oplus b$ towards 0. The correlation matrix describes the behaviour[4] that the strategy attains. For a given nonlocal game, and strategy, the value attained by the strategy is

$$\beta_q(G) = \sum_{(s,t) \in S \times T} \pi(s,t)(-1)^{f(s,t)} Q_{s,t}. \tag{69}$$

[4]A more general notion of correlation matrix can also be defined that is relevant for games beyond XOR games.

For question sets S and T, we can also define a *vector system* as two sets of unit vectors in some $\mathbb{R}^m$

$$\{|v_s\rangle : s \in S\} \quad \text{and} \quad \{|w_s\rangle : t \in T\}. \tag{70}$$

Tsirselson showed that, for any question sets S and T, there is an entangled strategy attaining a correlation matrix Q if and only is there is a vector system such that $Q_{s,t} = \langle v_s|w_t\rangle$ for all $(s,t) \in S \times T$. Therefore, an alternate way of defining the entangled value of an XOR game is as the supremium over all vector systems of

$$\sum_{(s,t)\in S\times T} \pi(s,t)(-1)^{f(s,t)}\langle v_s|w_t\rangle. \tag{71}$$

5.1 Converting from an entangled strategy to a vector system

Note that

$$\langle\psi|A_s \otimes B_t|\psi\rangle = \big(\langle\psi|A_s \otimes I\big)\big(I \otimes B_t|\psi\rangle\big) \tag{72}$$

and

$$\langle\psi|A_s \otimes I = \big(A_s \otimes I|\psi\rangle\big)^*. \tag{73}$$

Therefore, $\langle\psi|A_s \otimes B_t|\psi\rangle$ is the inner product between vectors $A_s \otimes I|\psi\rangle$ and $I \otimes B_t|\psi\rangle$. Thus, it suffices to set

$$|v_s\rangle = A_s \otimes I|\psi\rangle \tag{74}$$
$$|w_t\rangle = I \otimes B_s|\psi\rangle, \tag{75}$$

for all $s \in S$ and $t \in T$.

Note that if the local dimension of the entanglement is d then the vector system is in $\mathbb{C}^{d^2} \equiv \mathbb{R}^{2d^2}$.

5.2 Converting from a vector system to an entangled strategy

This is the more remarkable direction. The construction is based of a sequence of binary observables $U_1, \ldots, U_m$ with the property that, for all $j \neq k$, U_j and U_k anticommute (i.e., $U_jU_k = -U_kU_j$). For $m = 6$, a construction in terms of operators acting on $\mathbb{C}^{2^5}$

$$U_1 = Z \otimes I \otimes I \otimes I \otimes I \tag{76}$$
$$U_2 = X \otimes Z \otimes I \otimes I \otimes I \tag{77}$$
$$U_3 = X \otimes X \otimes Z \otimes I \otimes I \tag{78}$$
$$U_4 = X \otimes X \otimes X \otimes Z \otimes I \tag{79}$$
$$U_5 = X \otimes X \otimes X \otimes X \otimes Z \tag{80}$$
$$U_6 = X \otimes X \otimes X \otimes X \otimes X, \tag{81}$$

and this generalizes to any m in terms of operators on $\mathbb{C}^\ell$, where $\ell = 2^{m-1}$.

For any $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbb{R}^m$, with $\|\alpha\| = 1$, define the operator L_α as

$$L_\alpha = \sum_{k=1}^{m} \alpha_k U_k. \tag{82}$$

Define the maximally entangled state $|\psi\rangle$ on $\mathbb{C}^\ell$ as

$$|\psi\rangle = \frac{1}{\sqrt{\ell}} \sum_{k \in \mathbb{Z}_\ell} |k\rangle \otimes |k\rangle. \tag{83}$$

Lemma 5.1. *For each unit vector $\alpha \in \mathbb{R}^m$, L_α is a binary observable, and for each pair of unit vectors $\alpha, \beta \in \mathbb{R}^m$,*

$$\langle\psi| L_\alpha \otimes L_\beta^T |\psi\rangle = \alpha \cdot \beta. \tag{84}$$

Proof. Since each H_j is Hermitian and each α_j is real-valued, L_α is Hermitian. Also,

$$\begin{aligned}
(L_\alpha)^2 &= \left(\sum_{j=1}^{m} \alpha_j M_j \right)^2 && (85)\\
&= \sum_{j=1}^{m} \sum_{k=1}^{m} \alpha_j \alpha_k M_j M_k && (86)\\
&= \sum_{j=1}^{m} \alpha_j^2 L_j^2 + \sum_{j<k} \big(\alpha_j \alpha_k L_j L_k + \alpha_k \alpha_j L_k L_j \big) && (87)\\
&= \sum_{j=1}^{m} \alpha_j^2 I && (88)\\
&= I, && (89)
\end{aligned}$$

where we have used the fact that, for $j \neq k$, $L_j L_k + L_k L_j = 0$. Therefore, L_α is a binary observable.

Next, for $\alpha, \beta \in \mathbb{R}^m$ with $\|\alpha\| = \|\beta\| = 1$,

$$\begin{aligned}
\langle\psi| L_\alpha \otimes L_\beta^T |\psi\rangle &= \frac{1}{\ell} \mathrm{Tr}(L_\alpha L_\beta) && (90)\\
&= \frac{1}{\ell} \sum_{j=1}^{m} \sum_{k=1}^{m} \mathrm{Tr}(\alpha_j \beta_k M_j M_k) && (91)\\
&= \frac{1}{\ell} \sum_{j=1}^{m} \mathrm{Tr}(\alpha_j \beta_k I) + \frac{1}{\ell} \sum_{j<k}^{m} \mathrm{Tr}(\alpha_j \beta_k L_j L_k + \alpha_k \beta_j L_k L_j) && (92)\\
&= \frac{1}{\ell} \alpha_j \beta_j \ell + \frac{1}{\ell} \sum_{j<k}^{m} \alpha_j \beta_k \mathrm{Tr}(L_j L_k - L_j L_k) && (93)\\
&= \sum_{j=1}^{m} \alpha_j \beta_j && (94)\\
&= \alpha \cdot \beta. && (95)
\end{aligned}$$

□

From the above lemma, it is clear how to convert a vector system in $\mathbb{R}^m$ into an entangled strategy. The entangled state is $|\psi\rangle$, Alice's observables are

$$\{L_{|v_s\rangle} : s \in S\} \tag{96}$$

and Bob's observables are

$$\{L^T_{|w_t\rangle} : t \in T\}. \tag{97}$$

Note that the local dimension of $|\psi\rangle$ is $\ell = 2^{m-1}$, exponentially larger than m. If fact, this is not the optimal construction: there is a construction where this dimension is $\ell = 2^{\lceil (m-1)/2 \rceil}$ and it can be proven that this is *optimal* (in the sense that the dimension cannot be reduced any further).

6 Disentangling strategies by "rounding"

Rounding is a method that converts an entangled XOR-strategy with success probability $1 - \delta$ to a classical strategy with success probability at least $1 - \sqrt{\delta}$. Note that the classical strategy will generally have lower success probability than the entangled success probability, but it will nevertheless be large if δ is small.

We frequently view the consequences of rounding in the contrapositive form, where it can be used to prove interesting upper bounds on the best possible entangled strategy. If an upper bound on the best classical strategy is already known to be $1 - \epsilon$ then it can be deduced that the best quantum strategy cannot exceed $1 - \epsilon^2$; otherwise, rounding the quantum would lead to a classical strategy that contradicts the known upper bound. We consider such upper bounds in section 6.2.

6.1 Rounding procedure

We begin by explaining how the rounding procedure works for a nondegenerate XOR game $G = (S, T, \pi, f)$. Suppose we have a vector system in $\mathbb{R}^m$ for $S \otimes T$ achieving $\langle v_s | w_t \rangle = Q_{s,t}$ for all $(s, t) \in S \times T$. Recall that the vector system implies that there is a quantum strategy attaining success probability

$$\omega_q(G) = \sum_{(s,t) \in S \times T} \pi(s,t) \frac{1 + (-1)^{f(s,t)} \langle u_s | v_t \rangle}{2}. \tag{98}$$

The rounding procedure that we now describe next leads to a *probabilistic* classical strategy (where Alice and Bob use shared randomness). However, by convexity, this probabilistic strategy can be converted to a deterministic strategy without reducing its success probability.

The probabilistic classical strategy is based on a uniformly random unit vector $|r\rangle \in \mathbb{R}^m$. (Since the surface of the unit ball in $\mathbb{R}^m$ is compact, it has a uniform measure.) We assume that both Alice and Bob have a copy of $|r\rangle = (r_1, r_2, \ldots, r_m)$. The procedure for Alice is: on input $s \in S$, output

$$\begin{cases} 0 & \text{if } \langle r | v_s \rangle \geq 0 \\ 1 & \text{if } \langle r | v_s \rangle < 0. \end{cases} \tag{99}$$

Similarly, the procedure for Bob is: on input $t \in T$, output

$$\begin{cases} 0 & \text{if } \langle r|w_t\rangle \geq 0 \\ 1 & \text{if } \langle r|w_t\rangle < 0. \end{cases} \tag{100}$$

Fig. 1 illustrates the different regions in $\mathbb{R}^m$ that arise for a given $|v_s\rangle$ and $|w_t\rangle$, where the angle between $|v_s\rangle$ and $|w_t\rangle$ is $\theta \in [0, \pi]$ (thus $\langle v_s|w_t\rangle = \cos\theta$). The disk orthogonal to $|v_s\rangle$ is the region where $\langle r|v_s\rangle = 0$; one side of the disk is where $\langle r|v_s\rangle > 0$; the other side is where $\langle r|v_s\rangle < 0$. Similarly for the disc orthogonal to $|w_t\rangle$. The two discs together partition the sphere into the

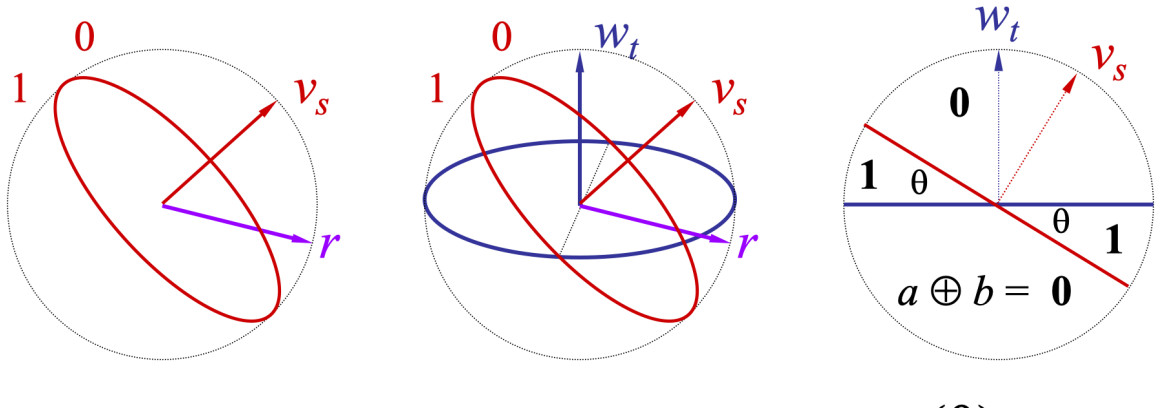

Figure 1: The different regions for $a \oplus b$ depending on where λ is.

regions for $|r\rangle$ that result in $a \oplus b = 0$ and those that result in $a \oplus b = 1$.

The probability that this classical procedure results in $a \oplus b = 0$ for a particular $(s, t) \in S \times T$ is $= 1 - \frac{\theta}{\pi}$. Therefore, the success probability for question (s, t) is

$$p_c(s,t) := \Pr[a \oplus b = f(s,t)] = \begin{cases} 1 - \frac{\theta}{\pi} & \text{if } f(s,t) = 0 \\ \frac{\theta}{\pi} & \text{if } f(s,t) = 1. \end{cases} \tag{101}$$

We analyze the success probability of the entangled strategy on question (s, t) in two cases. In the case where $f(s, t) = 0$,

$$p_q(s,t) = \frac{1+\cos\theta}{2} = \frac{1+\cos\big(\pi(1-p_c(s,t))\big)}{2} = \frac{1-\cos\big(\pi p_c(s,t)\big)}{2} = \sin^2\big(\tfrac{\pi}{2} p_s(s,t)\big). \tag{102}$$

In the case where $f(s, t) = 1$,

$$p_q(s,t) = \frac{1-\cos\theta}{2} = \frac{1-\cos\big(\pi p_c(s,t)\big)}{2} = \sin^2\big(\tfrac{\pi}{2} p_s(s,t)\big). \tag{103}$$

The left side of Fig. 2 plots the relationship between $p_q(s, t)$ (vertical axis) and $p_c(s, t)$ (horizontal axis).

For the entangled strategy and its rounded classical strategy, the success probabilities are

$$p_q = \sum_{(s,t)\in S\times T} \pi(s,t) p_q(s,t) \tag{104}$$

$$p_c = \sum_{(s,t)\in S\times T} \pi(s,t) p_c(s,t), \tag{105}$$

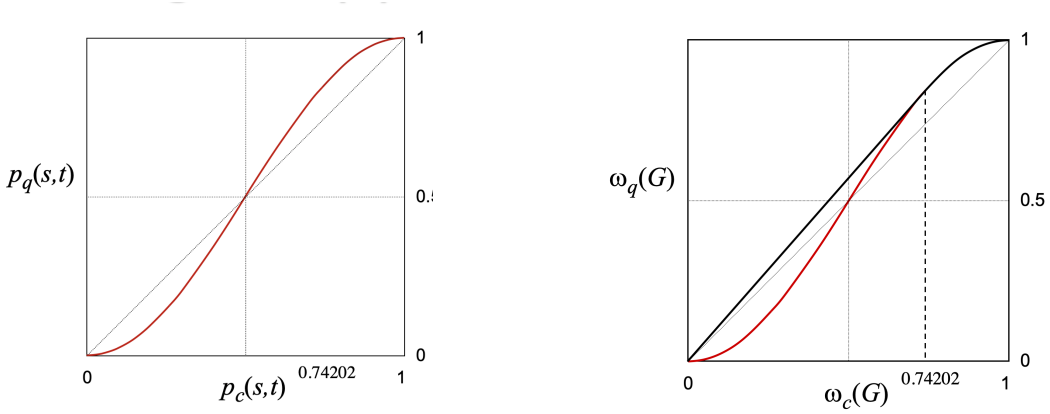

Figure 2: Relationship between $p_q(s,t)$ and $p_c(s,t)$ and relationship between $\omega_q(G)$ and $\omega_c(G)$.

respectively. We cannot apply Eq. (103) directly here, to relate p_q to p_c because the function $p \mapsto \sin^2(\frac{\pi}{2}p)$ is not concave and therefore need not apply to convex combinations. Instead, we can take the minimal concave function that upper bounds the curve, $g : [0,1] \to [0,1]$ as depicted in the right side of Fig. 2, defined as

$$g(x) = \begin{cases} \gamma_1 x & \text{if } 0 \leq x \leq \gamma_2 \\ \sin^2(\frac{\pi}{2}x) & \text{if } \gamma_2 < x \leq 1, \end{cases} \tag{106}$$

where $\gamma_1 \approx 1.1382$ and $\gamma_2 \approx 0.74202$, and relate $\omega_q(G)$ to $\omega_c(G)$ as

$$\omega_q(G) = \sum_{(s,t)\in S\times T} \pi(s,t) p_q(s,t) \tag{107}$$

$$\leq \sum_{(s,t)\in S\times T} \pi(s,t) g\big(p_c(s,t)\big) \tag{108}$$

$$\leq g\left(\sum_{(s,t)\in S\times T} \pi(s,t) p_c(s,t) \right) \tag{109}$$

$$\leq g(\omega_c(G)). \tag{110}$$

6.2 Revisiting the upper bounds for the CHSH and Odd Cycle games

Returning to the Odd Cycle game of section 4, since the classical value is $1 - \frac{1}{2n}$, the quantum value is at most $\sin^2(\frac{\pi}{2}(1 - \frac{1}{2n})) = \cos^2(\frac{\pi}{4n})$, which exactly matches the success probability obtained by the entangled strategy in section 4—therefore that is an optimal strategy and $\omega_q(G) = \cos^2(\frac{\pi}{4n})$.

This approach also provides an alternative proof that the strategy for CHSH in section 2.2 is optimal (since $\cos^2(\frac{\pi}{2}\frac{3}{4}) = \big(1 + \frac{1}{\sqrt{2}}\big)/2$).

Finally, it should be noted that rounding only produces upper bounds. For example, for the XOR game where $f(s,t) = (s_1 \wedge t_1) \oplus (s_2 \wedge t_2)$, it is known that $\omega_q(G) = \omega_c(G) = \frac{3}{4}$, so entangled success probability $\cos^2(\frac{\pi}{4n})$ is not attainable.

7 Characterizing perfect strategies for the Magic Square game

Recall from section 3 that any operator solution to the magic square constraints with binary observables acting on $\mathbb{C}^d$ can be converted into an entangled protocol using a maximally entangled state with local dimension d. What we show here is that *any* entangled strategy for the Magic Square game has this form. In fact, this characterization carries over to all *Binary Constraint System (BCS)* games, but we focus on the Magic Square game here for simplicity.

Theorem 7.1. *For any perfect entangled strategy for the Magic Square game using an entangled state in $\mathbb{C}^d \otimes \mathbb{C}^d$, there exist binary observables $A_1, \ldots, A_9 \in B(\mathbb{C}^d)$ such that, for each variable v_j occurring in Alice's constraint, she measures with respect to the corresponding binary observable A_j and, for the variable v_t that Bob is queried, he measures with respect to $(A_t)^T$, where the transpose is with respect to the Schmidt basis of the entangled state.*

Prior to proving Theorem 7.1, we prove a technical lemma about states with full Schmidt rank.

7.1 A basic property of states with full Schmidt rank

Recall that every state $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ has a Schmidt decomposition of the form

$$|\psi\rangle = \sum_{k=1}^{d} \lambda_k |\phi_k\rangle \otimes |\gamma_k\rangle, \tag{111}$$

where $|\phi_1\rangle, \ldots, |\phi_d\rangle$ and $|\gamma_1\rangle, \ldots, |\gamma_d\rangle$ are orthonormal bases for $\mathbb{C}^d$ and (without loss of generality) $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d \geq 0$. The Schmidt coefficients $\lambda_1, \lambda_2, \ldots, \lambda_d$ in the above form are unique and the *Schmidt rank* of $|\psi\rangle$ of such a decomposition is the number of non-zero Schmidt coefficients, counting multiciplicity (i.e., the maximum $k \in \{1, \ldots, d\}$ such that $\lambda_k > 0$). We say that $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ has *full Schmidt rank* if its Schmidt rank is d. The following lemma concerns states with full Schmidt rank.

Lemma 7.2. *If $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ has Schmidt decomposition*

$$|\psi\rangle = \sum_{k=1}^{d} \lambda_k |k\rangle \otimes |k\rangle, \tag{112}$$

such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d > 0$ (i.e., $|\psi\rangle$ has full Schmidt rank) and $A, B \in \mathbb{C}^{d \times d}$ such that $A \otimes I|\psi\rangle = I \otimes B|\psi\rangle$ then $B = A^T$.

Proof. Since $A \otimes I|\psi\rangle = I \otimes B|\psi\rangle$,

$$\sum_{k=1}^{d} \lambda_k (A|k\rangle) \otimes |k\rangle = \sum_{k=1}^{d} \lambda_k |k\rangle \otimes (B|k\rangle). \tag{113}$$

The Schmidt decomposition has the property that the space associated with each Schmidt coefficient is unique. Thus, if the multiplicity of the first Schmidt coefficient is r (i.e., $\lambda_1 = \cdots = \lambda_r > \lambda_{r+1}$) then

$$\operatorname{span}\big(A|1\rangle, \ldots, A|r\rangle\big) = \operatorname{span}\big(|1\rangle, \ldots, |r\rangle\big) = \operatorname{span}\big(B|1\rangle, \ldots, B|r\rangle\big). \tag{114}$$

Therefore, A and B have the block structure

$$A = \begin{bmatrix} A^{(1)} & 0 \\ 0 & A' \end{bmatrix} \qquad B = \begin{bmatrix} B^{(1)} & 0 \\ 0 & B' \end{bmatrix}, \tag{115}$$

where $A^{(1)}$ and $B^{(1)}$ act on $\text{span}\big(|1\rangle, \ldots, |r\rangle\big)$. From Eq. (113), we have

$$\lambda_1 \sum_{k=1}^{r} (A|k\rangle) \otimes |k\rangle = \lambda_1 \sum_{k=1}^{r} |k\rangle \otimes (B|k\rangle), \tag{116}$$

which implies that

$$\big(A^{(1)} \otimes I\big)|\psi^{(1)}\rangle = \big(I \otimes B^{(1)}\big)|\psi^{(1)}\rangle, \tag{117}$$

where

$$|\psi^{(1)}\rangle = \frac{1}{\sqrt{r}} \sum_{k=1}^{r} |k\rangle \otimes |k\rangle. \tag{118}$$

Since $|\psi^{(1)}\rangle$ is the maximally entangled state (in $\mathbb{C}^r \otimes \mathbb{C}^r$), applying Lemma 1.1 from section 1.6, we obtain

$$\big(I \otimes \big(A^{(1)}\big)^T\big)|\psi^{(1)}\rangle = \big(I \otimes B^{(1)}\big)|\psi^{(1)}\rangle, \tag{119}$$

which implies $B^{(1)} = \big(A^{(1)}\big)^T$.

By continuing this process for the other Schmidt coefficients (more formally, this would be by induction), we deduce that A and B have block decompositions (in the computational basis) of the form

$$A = \begin{bmatrix} A^{(1)} & 0 & 0 & \cdots \\ 0 & A^{(2)} & 0 & \cdots \\ 0 & 0 & A^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \qquad B = \begin{bmatrix} B^{(1)} & 0 & 0 & \cdots \\ 0 & B^{(2)} & 0 & \cdots \\ 0 & 0 & B^{(3)} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{120}$$

(where $A^{(j)}$ and $B^{(j)}$ act on the space associated with the j^{th} distinct Schmidt coefficient) and $B^{(j)} = (A^{(j)})^T$. Therefore $B = A^T$. □

Corollary 7.3. *Let $|\psi\rangle = \sum_{k=1}^{d} \lambda_k |k\rangle \otimes |k\rangle$ be a state with full Schmidt rank, $A, B \in \mathbb{C}^{d\times d}$ be unitary, and $\langle\psi|A \otimes B|\psi\rangle = 1$. Then $B = A^T$.*

Proof. Since $(A \otimes I)|\psi\rangle$ and $(I \otimes B)|\psi\rangle$ are two unit vectors with inner product 1, it follows that $(A \otimes I)|\psi\rangle = (I \otimes B)|\psi\rangle$ and we can apply Lemma 7.2. □

7.2 Proof of Theorem 7.1

For each $t \in \{1, 2, \ldots, 9\}$ Bob is returning a bit, so there is a binary observable associated with t. Therefore, we can picture Bob's strategy as a matrix of binary observables

$$\begin{array}{ccc} B_1 & B_2 & B_3 \\ B_4 & B_5 & B_6 \\ B_7 & B_8 & B_9 \end{array} \tag{121}$$

On Alice's side, it's more complicated. For each constraint $s \in \{1, 2, \ldots, 6\}$, Alice returns three bits, and therefore performs an eight-outcome measurement. In general, a POVM measurement can be performed; however, we assume[5] that this is a projective measurement, with eight complete orthogonal projectors: Π_{000}, Π_{001}, Π_{010}, Π_{011}, Π_{100}, Π_{101}, Π_{110}, Π_{111}. From these projectors, we can define these binary observables

$$A_1^{(s)} = \Pi_{000} + \Pi_{001} + \Pi_{010} + \Pi_{011} - \Pi_{100} - \Pi_{101} - \Pi_{110} - \Pi_{111} \tag{122}$$

$$A_2^{(s)} = \Pi_{000} + \Pi_{001} - \Pi_{010} - \Pi_{011} + \Pi_{100} + \Pi_{101} - \Pi_{110} - \Pi_{111} \tag{123}$$

$$A_3^{(s)} = \Pi_{000} - \Pi_{001} + \Pi_{010} - \Pi_{011} + \Pi_{100} - \Pi_{101} + \Pi_{110} - \Pi_{111}. \tag{124}$$

These are three commuting binary observables, and $A_1^{(s)}$ can be interpreted as the first bit that Alice outputs, $A_2^{(s)}$ the second bit and $A_3^{(s)}$ the third bit. With some relabeling of the subscripts, the measurements of Alice's side are of the form

$$\begin{array}{ccccccc} A_1^{(1)} & & & A_2^{(1)} & & & A_3^{(1)} \\ & A_1^{(4)} & & & A_2^{(5)} & & & A_3^{(6)} \\ A_4^{(2)} & & & A_5^{(2)} & & & A_6^{(2)} \\ & A_4^{(4)} & & & A_5^{(5)} & & & A_6^{(6)} \\ A_7^{(3)} & & & A_8^{(3)} & & & A_9^{(3)} \\ & A_7^{(4)} & & & A_8^{(5)} & & & A_9^{(6)} \end{array} \tag{125}$$

To interpret this, $A^{(1)}$ is the observable that Alice associates with v_1 if it arises in the context of constraint 1 ($v_1 \oplus v_2 \oplus v_3 = 0$); whereas $A^{(4)}$ is the observable that Alice associates with v_1 if it arises in the context of constraint 4 ($v_1 \oplus v_4 \oplus v_7 = 0$). In general, it is conceivable that Alice and uses different observables in the two cases. We will next prove that this cannot occur for a perfect strategy.

We have that, for all $j \in \{1, 2, \ldots, 9\}$ and Alice and Bob always return consistent answers for v_j, which implies (for the constraints s and s' that contain variable v_j)

$$\langle\psi|A_j^{(s)} \otimes B_j|\psi\rangle = 1 = \langle\psi|A_j^{(s')} \otimes B_j|\psi\rangle, \tag{126}$$

which, by Corollary 7.3, implies $A_j^{(s)} = (B_j)^T = A_j^{(s')}$. This permits us to denote the observable simply as A_j.

[5]The general case in handled in Ref. [7].

Therefore, Alice and Bob's strategies are, respectively, based on binary observables

$$\begin{array}{ccc} A_1 & A_2 & A_3 \\ A_4 & A_5 & A_6 \\ A_7 & A_8 & A_9 \end{array} \qquad\qquad \begin{array}{ccc} (A_1)^T & (A_2)^T & (A_3)^T \\ (A_4)^T & (A_5)^T & (A_6)^T \\ (A_7)^T & (A_8)^T & (A_9)^T, \end{array} \tag{127}$$

and we have that they are commuting within each row and column.

Must these observables satisfy the constraints? Consider the first constraint, that $A_1A_2A_3 = I$. Note that the probability that, for this constraint, the product of Alice's three output bits is $+1$ is given by

$$\frac{1 + \langle\psi|(A_1A_2A_3) \otimes I|\psi\rangle}{2}. \tag{128}$$

This is 1 if and only if $\langle\psi|(A_1A_2A_3) \otimes I|\psi\rangle = 1$, which, by Corollary 7.3, implies $A_1A_2A_3 = I^T = I$. The other constraints are handled similarly. This completes the proof of Theorem 7.1.

7.3 Generalization to arbitrary binary constraint systems

It is straightforward to extend the proof techniques to arbitrary binary constraint system games. The details can be found in [7].

Theorem 7.4. *For any binary constraint system (BCS) with n binary variables and m constraints, there exists a perfect entangled strategy using an entangled state in $\mathbb{C}^d \otimes \mathbb{C}^d$ if and only there exist binary observables $A_1, \ldots, A_n \in B(\mathbb{C}^d)$ (corresponding to the n variables) such that, for each constraint, the observables corresponding to its variables commute and their product is $+I$ or $-I$ (appropriately, depending on the value of the constraint).*

8 Rigidity of Magic Square game (exact case)

It can be shown for various games that, in a certain sense, all entangled strategies that achieve maximum success probability for them have a specific form. For example, this can be shown for the CHSH game and also for the Magic Square game. Here we show this for the Magic Square game.

Theorem 8.1. *For any perfect strategy for the Magic Square game, there is an orthonormal basis with respect to which the operators are of the form*

$$\begin{array}{ccc} (I \otimes Z) \otimes I_m & (Z \otimes I) \otimes I_m & (Z \otimes Z) \otimes I_m \\ (X \otimes I) \otimes I_m & (I \otimes X) \otimes I_m & (X \otimes X) \otimes I_m \\ (X \otimes Z) \otimes I_m & (Z \otimes X) \otimes I_m & (Y \otimes Y) \otimes I_m \end{array} \tag{129}$$

and the entangled state is of the form

$$|\psi\rangle = \left(\tfrac{1}{\sqrt{2}}|00\rangle_{A_1B_1} + \tfrac{1}{\sqrt{2}}|11\rangle_{A_1B_1}\right) \otimes \left(\tfrac{1}{\sqrt{2}}|00\rangle_{A_2B_2} + \tfrac{1}{\sqrt{2}}|11\rangle_{A_2B_2}\right) \otimes |\phi\rangle_{A_3B_3} \tag{130}$$

for some arbitrary state $|\phi\rangle_{A_3B_3} \in \mathbb{C}^m \otimes \mathbb{C}^m$.

Therefore, the strategy for the Magic Square game in section 3 is unique (up to a basis change and an additional register that is ignored by the measurements)!

Proof. Recall that the commutativity constraints of the magic square game imply that any two operators in the same row or column must commute. What about two observables that are not in the same row or columns, such as A_2 and A_4? It can be shown that they must *anticommute* (i.e., $A_2A_4 = -A_4A_2$). To see why this is so, note that

$$A_3 = A_1A_2 \tag{131}$$
$$A_6 = A_4A_5 \tag{132}$$

so the last column implies $-I = A_3A_6A_9 = (A_1A_2)(A_4A_5)A_9$. Similarly, we have

$$A_7 = A_1A_4 \tag{133}$$
$$A_8 = A_2A_5 \tag{134}$$

so the last row implies $-I = A_7A_8A_9 = (A_1A_4)(A_2A_5)A_9$. Combining the above, we have $A_1A_2A_4A_5A_9 = -A_1A_4A_2A_5A_9$, which implies $A_2A_4 = -A_4A_2$.

By symmetry, *any* two observables that are not in the same row or column must anticommute.

Whenever two binary observables anticommute they have a special form, given by the following lemma.

Lemma 8.2. *If A and B are binary observables acting on $\mathbb{C}^d$ that anticommute (i.e., $AB = -BA$) then d must be even and there exists a unitary U with respect to which*

$$U^*AU = Z \otimes I_m \tag{135}$$
$$U^*BU = X \otimes I_m, \tag{136}$$

where X and Z are the 2×2 Pauli matrices and I_m denotes the $m \times m$ identity operator for $m = \frac{d}{2}$.

Proof. Since A is a binary observable, there exists a coordinate system in which A and B have block structure

$$A = \begin{bmatrix} I_{m_1} & 0 \\ 0 & -I_{m_2} \end{bmatrix} \qquad\qquad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}. \tag{137}$$

The anticommutativity implies that $B_{11} = -B_{11} = 0$ and $B_{22} = -B_{22} = 0$. Since B is Hermitian, $B_{21} = B_{12}^*$. The fact that $B^2 = I$ implies that

$$B_{12}B_{12}^* = I_{m_1} \tag{138}$$
$$B_{12}^*B_{12} = I_{m_2}, \tag{139}$$

which implies that $m_1 = m_2$ and B_{12} is unitary (one way of seeing this is that $m_1 = \mathrm{Tr}(B_{12}B_{12}^*) = \mathrm{Tr}(B_{12}^*B_{12}) = m_2$). Thus (setting $C = B_{12}$) we have

$$B = \begin{bmatrix} 0 & C \\ C^* & 0 \end{bmatrix} \tag{140}$$

for a unitary C. If we set

$$U = \begin{bmatrix} I & 0 \\ 0 & C^* \end{bmatrix} \tag{141}$$

then

$$U^*AU = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} = Z \otimes I_m \tag{142}$$

$$U^*BU = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} = X \otimes I_m. \tag{143}$$

□

Lemma 8.3. *A commutes with $Z \otimes I_m$ and $X \otimes I_m$ if and only if $A = I \otimes B$ for some $m \times m$ matrix B.*

Proof. Commuting with $Z \otimes I$ implies that A has block structure

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \tag{144}$$

and commuting with $X \otimes I$ implies that $A_{11} = A_{22}$. Therefore, setting $B = A_{11}$, we have

$$A = \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix} = I \otimes B. \tag{145}$$

□

We now return to the proof of Theorem 8.1. Since A_2 and A_4 anticommute, they can be expressed as $A_2 = Z \otimes I_n$ and $A_4 = X \otimes I_n$. Since A_1 commutes with A_2 and A_4, we have $A_1 = I \otimes B$. Similarly, we have $A_5 = I \otimes C$.

So far, the structure of the four observables A_1, A_2, A_4, A_5 is the following.

$I \otimes B$	$Z \otimes I_n$	
$X \otimes I_n$	$I \otimes C$	

Since A_1 and A_5 anticommute, we must have $B = Z \otimes I_m$ and $C = X \otimes I_m$ ($m = \frac{n}{2}$) in some coordinate system. Therefore the table becomes

$I \otimes (Z \otimes I_m)$	$Z \otimes (I \otimes I_m)$	
$X \otimes (I \otimes I_m)$	$I \otimes (X \otimes I_m)$	

Finally, we can fill in the other entries of the table from the constraints as

$I \otimes Z \otimes I_m$	$Z \otimes I \otimes I_m$	$Z \otimes Z \otimes I_m$
$X \otimes I \otimes I_m$	$I \otimes X \otimes I_m$	$X \otimes X \otimes I_m$
$X \otimes Z \otimes I_m$	$Z \otimes X \otimes I_m$	$Y \otimes Y \otimes I_m$

The next step is to show that the entanglement is of the correct form. Consider what happens if both Alice and Bob both measure with respect to the $Z \otimes I \otimes I_m$ observable. Since their answers must be consistent, it must hold that

$$\langle\psi|(Z \otimes I \otimes I_m) \otimes (Z \otimes I \otimes I_m)|\psi\rangle = 1, \tag{146}$$

so $(Z \otimes I \otimes I_m) \otimes (Z \otimes I \otimes I_m)|\psi\rangle = |\psi\rangle$. Therefore, the first qubit of Alice and the first qubit of Bob are in a state that is invariant under $Z \otimes Z$. Similarly, by considering what happens if Alice and Bob each measure with respect to the $X \otimes I \otimes I_m$ observable, the first qubit of Alice and the first qubit of Bob are in a state that is invariant under $X \otimes X$. The only 2-qubit state that is invariant under both $Z \otimes Z$ and $X \otimes X$ is $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$.

By a similar argument considering the observables $I \otimes Z \otimes I_m$ and $I \otimes X \otimes I_m$, it can be argued that Alice and Bob's respective second qubits are also in state $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$.

This completes the proof of Theorem 8.1. □

9 Binary linear system games

A *binary linear system game (BLS game)* consists of n $\{0,1\}$-valued variables $v_1, \ldots, v_n$ and m constraints, each of which specifying whether a mod-2 sum of a subset of the variables is 0 or 1. In multiplicative form, the variables are $A_1, \ldots, A_n$ and each constraint specifies whether or not a product of a subset of the variables is I or $-I$.

The Magic Square is an example of such a game. Illustrated in Figure 3 is the Magic Square, Magic Pentagram, and Four Lines game. From Theorem 7.4, there is a perfect strategy for such a

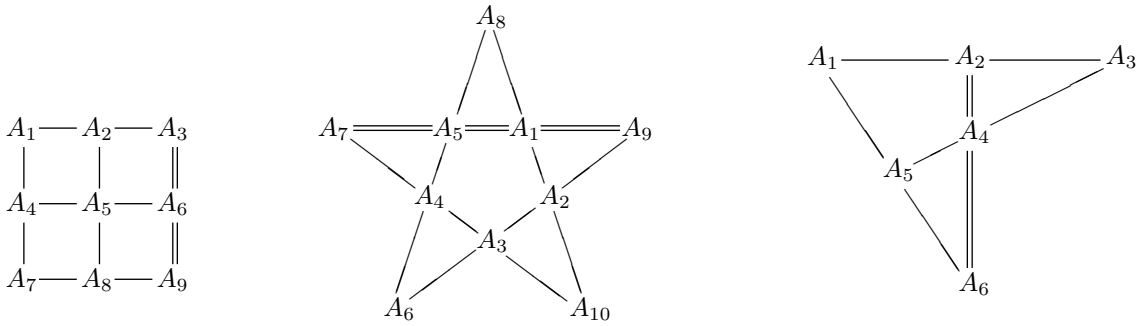

Figure 3: Structure of three BLS games: (a) magic square (left), (b) magic pentagram (middle), and (c) four-lines (right). Each straight line indicates a product constraint on its variables of $+I$ for single lines, and $-I$ for double lines.

game if and only if the multiplicative form of the game has an *operator solution*, where the variables are assigned to matrices such that, for each equation, the operators associated with its variables:

- commute; and
- have product equal to $+I$ or $-I$ (as per the equation).

9.1 Analysis of BLS games with mutiplicity 2

A BLS game has *multiplicity 2* if every variable appears in exactly two constraints. The Magic Square, Magic Pentagram, and Four Lines game (figure 3) all have multiplicity 2.

Ref. [1] shows an elegant way of analyzing any BLS game of multiplicity 2, to either find a perfect strategy or show that none exists. We refer to [1] for the details, and provide only a brief sketch here.

The idea is to first construct a *constraint graph* associated with a BLS game, defined as follows. Each constraint in the BLS corresponds to a vertex and each variable corresponds to an edge

connecting the two vertices corresponding to the constraints where it appears. For example: for the Magic Square, the graph is $K_{3,3}$ (the complete bipartite graph on two sets of size 3); for Magic Pentagram, the graph is K_5 (the complete graph on 5 vertices); and for the Four Lines, the graph is K_4. Each node of the constraint graph is labelled $+1$ or -1, depending if the associated product is $+I$ or $-I$.

If the graph is not connected then each connected component corresponds to a separate BLS game, and the original game has a perfect strategy if and only if each of it connected components has a perfect strategy. Henceforth, we only consider BLS games where the constraint graph is connected.

Given any operator solution to a BLS, if the value of A_j is multiplied by -1 then it is a solution to a new BLS where the two constraints containing A_j are sign flipped (between $+I$ and $-I$). By a series of such moves, any BLS can be converted to one in which there is either a single node with -1 label or no nodes with -1 labels. In the first case there is a trivial classical solution (set each variable to $+1$). The second case is the nontrivial case (where there is no classical solution).

What [1] shows is that (in the nontrivial case):

- If the constraint graph is planar then there is no operator solution. This is shown by proving that there is a systematic way of reducing the equations $I = -I$, a contradiction. The details of this are in section III.A of [1].

- If the constraint graph is not planar then there is an operator solution. This is shown using the fact that every nonplanar graph contains $K_{3,3}$ or K_5 as a *topological minor*[6] and that there are operator solutions for those two constraint graphs. The details of this are in section III.B of [1].

9.2 Analysis of BLS games with mutiplicity ≤ 2

Suppose that each variable appears in *at most* two constraints, but at least one variable appears in only a single constraint. We can reduce this to a BLS where the multiciplity of each variable is exactly 2 as follows. If the original system has variables $A_1, \ldots, A_n$, let the new system have variables $A_1, \ldots, A_n, B_1, \ldots, B_n$. For each constraint, say, $A_i A_j A_k = (-1)^b I$ in the original BLS, let the new BLS have that constraint plus its "twin" $B_i B_j B_k = (-1)^b I$. So far, we have essentially two copies of the original BLS. Also, for each variable A_i that has multiplicity 1 in the original BLS, let the new BLS also include the constraint $A_i B_i = I$ (which increases the multiplicity of both A_i and B_i from 1 to 2).

The new system is a BLS with multiplicity exactly 2 for each variable. Also, there is a solution to the original BLS if and only if there is a solution to the new BLS. To see why this is so, first suppose that the original BLS has an operator solution. Then, setting each $B_i = A_i$, results in a solution to the new BLS. For the other direction, if there is a solution to the new system then, by simply discarding $B_1, \ldots, B_n$, we obtain a solution to the original system.

Therefore, each such BLS can be analyzed by first converting it to one with multiplicity exactly 2 and then applying the techniques in section 9.1.

[6] Defined in Def. 25 of [1].

10 Rigidity of CHSH (extremal case)

In this section, we prove the following.

Theorem 10.1. *For any entangled strategy for the CHSH game that attains the maximum possible success probability of $\left(1+\frac{1}{\sqrt{2}}\right)/2$, there is an orthonormal basis with respect to which Alice and Bob's observables are of the form*

$$A_0 = Z \otimes I_m \tag{147}$$
$$A_1 = X \otimes I_m \tag{148}$$
$$B_0 = H \otimes I_m \tag{149}$$
$$B_1 = (ZHZ) \otimes I_m \tag{150}$$

and the entangled state is of the form

$$|\psi\rangle = \left(\tfrac{1}{\sqrt{2}}|00\rangle + \tfrac{1}{\sqrt{2}}|11\rangle\right) \otimes |\Phi_{\text{junk}}\rangle \tag{151}$$

for some arbitrary state $|\Phi_{\text{junk}}\rangle \in \mathbb{C}^m \otimes \mathbb{C}^m$.

10.1 Preliminary: a special property of two-outcome POVM measurements

In general, a POVM measurement can be simulated by a projective measurement in la larger Hilbert space (the so-called Stinespring dilation). For example, these three operators on $\mathbb{C}^2$

$$E_1 = \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & 0 \end{pmatrix}, \qquad E_2 = \begin{pmatrix} \frac{1}{6} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{2} \end{pmatrix}, \qquad E_3 = \begin{pmatrix} \frac{1}{6} & \frac{-1}{\sqrt{3}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{2} \end{pmatrix} \tag{152}$$

are the elements of a POVM measurement, and this measurement can be simulated by these projective measurements in $\mathbb{C}^3$

$$\Pi_1 = \begin{pmatrix} \frac{2}{3} & 0 & \frac{\sqrt{2}}{3} \\ 0 & 0 & 0 \\ \frac{\sqrt{2}}{3} & 0 & \frac{1}{3} \end{pmatrix}, \quad \Pi_2 = \begin{pmatrix} \frac{1}{6} & \frac{-1}{\sqrt{12}} & \frac{-1}{\sqrt{18}} \\ \frac{-1}{\sqrt{12}} & \frac{1}{2} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{18}} & \frac{1}{\sqrt{6}} & \frac{1}{3} \end{pmatrix}, \quad \Pi_3 = \begin{pmatrix} \frac{1}{6} & \frac{1}{\sqrt{12}} & \frac{-1}{\sqrt{18}} \\ \frac{1}{\sqrt{12}} & \frac{1}{2} & \frac{-1}{\sqrt{6}} \\ \frac{-1}{\sqrt{18}} & \frac{-1}{\sqrt{6}} & \frac{1}{3} \end{pmatrix}, \tag{153}$$

where the qubit to be measured is in $\text{span}(|0\rangle, |1\rangle)$ and the projectors act on the larger space $\text{span}(|0\rangle, |1\rangle, |2\rangle)$. It turns out that, for this example, it is impossible to construct the projectors in the original space $\text{span}(|0\rangle, |1\rangle)$ (i.e., the dimension increase in necessary). In fact, it is not even possible to construct a probabilistic mixture of projectors on $\text{span}(|0\rangle, |1\rangle)$ that simulate the above POVM measurement.

Why do we care about this? In section 7.1, we have some results that apply for projective measurements on entangled states with full Schmidt rank. For any entangled strategy, we can, without loss of generality, assume that the entanglement used has full Schmidt rank. However, the local measurements can, in general, be POVM measurements. If we dilate the measurements so as to be projective then we introduce additional dimensions in the local spaces and the entanglement will not have full Schmidt rank for these larger local spaces. So it seems that we can *either* consider

the entanglement to be of full Schmidt rank and the measurements to be POVMs *or* drop the full Schmidt rank property and assume projective measurements.

For the special case of *two-outcome measurements*, we can have the best of both worlds. This is because any two-outcome POVM measurement can be simulated as a probabilistic mixture of two-outcome projective measurements.

To see why this is so, let E_0, E_1 be the elements of an arbitrary two-outcome POVM measurement. Since $E_0 \geq 0$, we can write E_0 as a diagonal matrix (in some coordinate system) so that

$$E_0 = \begin{pmatrix} \mu_1 & & & \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_d \end{pmatrix} \quad \text{and} \quad E_1 = I - E_0 = \begin{pmatrix} 1-\mu_1 & & & \\ & 1-\mu_2 & & \\ & & \ddots & \\ & & & 1-\mu_d \end{pmatrix}, \tag{154}$$

where $\mu_1, \mu_2, \ldots, \mu_d \in [0,1]$. Intuitively, performing this measurement is equivalent to first generating Π_0 as a diagonal matrix where entry (k,k) is set to 1 with probability μ_k, and to 0 with probability $1-\mu_k$ (independently for each k). And then $\Pi_1 = I - \Pi_0$. Then the projective measurement (Π_0, Π_1) is applied.

The result is what we can nickname the "Two-outcome POVMs are boring Lemma":

Lemma 10.2. *For any two-outcome POVM measurement on space $\mathbb{C}^d$, there exists an ensemble of binary observables $\{A^{(r)} : r \subseteq \{1,2,\ldots,d\}\}$ and a probability measure on the set of all subsets of $\{1,2,\ldots,d\}$, such that sampling $r \subseteq \{1,2,\ldots,d\}$ and then measuring according to the observable $A^{(r)}$ is equivalent to performing the original POVM measurement.*

This will be very useful for the CHSH game where Alice and Bob's perform two-outcome measurements. However, it should be remembered that, for general nonlocal games, there is no such lemma when the number of outcomes is more than two.

10.2 Proof of Theorem 10.1

Lemma 10.2 enables us to prove the rigidity of CHSH in the extremal case quite easily.

Recalling Eq. (38), if $A_0^{(r)}$, $A_1^{(r)}$, $B_0^{(r)}$, $B_1^{(r)}$ are the observables as a function of the randomly generated $r \in \Omega$ then, for each $r \in \Omega$ that arises with non-zero probability, it must hold that

$$\frac{1}{4}\langle\psi| A_0^{(r)} \otimes B_0^{(r)} + A_0^{(r)} \otimes B_1^{(r)} + A_1^{(r)} \otimes B_0^{(r)} - A_1^{(r)} \otimes B_1^{(r)} |\psi\rangle = \frac{1}{\sqrt{2}}. \tag{155}$$

(This is by an averaging argument: since the right side is always at most $1/\sqrt{2}$, if, for some r the value was less than $1/\sqrt{2}$ then the weighted average over all r would be less than $1/\sqrt{2}$.) Henceforth, to reduce clutter, we omit the (r) superscripts. This is equivalent to the following expression

$$\langle\psi| \begin{bmatrix} \frac{1}{\sqrt{2}} A_0 \otimes I & \frac{1}{\sqrt{2}} A_1 \otimes I \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} I \otimes \frac{B_0 + B_1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} I \otimes \frac{B_0 - B_1}{\sqrt{2}} \end{bmatrix} |\psi\rangle = 1, \tag{156}$$

where the bracketed items are block matrices consisting of two square blocks, stacked horizontally or vertically. It is straightforward to verify that the block matrices

$$M_A = \begin{bmatrix} \frac{1}{\sqrt{2}} A_0 \otimes I \\ \\ \frac{1}{\sqrt{2}} A_1 \otimes I \end{bmatrix} \qquad \text{and} \qquad M_B = \begin{bmatrix} \frac{1}{\sqrt{2}} I \otimes \frac{B_0 + B_1}{\sqrt{2}} \\ \\ \frac{1}{\sqrt{2}} I \otimes \frac{B_0 - B_1}{\sqrt{2}} \end{bmatrix} \tag{157}$$

satisfy $M_A^* M_A = I = M_B^* M_B$ and are therefore isometries from $\mathbb{C}^d$ to $\mathbb{C}^{2d}$. If we define vectors $|v\rangle = M_A|\psi\rangle$ and $|w\rangle = M_B|\psi\rangle$ then we have $\||v\rangle\| = \||w\rangle\| = 1$ and $\langle v|w\rangle = 1$, from which it follows that $|v\rangle = |w\rangle$. Therefore,

$$A_0 \otimes I|\psi\rangle = I \otimes \frac{B_0 + B_1}{\sqrt{2}}|\psi\rangle \tag{158}$$

$$A_1 \otimes I|\psi\rangle = I \otimes \frac{B_0 - B_1}{\sqrt{2}}|\psi\rangle. \tag{159}$$

Since $|\psi\rangle$ has full Schmidt rank, by Lemma 7.2, we have

$$\frac{B_0 + B_1}{\sqrt{2}} = A_0^T \qquad \text{and} \qquad \frac{B_0 - B_1}{\sqrt{2}} = A_1^T, \tag{160}$$

where the transpose is with respect to the Schmidt basis coordinate system. By direct calculation,

$$\left(\frac{B_0 + B_1}{\sqrt{2}}\right)\left(\frac{B_0 - B_1}{\sqrt{2}}\right) = \frac{B_1 B_0 - B_0 B_1}{2} = -\left(\frac{B_0 - B_1}{\sqrt{2}}\right)\left(\frac{B_0 + B_1}{\sqrt{2}}\right), \tag{161}$$

which implies that A_0 and A_1 anticommute. Therefore, by Lemma 8.2, there exists a unitary U such that

$$U^* A_0 U = Z \otimes I_m \tag{162}$$

$$U^* A_1 U = X \otimes I_m. \tag{163}$$

We also have (using the fact that $(U^* A U)^T = U^T A^T (U^T)^*$)

$$(U^T) B_0 (U^T)^* = (U^T) \frac{A_0^T + A_1^T}{\sqrt{2}} (U^T)^* = \frac{Z + X}{\sqrt{2}} \otimes I_m = H \otimes I_m \tag{164}$$

$$(U^T) B_0 (U^T)^* = (U^T) \frac{A_0^T - A_1^T}{\sqrt{2}} (U^T)^* = \frac{Z - X}{\sqrt{2}} \otimes I_m = (ZHZ) \otimes I_m. \tag{165}$$

In summary, there exists a local unitary coordinate system transformation of the form $U \otimes U^T$ so that the local observables are

$$A_0 = Z \otimes I_m \tag{166}$$

$$A_1 = X \otimes I_m \tag{167}$$

$$B_0 = H \otimes I_m \tag{168}$$

$$B_1 = (ZHZ) \otimes I_m. \tag{169}$$

Also, the entanglement used is a state $|\psi\rangle \in \mathbb{C}^d$ such that

$$\langle\psi|\Big(\frac{A_0\otimes B_0 + A_0\otimes B_1 + A_1\otimes B_0 - A_1\otimes B_1}{4}\Big)\otimes I_m|\psi\rangle = \frac{1}{\sqrt{2}}. \tag{170}$$

Since

$$\Big(\frac{A_0\otimes B_0 + A_0\otimes B_1 + A_1\otimes B_0 - A_1\otimes B_1}{4}\Big)\otimes I_m = \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}\otimes I_m \tag{171}$$

has eigenvalues $\frac{1}{\sqrt{2}}$, 0, 0, $-\frac{1}{\sqrt{2}}$, it follows that $|\psi\rangle$ is a $\frac{1}{\sqrt{2}}$-eigenvector, which is

$$\Big(\tfrac{1}{\sqrt{2}}|00\rangle + \tfrac{1}{\sqrt{2}}|11\rangle\Big)\otimes|\Phi_{\text{junk}}\rangle. \tag{172}$$

This completes the proof of Theorem 10.1.

11 Robust rigidity of CHSH (approximately extremal case)

In this section, we prove the following, which shows that every protocol that achieves success probability *close* to optimal, is essentially close to the strategy in section 2.1, in some local coordinate systems. The first proof of this (albeit with bounds that are less qualitatively tight) is in [14].

Theorem 11.1 ([14])**.** *For any entangled strategy $(|\psi\rangle, A_0, A_1, B_0, B_1)$ with local dimension d for the CHSH game that attains success probability of at least $\big(1+\frac{1}{\sqrt{2}}-\epsilon\big)/2$, and local isometries $V_A, V_B : \mathbb{C}^d \to \mathbb{C}^2\otimes\mathbb{C}^d$ $(=\mathbb{C}^{2d})$ such that:*

- $\Big\|(V_A\otimes V_B)|\psi\rangle - \Big(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\Big)\otimes|\Phi_{\text{junk}}\rangle\Big\| \in O(\sqrt{\epsilon})$
- $\Big\|(V_A\otimes V_B)(A_0\otimes I)|\psi\rangle - (Z\otimes I)\Big(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\Big)\otimes|\Phi_{\text{junk}}\rangle\Big\| \in O(\sqrt{\epsilon})$
- $\Big\|(V_A\otimes V_B)(A_1\otimes I)|\psi\rangle - (X\otimes I)\Big(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\Big)\otimes|\Phi_{\text{junk}}\rangle\Big\| \in O(\sqrt{\epsilon})$
- $\Big\|(V_A\otimes V_B)(I\otimes B_0)|\psi\rangle - (I\otimes H)\Big(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\Big)\otimes|\Phi_{\text{junk}}\rangle\Big\| \in O(\sqrt{\epsilon})$
- $\Big\|(V_A\otimes V_B)(I\otimes B_1)|\psi\rangle - (I\otimes (ZHZ))\Big(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\Big)\otimes|\Phi_{\text{junk}}\rangle\Big\| \in O(\sqrt{\epsilon})$

for some arbitrary state $|\Phi_{\text{junk}}\rangle \in \mathbb{C}^d\otimes\mathbb{C}^d$.

To compare with Theorem 10.1, note that $V_A A_0 V_A^*$ and $V_A A_1 V_A^*$ need not be close to $Z\otimes I_m$ and $X\otimes I_m$ as operators (on $\mathbb{C}^{2d}$). One reason for this is that, on the part local spaces corresponding to Schmidt coefficients zero, A_0 and A_1 can be arbitrary, and need not approximately anticommute. In the case of Theorem 10.1, by working with entangled states of full rank, we effectively truncated the local spaces so that there are no zero Schmidt coefficients. We can also do that in our present

context; however, there may be some nonzero Schmidt coefficients that are nevertheless very small, and A_0 and A_1 may be far from anticommuting on the corresponding spaces.

Another issue is that, even if $\|A_0A_1 + A_1A_0\|$ is small in the operator norm, there need not exist anticommuting observables $\tilde{A}_0$ and $\tilde{A}_1$ such that $\|A_0 - \tilde{A}_0\|$ and $\|A_1 - \tilde{A}_1\|$ are small (see [19]).

Our approach to proving Theorem 11.1 uses techniques similar to those in the original proof [14], along with techniques in [18]. We utilize a special norm (actually a semi-norm) that is relevant to our setting, which is explained in the next section.

11.1 Inner products and norms relative to a bipartite quantum state

Here we define measures of distance between two isometries $U, V : \mathbb{C}^d \to \mathbb{C}^{d'}$ for $d \leq d'$ that are relevant when they are applied to one component of a bipartite state $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$. Thus, it is the distance between the states $U \otimes I|\psi\rangle$ and $V \otimes I|\psi\rangle$ that we care about.

The inner product between $U \otimes I|\psi\rangle$ and $V \otimes I|\psi\rangle$ is $\langle\psi|U^*V \otimes I|\psi\rangle$, and this expression simplifies as in the following lemma.

Lemma 11.2. *$\langle\psi|U^*V \otimes I|\psi\rangle = \mathrm{Tr}(U^*V\sigma)$, where $\sigma = \mathrm{Tr}_2\big(|\psi\rangle\langle\psi|\big)$ (the reduced density operator of state $|\psi\rangle$).*

The proof of Lemma 11.2 is left as an exercise for the reader.

With the above in mind, the following definition is natural.

Definition 11.3. *Relative to a density operator σ, define the semi-inner product*

$$\langle U, V\rangle_\sigma = \mathrm{Tr}(U^*V\sigma). \tag{173}$$

(Note the resemblance to the Frobenius inner product defined in section 1.5.) We can also define the corresponding semi-norm[7] as follows.

Definition 11.4. *Relative to a density operator σ, define the semi-norm*

$$\|W\|_\sigma = \sqrt{\langle W, W\rangle_\sigma}. \tag{174}$$

A useful way of thinking about the meaning of $\|U - V\|_\sigma$ is as the Euclidean distance between $U \otimes I|\psi\rangle$ and $V \otimes I|\psi\rangle$.

Exercise 11.5. *Show that $\|U - V\|_\sigma$ is the Euclidean distance between $U \otimes I|\psi\rangle$ and $V \otimes I|\psi\rangle$.*

We can relate the norm to the inner product by the following (where $\Re(\zeta) = \frac{1}{2}(\zeta + \overline{\zeta})$, which is real part of $\zeta \in \mathbb{C}$).

Lemma 11.6.

$$\Re\big(\langle U, V\rangle_\sigma\big) = 1 - \tfrac{1}{2}\|U - V\|_\sigma^2. \tag{175}$$

Proof. The proof is straightforward using the fact that

$$\Re\big(\langle U, V\rangle_\sigma\big) = \tfrac{1}{2}\big(\langle U, V\rangle_\sigma + \langle V, U\rangle_\sigma\big) \tag{176}$$

and $\langle U, U\rangle_\sigma = 1 = \langle V, V\rangle_\sigma$. □

[7] We use the terminology *semi-norm* and *semi-inner product*, because it is possible for a non-zero W to have the property that $\|W\|_\sigma = 0$.

11.2 Approximate anticommuting

Here we show that, if a strategy for CHSH is close to optimal then its observables are close to anticommuting (where *close* is with respect to the norm defined in the previous section). The following is from [18].

Theorem 11.7 ([18])**.** *Let $A_0, A_1, B_0, B_1 \in \mathbb{C}^{d\times d}$ and $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ be a strategy for CHSH that attains bias at least $\frac{1}{\sqrt{2}} - \epsilon$. Then, for $\sigma = \mathrm{Tr}_2|\psi\rangle\langle\psi|$,*

$$\big\|A_1A_0 + A_0A_1\big\|_\sigma^2 \leq 32\sqrt{2}\,\epsilon \tag{177}$$

$$\big\|B_0B_1 + B_1B_0\big\|_\sigma^2 \leq 32\sqrt{2}\,\epsilon. \tag{178}$$

Proof. Our starting condition is

$$\tfrac{1}{4}\langle\psi|A_0\otimes B_0 + A_0\otimes B_1 + A_1\otimes B_0 - A_1\otimes B_1|\psi\rangle \geq \tfrac{1}{\sqrt{2}} - \epsilon. \tag{179}$$

Using the identity

$$\left(\frac{A_0\otimes B_0 + A_0\otimes B_1 + A_1\otimes B_0 - A_1\otimes B_1}{2}\right)^2 \tag{180}$$

$$= I\otimes I + \left(\frac{A_1A_0 - A_0A_1}{2}\right)\otimes\left(\frac{B_0B_1 - B_1B_0}{2}\right), \tag{181}$$

we can deduce

$$\langle\psi|\left(\frac{A_1A_0 - A_0A_1}{2}\right)\otimes\left(\frac{B_0B_1 - B_1B_0}{2}\right)|\psi\rangle \tag{182}$$

$$= \langle\psi|\left(\frac{A_0\otimes B_0 + A_0\otimes B_1 + A_1\otimes B_0 - A_1\otimes B_1}{2}\right)^2|\psi\rangle - 1 \tag{183}$$

$$\geq \big(2(\tfrac{1}{\sqrt{2}} - \epsilon)\big)^2 - 1 \tag{184}$$

$$\geq 1 - 4\sqrt{2}\,\epsilon \tag{185}$$

(where, in Eq. (184), we are using the fact that $\langle\psi|M|\psi\rangle^2 \leq \langle\psi|M^2|\psi\rangle$ for any Hermitian M). Applying the Cauchy-Schwarz inequality and the fact that $\|\frac{A_1A_0-A_0A_1}{2}\| \leq 1$, we have

$$\langle\psi|\left(\frac{A_1A_0 - A_0A_1}{2}\right)^2\otimes I\,|\psi\rangle \geq \big(1 - 4\sqrt{2}\epsilon\big)^2 \geq 1 - 8\sqrt{2}\epsilon. \tag{186}$$

Using the fact that, for any unit vectors $|v\rangle$ and $|w\rangle$, $\big\|\frac{|v\rangle+|w\rangle}{2}\big\|^2 + \big\|\frac{|v\rangle-|w\rangle}{2}\big\|^2 = 1$, it follows that

$$\langle\psi|\left(\frac{A_1A_0 + A_0A_1}{2}\right)^2\otimes I\,|\psi\rangle \leq 8\sqrt{2}\,\epsilon. \tag{187}$$

We can similarly derive

$$\langle\psi|\,I\otimes\left(\frac{B_0B_1 + B_1B_0}{2}\right)^2|\psi\rangle \leq 8\sqrt{2}\,\epsilon. \tag{188}$$

Since

$$\left\|A_1A_0 + A_0A_1\right\|_\sigma^2 = \langle\psi|(A_1A_0 + A_0A_1)^2 \otimes I|\psi\rangle \tag{189}$$

$$\left\|B_0B_1 + B_1B_0\right\|_\sigma = \langle\psi|I \otimes (B_0B_1 + B_1B_0)^2|\psi\rangle, \tag{190}$$

the conditions of the theorem are proven. □

It is useful to also express the distance between A_0A_1 and $-A_1A_0$ and via the inner product $\langle\cdot,\cdot\rangle_\sigma$ in the following corollary.

Corollary 11.8. *Let $A_0, A_1, B_0, B_1 \in \mathbb{C}^{d\times d}$ and $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ be a strategy for CHSH that attains bias at least $\frac{1}{\sqrt{2}} - \epsilon$. Then, for $\sigma = \mathrm{Tr}_2|\psi\rangle\langle\psi|$,*

$$\Re\langle -A_1A_0, A_0A_1\rangle_\sigma \geq 1 - 16\sqrt{2}\,\epsilon \tag{191}$$

$$\Re\langle -B_1B_0, B_0B_1\rangle_\sigma \geq 1 - 16\sqrt{2}\,\epsilon. \tag{192}$$

11.3 Canonical form for approximately anticommuting observables

Here, we show an analogue of Lemma 8.2 for the case of *approximately anticommuting* binary observables: namely, that in some coordinate system A_0 and A_1 are *close to* $Z \otimes I_d$ and $X \otimes I_d$; however, we need to double the dimension of the space on which the operators act (i.e., add one qubit). We first do this by a unitary change of basis, which requires us to extend A_0 and A_1 to observables in the larger space. Then we modify or change of basis to an isometry so as to transform the original A_0 and A_1. (Similar results hold for B_0 and B_1.)

11.3.1 Canonical form via a unitary transformation

Our coordinate system transformation is based on U_A on Alice's side (and a similarly defined U_B on Bob's side), which is easily described by the following circuit, illustrated in figure 4. To get

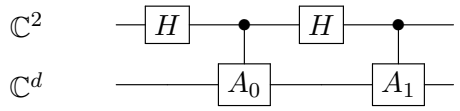

Figure 4: Definition of $U_A : \mathbb{C}^2 \otimes \mathbb{C}^d \to \mathbb{C}^2 \otimes \mathbb{C}^d$.

some intuition about what this does, consider the case where A_0 and A_1 actually anticommute and hence, by Lemma 8.2, are expressible as $Z \otimes I_{d/2}$ and $X \otimes I_{d/2}$ (respectively) in some coordinate system. In that case, it is a simple exercise to show that the circuit simplifies to two controlled-NOT gates and has the property that, when the first qubit is initialized to state $|0\rangle$, it performs a swap between the first qubit and the first qubit of the second register—so the net effect is similar to a swap gate.

Under the coordinate system transform induced by U_A, Lemma 11.9 roughly states that A_0 becomes $Z \otimes I_d$ and Lemma 11.10 states that A_1 becomes an operator close to $X \otimes I_d$. Due to the additional qubit, A_0 and A_1 are extended to operators acting on the larger space. A natural extension is as $I \otimes A_0$ and $I \otimes A_1$. However, Lemma 11.9 is actually based on the extension $Z \otimes A_0$ instead of $I \otimes A_0$; ignore this distinction for the time being (it is resolved in the next section, where our coordinate system transformation is by isometries, so no extension is needed).

Lemma 11.9.

$$U_A^*(Z \otimes I_d)U_A = Z \otimes A_0. \tag{193}$$

Proof. The circuit diagram for $U_A^*(Z \otimes I_d)U_A$ is illustrated in figure 5, and we show that this is equivalent to $Z \otimes A_0$. We begin by noting that the net effect of the three middle gates (the

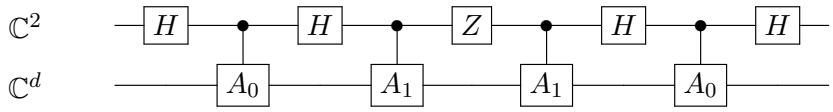

Figure 5: $U_A^*(Z \otimes I_d)U_A$ expressed as a circuit.

controlled-A_1, $Z \otimes I_d$, controlled-A_1 sequence) is to compute the operation specified by the block matrix

$$\begin{bmatrix} I_d & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} I_d & 0 \\ 0 & -I_d \end{bmatrix} \begin{bmatrix} I_d & 0 \\ 0 & A_1 \end{bmatrix} = \begin{bmatrix} I_d & 0 \\ 0 & -I_d \end{bmatrix} = Z \otimes I_d. \tag{194}$$

Next, working outward from these gates to include the two adjacent Hadamard gates, results in a subcircuit that computes $X \otimes I_d$. Next, including the two controlled-A_0 gates, we obtain the block matrix

$$\begin{bmatrix} I_d & 0 \\ 0 & A_0 \end{bmatrix} \begin{bmatrix} 0 & I_d \\ I_d & 0 \end{bmatrix} \begin{bmatrix} I_d & 0 \\ 0 & A_0 \end{bmatrix} = \begin{bmatrix} 0 & A_0 \\ A_0 & 0 \end{bmatrix} = X \otimes A_0 \tag{195}$$

Finally, conjugating by the outer Hadamard gates, we obtain $Z \otimes A_0$. □

Lemma 11.10.

$$\|U_A^*(X \otimes I_d)U_A - I \otimes A_1\|_\rho^2 \leq c\epsilon, \tag{196}$$

where $\rho = |0\rangle\langle 0| \otimes \sigma$ *and* $c = 16\sqrt{2}$.

Proof. We begin by expressing $E = U_A^*(X \otimes I_d)U_A$ in the form of a 2×2 block matrix. The circuit diagram for E is illustrated in figure 6, First, note that the net effect of the three middle gates (the

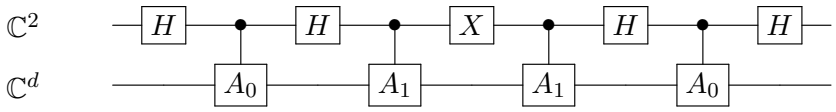

Figure 6: $E = U_A^*(X \otimes I_d)U_A$ expressed as a circuit.

controlled-A_1, $X \otimes I_d$, controlled-A_1 sequence) is to compute

$$\begin{bmatrix} I_d & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 0 & I_d \\ I_d & 0 \end{bmatrix} \begin{bmatrix} I_d & 0 \\ 0 & A_1 \end{bmatrix} = \begin{bmatrix} 0 & A_1 \\ A_1 & 0 \end{bmatrix} = X \otimes A_1. \tag{197}$$

Next, working outward from these gates to include the two adjacent Hadamard gates, results in a subcircuit that computes $Z \otimes A_1$. Next, including the two controlled-A_0 gates, we obtain the block matrix

$$\begin{bmatrix} I_d & 0 \\ 0 & A_0 \end{bmatrix} \begin{bmatrix} A_1 & 0 \\ 0 & -A_1 \end{bmatrix} \begin{bmatrix} I_d & 0 \\ 0 & A_0 \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & -A_0 A_1 A_0 \end{bmatrix}. \tag{198}$$

Finally, conjugating by the outer Hadamard gates, we obtain

$$E = \begin{bmatrix} \frac{1}{\sqrt{2}}I_d & \frac{1}{\sqrt{2}}I_d \\ \frac{1}{\sqrt{2}}I_d & -\frac{1}{\sqrt{2}}I_d \end{bmatrix} \begin{bmatrix} A_1 & 0 \\ 0 & -A_0A_1A_0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}}I_d & \frac{1}{\sqrt{2}}I_d \\ \frac{1}{\sqrt{2}}I_d & -\frac{1}{\sqrt{2}}I_d \end{bmatrix} \tag{199}$$

$$= \frac{1}{2}\begin{bmatrix} A_1 - A_0A_1A_0 & A_1 + A_0A_1A_0 \\ A_1 + A_0A_1A_0 & A_1 - A_0A_1A_0 \end{bmatrix}. \tag{200}$$

We are interested in how well E approximates

$$I \otimes A_1 = \begin{bmatrix} A_1 & 0 \\ 0 & A_1 \end{bmatrix}, \tag{201}$$

which we calculate as

$$\langle E, I \otimes A_1\rangle_\rho = \tfrac{1}{2}\mathrm{Tr}\left(\begin{bmatrix} A_1 - A_0A_1A_0 & A_1 + A_0A_1A_0 \\ A_1 + A_0A_1A_0 & A_1 - A_0A_1A_0 \end{bmatrix} \begin{bmatrix} A_1 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} \sigma & 0 \\ 0 & 0 \end{bmatrix}\right) \tag{202}$$

$$= \tfrac{1}{2}\mathrm{Tr}\big((I - A_0A_1A_0A_1)\sigma\big) \tag{203}$$

$$= \tfrac{1}{2} + \tfrac{1}{2}\mathrm{Tr}\big(-A_0A_1A_0A_1\sigma\big) \tag{204}$$

$$= \tfrac{1}{2} + \tfrac{1}{2}\big\langle -A_1A_0, A_0A_1\big\rangle_\sigma. \tag{205}$$

Therefore,

$$\Re\langle E, I \otimes A_1\rangle_\rho = \tfrac{1}{2} + \tfrac{1}{2}\Re\big\langle -A_1A_0, A_0A_1\big\rangle_\sigma \geq \tfrac{1}{2} + \tfrac{1}{2}(1 - c\epsilon) = 1 - \tfrac{1}{2}c\epsilon. \tag{206}$$

It follows that

$$\|E - I \otimes A_1\|_\rho^2 = 2\Big(1 - \big(1 - \tfrac{1}{2}c\epsilon\big)\Big) = c\epsilon. \tag{207}$$

□

11.3.2 Canonical form via an isometric transformation

The coordinate system change in the previous section is expressed in terms of local unitary operators $U_A, U_B : \mathbb{C}^2 \otimes \mathbb{C}^d \to \mathbb{C}^2 \otimes \mathbb{C}^d$; however, a more natural way of expressing our canonical form is in terms of local isometries $V_A, V_B : \mathbb{C}^d \to \mathbb{C}^2 \otimes \mathbb{C}^d$.

We begin by rewriting Eq. (193) as

$$(Z \otimes I_d)U_A = U_A(Z \otimes A_0) \tag{208}$$

and Eq. (196) as

$$\|(X \otimes I_d)U_A - U_A(I \otimes A_1)\|_\rho^2 \leq c\epsilon, \tag{209}$$

where $\rho = |0\rangle\langle 0| \otimes \sigma$ and $c = 16\sqrt{2}$.

Define the isometry $S : \mathbb{C}^d \to \mathbb{C}^2 \otimes \mathbb{C}^d$ as $S = |0\rangle \otimes I_d$. Expressed as a block matrix,

$$S = \begin{bmatrix} I_d \\ 0_d \end{bmatrix}. \tag{210}$$

Define $V_A, V_B : \mathbb{C}^d \to \mathbb{C}^2 \otimes \mathbb{C}^d$ as

$$V_A = U_A S \tag{211}$$
$$V_B = U_B S. \tag{212}$$

Since

$$U_A(Z \otimes A_0)S = U_A(Z \otimes A_0)(|0\rangle \otimes I_d) = U_A(|0\rangle \otimes A_0) = U_A S A_0 = V_A A_0, \tag{213}$$

we can deduce from Eq. (208) (by multiplying both sides on the right by S) that

$$(Z \otimes I_d)V_A = (Z \otimes I_d)U_A S = U_A(Z \otimes A_0)S = V_A A_0. \tag{214}$$

So, in particular, we have

$$\|(Z \otimes I_d)V_A - V_A A_0\|_\sigma^2 = 0. \tag{215}$$

Similary, since $U_A(I \otimes A_1)S = U_A(|0\rangle \otimes A_1) = U_A S A_1 = V_A A_1$, we can deduce from Eq. (209) that

$$\|(X \otimes I_d)V_A - V_A A_1\|_\sigma^2 \leq c\epsilon. \tag{216}$$

From this (and similar relationships involving B_0 and B_1), it follows that

$$\Big\|(V_A \otimes V_B)(A_0 \otimes I_d)|\psi\rangle - \big((Z \otimes I_d) \otimes (I \otimes I_d)\big)(V_A \otimes V_B)|\psi\rangle\Big\| \in O(\sqrt{\epsilon}) \tag{217}$$
$$\Big\|(V_A \otimes V_B)(A_1 \otimes I_d)|\psi\rangle - \big((X \otimes I_d) \otimes (I \otimes I_d)\big)(V_A \otimes V_B)|\psi\rangle\Big\| \in O(\sqrt{\epsilon}) \tag{218}$$
$$\Big\|(V_A \otimes V_B)(I_d \otimes B_0)|\psi\rangle - \big((I \otimes I_d) \otimes (Z \otimes I_d)\big)(V_A \otimes V_B)|\psi\rangle\Big\| \in O(\sqrt{\epsilon}) \tag{219}$$
$$\Big\|(V_A \otimes V_B)(I_d \otimes B_1)|\psi\rangle - \big((I \otimes I_d) \otimes (X \otimes I_d)\big)(V_A \otimes V_B)|\psi\rangle\Big\| \in O(\sqrt{\epsilon}). \tag{220}$$

11.4 Form of the entangled state

Since we have (with some rearranging of the registers)

$$\langle\psi|(V_A^* \otimes V_B^*)\Big(\frac{(Z \otimes Z + Z \otimes X + X \otimes Z - X \otimes X) \otimes (I_d \otimes I_d)}{4}\Big)(V_A \otimes V_B)|\psi\rangle \geq \frac{1}{\sqrt{2}} - \epsilon, \tag{221}$$

it follows that the first two qubits of $(V_A \otimes V_B)|\psi\rangle$ are in a state that is $O(\epsilon)$-close (in Euclidean distance) to a maximum eigenvector of

$$\frac{Z \otimes Z + Z \otimes X + X \otimes Z - X \otimes X}{4}. \tag{222}$$

This maximum eigenvector is the maximally entangled state that arises in the standard CHSH strategy in symmetrized form[8] which is

$$(I \otimes R)\frac{|00\rangle + |11\rangle}{\sqrt{2}} = \cos\Big(\frac{\pi}{8}\Big)\frac{|00\rangle - |11\rangle}{\sqrt{2}} + \sin\Big(\frac{\pi}{8}\Big)\frac{|01\rangle + |10\rangle}{\sqrt{2}}, \tag{223}$$

[8]The *symmetrized form* is where Alice and Bob both employ observables Z and X, instead of Bob employing H and ZHZ. The conversion between the two protocols is via R (i.e., $H = RZR$ and $ZHZ = RXR$).

where

$$R = \begin{pmatrix} \cos(\frac{\pi}{8}) & \sin(\frac{\pi}{8}) \\ \sin(\frac{\pi}{8}) & -\cos(\frac{\pi}{8}) \end{pmatrix}. \tag{224}$$

Therefore,

$$\left\| (V_A \otimes V_B)|\psi\rangle - \Big(\big(I \otimes R\big)\big(\tfrac{1}{\sqrt{2}}|00\rangle + \tfrac{1}{\sqrt{2}}|11\rangle\big)\Big) \otimes |\Phi_{\text{junk}}\rangle \right\| \in O(\sqrt{\epsilon}) \tag{225}$$

for some arbitrary state $|\Phi_{\text{junk}}\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$.

11.5 Completion of the robust rigidity proof (Theorem 11.1)

The difference between Eqns. (225), (217), (218), (219), and (220) and the conditions in Theorem 11.1 is merely that the former equations are in terms of the symmeterized canonical strategy for CHSH; whereas the theorem is stated in terms of the "standard" canonical form. If we change the isometry on Bob's side from V_B to $(R \otimes I_d)V_B$, we obtain the conditions as stated in the theorem.

11.6 An alternate approach using the Gowers-Hatami Theorem

See [18].

12 Nonlocal games with quantum input

The main purpose of this section is to exhibit a nonlocal game where the success probability only arises in the limit of infinite entanglement. That is, for some fixed nonlocal game G, there exists no finite d such that there is an optimal strategy using entanglement in $\mathbb{C}^d \otimes \mathbb{C}^d$; there is always a better strategy using a larger d.

Although there exists such a nonlocal game our in framework (as defined back in section 1.3), what we do here is extend the nonlocal game model and show that a fairly simple game exists in the extended model with the aforementioned property.

The extended model is a variant of the nonlocal game model, where the inputs are bipartite quantum states. As with our original definition, there are finitely many possible inputs; however, rather than being a finite subset of $S \times T$, the inputs are quantum states from a finite subset $\{|\phi_0\rangle, \ldots, |\phi_{m-1}\rangle\}$, where $|\phi_0\rangle, \ldots, |\phi_{m-1}\rangle \in \mathbb{C}^\ell \otimes \mathbb{C}^\ell$. The outputs are classical $(a, b) \in A \times B$ (for finite sets A and B). (In our example, $\ell = 3$, these are *two* possible inputs from $\mathbb{C}^\ell \otimes \mathbb{C}^\ell$, and the outputs are single bits.).

When we consider *entangled strategies* for quantum-input games, there can be a *resource* entangled state, which is some state $|\psi\rangle_{AB} \in \mathbb{C}^d \otimes \mathbb{C}^d$. Note that this resource entanglement is distinguished from any entanglement that might be embodied by the input data that Alice and Bob receive. Thus, for a general entangled strategy, Alice and Bob are performing local measurements on $|\phi_k\rangle_{AB} \otimes |\psi\rangle_{AB}$, where $|\phi_k\rangle_{AB}$ is the input data, and $k \in \{0, \ldots, m-1\}$ is unknown to them.

12.1 A simple illustrative example of a quantum-input nonlocal game

The quantum-input nonlocal game presented here helps illustrate our definitions.

Consider the quantum-input nonlocal game where the inputs are qubits, the outputs are bits, and the possible input states are

$$|\phi_0\rangle_{AB} = \tfrac{1}{\sqrt{2}}|0\rangle_A|0\rangle_B + \tfrac{1}{\sqrt{2}}|1\rangle_A|1\rangle_B \tag{226}$$

$$|\phi_1\rangle_{AB} = \tfrac{1}{\sqrt{2}}|0\rangle_A|0\rangle_B - \tfrac{1}{\sqrt{2}}|1\rangle_A|1\rangle_B, \tag{227}$$

where Alice receives the first qubit as her input and Bob receives the second qubit as his input. Let $a, b \in \{0, 1\}$ denote Alice and Bob's respective output bits. The winning condition is defined as:

$$a \oplus b = \begin{cases} 0 & \text{if the input is } |\phi_0\rangle \\ 1 & \text{if the input is } |\phi_1\rangle. \end{cases} \tag{228}$$

For the above example, it is not difficult to prove that neither Alice nor Bob can acquire any information about k from their local measurements. Moreover, the most trivial strategy, where they measure in the computational basis succeeds with probability only $\frac{1}{2}$.

Nevertheless, there is a simple strategy, that requires no resource entanglement: Alice and Bob each measure their qubit in the Hadamard basis.

12.2 A quantum-input nonlocal game that requires infinite entanglement

Consider the quantum-input nonlocal game (that originated in [15] and is a simplification of the earlier result in [13]) where the inputs are qutrits, the outputs are bits, and the possible input states are

$$|\phi_0\rangle_{AB} = \tfrac{1}{\sqrt{2}}|0\rangle_A|0\rangle_B + \tfrac{1}{2}\big(|1\rangle_A|1\rangle_B + |2\rangle_A|2\rangle_B\big) \tag{229}$$

$$|\phi_1\rangle_{AB} = \tfrac{1}{\sqrt{2}}|0\rangle_A|0\rangle_B - \tfrac{1}{2}\big(|1\rangle_A|1\rangle_B + |2\rangle_A|2\rangle_B\big). \tag{230}$$

Let $a, b \in \{0, 1\}$ denote Alice and Bob's respective output bits. The winning condition is defined as:

$$a \oplus b = \begin{cases} 0 & \text{if the input is } |\phi_0\rangle \\ 1 & \text{if the input is } |\phi_1\rangle. \end{cases} \tag{231}$$

At first glance, this game may look like a slight variant of the quantum-input nonlocal game of the previous section; however, the following two theorems show that this game has a remarkable property.

Theorem 12.1. *For any $\epsilon > 0$, there exists an extangled strategy attaining success probability at least $1 - \epsilon$.*

Theorem 12.2. *The local dimension of the resource entanglement required to attain success probability $1 - \epsilon$ approaches ∞ as $\epsilon \to 0$.*

Prior to proving these theorems, we introduce a concept called *embezzlement* in the next section. Following this, we will prove the theorems.

12.3 Embezzlement of entanglement

It is well known that an entangled quantum state cannot be produced by local operations alone. Embezzlement [17] is a process where an entangled state is produced by local operations from a *catalyst* state in a manner in which the catalyst is *almost* undisturbed. Associated with each such process is a precision parameter $\epsilon > 0$ and the catalyst is preserved within fidelity $1 - \epsilon$.

The entanglement entropy of the combined state that is produced cannot exceed that of the catalyst. The name *embezzlement* reflects the fact that the protocol "steals" entanglement from the catalyst in order to produce entanglement elsewhere, but in a manner that is difficult to detect. A formal definition of embezzlement (of a $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$ state) is as follows.

Definition 12.3. *For $\epsilon \geq 0$, an ϵ-approximate embezzlement strategy for $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$ is a triple $(U_A, U_B, |\psi\rangle_{AB})$ where $U_A, U_B \in \mathbb{C}^2 \otimes \mathbb{C}^{d \times d}$ are unitary and $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ such that the fidelity between*

$$|\phi_{\text{ideal}}\rangle_{AB} = \big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big) \otimes |\psi\rangle_{AB} \tag{232}$$

and

$$|\phi_{\text{actual}}\rangle_{AB} = (U_A \otimes U_B)\big(|00\rangle_{AB} \otimes |\psi\rangle_{AB}\big) \tag{233}$$

is at least $1 - \epsilon$.

The above definition is for the target state $\frac{1}{\sqrt{2}}|00\rangle_{AB} + \frac{1}{\sqrt{2}}|11\rangle_{AB}$; there is a similar definition any other target state[9].

12.3.1 Strategy for embezzling $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$

Here we show that, for any $\epsilon > 0$, there exists a simple construction of an ϵ-approximate embezzlement strategy for $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$. Our construction uses entanglement dimension $\exp\big(O(1/\epsilon)\big)$.

Let $m \in \mathbb{N}$ be given (we will set the value of m as a function of ϵ later). For each $k \in \{0, 1, \ldots, m\}$, define the 2-qubit bipartite state

$$|\Theta_k\rangle = \cos\Big(\frac{k\pi}{4m}\Big)|00\rangle_{AB} + \sin\Big(\frac{k\pi}{4m}\Big)|11\rangle_{AB}, \tag{234}$$

where Alice has the first qubit and Bob has the second qubit.

Note that $|\Theta_0\rangle = |00\rangle_{AB}$ and $|\Theta_m\rangle = \frac{1}{\sqrt{2}}|00\rangle_{AB} + \frac{1}{\sqrt{2}}|11\rangle_{AB}$. Also, for all $k \in \{1, \ldots, m\}$,

$$\langle\Theta_k|\Theta_{k-1}\rangle = \cos(\pi/4m) > 1 - c/m^2, \tag{235}$$

for some constant $c > 0$.

Let Alice and Bob's shared entanglement consist of m qubits each, in states $|\Theta_1\rangle, \ldots, |\Theta_m\rangle$. I.e., their shared state is $|\Psi\rangle_{AB} = |\Theta_1\rangle \otimes \cdots \otimes |\Theta_m\rangle$.

We will define local unitaries U_A and U_B (each acting on $m + 1$ qubits) such that

$$(U_A \otimes U_B)|00\rangle_{AB} \otimes |\Psi\rangle_{AB} = \big(|00\rangle_{AB} + |11\rangle_{AB}\big) \otimes |\Psi'\rangle_{AB}, \tag{236}$$

[9] In fact, [17] considers a notion of a "universal" embezzling state.

where $\langle\Psi|\Psi'\rangle > 1 - c/m$. Let U_A and U_B each be the right cyclic shift (of the $m+1$ qubits), i.e., the mapping on computational basis states $|x_0, x_1, x_2, \ldots, x_m\rangle \mapsto |x_m, x_0, x_1, \ldots, x_{m-1}\rangle$. Then, since

$$(U_A \otimes U_B)|\Theta_0\rangle \otimes |\Theta_1\rangle \otimes \cdots \otimes |\Theta_m\rangle = |\Theta_m\rangle \otimes |\Theta_0\rangle \otimes \cdots \otimes |\Theta_{m-1}\rangle, \tag{237}$$

it follows that

$$(U_A \otimes U_B)|00\rangle_{AB} \otimes |\Psi\rangle_{AB} = \big(|00\rangle_{AB} + |11\rangle_{AB}\big) \otimes |\Psi'\rangle_{AB}, \tag{238}$$

where $|\Psi'\rangle_{AB} = |\Theta_0\rangle \otimes \cdots \otimes |\Theta_{m-1}\rangle$. Note that

$$\langle\Psi|\Psi'\rangle = \langle\Theta_1|\Theta_0\rangle \cdots \langle\Theta_m|\Theta_{m-1}\rangle > \big(1 - c/m^2\big)^m > 1 - c/m. \tag{239}$$

In order to obtain fidelity $1 - \epsilon$, it suffices to set $m = c/\epsilon$ in the above, in which case the local dimension of the catalyst state is $2^{c/\epsilon} = \exp\big(O(1/\epsilon)\big)$.

12.3.2 Entanglement cost of embezzling $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$

Here we show that, in order to attain success probability $\epsilon > 0$, the local dimension of the catalyst must be $\exp\big(\Omega(1/\sqrt{\epsilon})\big)$ (i.e., $\geq \exp\big(c/\sqrt{\epsilon}\big)$ for some constant $c > 0$).

Our proof is based on the continuity of the von Neumann entropy function as embodied in the so-called Fannes inequality, which states that, for any density operators $\rho, \sigma \in \mathbb{C}^{d \times d}$,

$$|S(\rho) - S(\sigma)| \leq \delta \log d + O\big(\delta \log(1/\delta)\big), \tag{240}$$

where $S(\cdot)$ is the von Neumann entropy function, and $\delta = |\rho - \sigma|_1$ (the trace distance between δ and σ). Note that, for any fixed d, the right side of Eq (240) approaches zero as $\delta \to 0$ (where we use the fact that $\lim_{\delta \to 0} \delta \log(1/\delta) = 0$). Intuitively, Eq. (240) says that, as two states approach each other in trace distance, their von Neumann entropies also approach each other.

If terms of fidelity, if the fidelity between ρ and σ is $\geq 1 - \epsilon$ then their trace distance δ satisfies $\delta \leq \sqrt{1 - (1 - \epsilon)^2} \leq \sqrt{2\epsilon}$. Therefore, two states being close to each other in terms of fidelity also have von Neumann entropies close to each other.

Now consider the states

$$\rho = \mathrm{Tr}_B\Big(|\phi_{\text{ideal}}\rangle\langle\phi_{\text{ideal}}|_{AB}\Big) \tag{241}$$

$$\sigma = \mathrm{Tr}_B\Big(|\phi_{\text{actual}}\rangle\langle\phi_{\text{actual}}|_{AB}\Big), \tag{242}$$

where $|\phi_{\text{ideal}}\rangle_{AB}$ and $|\phi_{\text{actual}}\rangle_{AB}$ are as in defnition 12.3. That is, ρ is the density operator of the idealized state that the embezzlement strategy is approximating, but with all the data on Bob's side traced out. And σ is the density operator of the actual state that is the output of the embezzlement strategy, again with all the data on Bob's side traced out. We now show that, for catalyst state $|\psi\rangle_{AB}$,

$$S(\rho) = 1 + S\Big(\mathrm{Tr}_B\big(|\psi\rangle\langle\psi|_{AB}\big)\Big) \tag{243}$$

$$S(\sigma) = S\Big(\mathrm{Tr}_B\big(|\psi\rangle\langle\psi|_{AB}\big)\Big). \tag{244}$$

Eq. (243) holds because

$$\rho = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \otimes \mathrm{Tr}_B\big(|\psi\rangle\langle\psi|_{AB}\big) \tag{245}$$

and Eq. (244) holds because local unitary operations U_A and U_B do not affect the Schmidt coefficients—and hence do not affect the value of $S(\cdot)$.

Therefore, for an ϵ-approximate embezzlement strategy, the left side of inequality (240) is 1 whereas, the right side approaches 0 as $\epsilon \to 0$. How is that possible? It is possible, because of the dimensional factor $\log d$ in Eq. (240). As $\epsilon \to 0$, it must be the case the $d \to \infty$ in order for Eq. (240) to hold. Quantitatively, we have

$$1 \leq \sqrt{\epsilon} \log d + O(\sqrt{\epsilon} \log \epsilon), \tag{246}$$

from which $d \geq \exp\big(\Omega(1/\sqrt{\epsilon})\big)$ can be deduced.

12.4 Proof of Theorem 12.1

Here we prove Theorems 12.1 from section 12.2.

We can express the inputs to the game in section 12.2 in terms of two qubits rather than trits:

$$|\phi_0\rangle_{AB} = \tfrac{1}{\sqrt{2}}|00\rangle_{AB}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big) \tag{247}$$

$$|\phi_1\rangle_{AB} = \tfrac{1}{\sqrt{2}}|00\rangle_{AB}|00\rangle_{AB} - \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big). \tag{248}$$

This is equivalent to the game in section 12.2.

A way of interpreting the states in Eqns. (247)(248) is that the first qubit of Alice/Bob indicates whether or not their second qubit is entangled or in a state $|0\rangle$. This suggests the following strategy, based on embezzlement. Let $(U_A, U_B, |\psi\rangle_{AB})$ be an an ϵ-approximate embezzlement strategy and let Alice and Bob's resource-entanglement be $|\psi\rangle_{AB}$. The initial state of the system (the input data and the resource-entanglement) is

$$|\phi_k\rangle_{AB}|\psi\rangle_{AB} = \Big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB}|00\rangle_{AB} + (-1)^k \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big)\Big)|\psi\rangle_{AB} \tag{249}$$

where $k \in \{0, 1\}$ is unknown to Alice and Bob. First, Alice (respectively, Bob) performs a controlled-U_A (resp. controlled-U_B), controlled on her (resp. his) first qubit being in state $|0\rangle$. Thus Alice and Bob's unitary operations are

$$\begin{bmatrix} U_A & 0 \\ 0 & I \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} U_B & 0 \\ 0 & I \end{bmatrix}. \tag{250}$$

The system after this operation is in state that is approximately

$$\Big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB}\big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big) + (-1)^k \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big)\Big)|\psi\rangle_{AB} \tag{251}$$

$$= \Big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + (-1)^k \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\Big)\Big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\Big)|\psi\rangle_{AB} \tag{252}$$

within fidelity $1 - \epsilon/2$. Next, Alice and Bob measure their first qubits in the Hadamard basis, to produce outputs a and b such that $a \oplus b = k$ with probability $1 - O(\sqrt{\epsilon})$. The smaller the

proximity parameter ϵ of their embezzlement strategy, the closer the success probability is to 1. This completes the proof of Theorem 12.1.

Note that the above strategy requires entanglement dimension approaching ∞ as $\epsilon \to 0$. So far, we have not ruled out the possibility of there being a different strategy for the game that attains the same success probability with entanglement size bounded by a constant; however, we next give a proof of Theorem 12.2.

12.5 Proof of Theorem 12.2

Consider any protocol for the game in section 12.2. Without loss of generality, this measurement can be expressed as a pair of local unitaries V_A and V_B followed by Alice and Bob each measuring the first qubit of their local system. The initial state of the system is $|\phi_k\rangle_{AB}|\psi\rangle_{AB}$, where $|\phi_k\rangle_{AB}$ is the input state (with $k \in \{0,1\}$ unknown) and $|\psi\rangle_{AB}$ is the resource entanglement.

Also, after applying $V_A \otimes V_B$, but before measuring, the state is of the form

$$(V_A \otimes V_B)|\phi_k\rangle_{AB}|\psi\rangle_{AB} \tag{253}$$

$$= \alpha_{00}|00\rangle_{AB}|S_{00}\rangle_{AB} + \alpha_{01}|01\rangle_{AB}|S_{01}\rangle_{AB} + \alpha_{10}|10\rangle_{AB}|S_{10}\rangle_{AB} + \alpha_{11}|11\rangle_{AB}|S_{11}\rangle_{AB}, \tag{254}$$

for some states $|S_{00}\rangle_{AB}$, $|S_{01}\rangle_{AB}$, $|S_{10}\rangle_{AB}$, $|S_{11}\rangle_{AB}$, and some $\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11} \geq 0$ such that $\alpha_{00}^2 + \alpha_{01}^2 + \alpha_{10}^2 + \alpha_{11}^2 = 1$. Since the probability that the protocol succeeds is at least $1 - \epsilon$, it must hold that

$$\begin{cases} \alpha_{00}^2 + \alpha_{11}^2 \geq 1 - \epsilon & \text{and} \quad \alpha_{01}^2 + \alpha_{10}^2 \leq \epsilon & \text{if } k = 0 \\ \alpha_{00}^2 + \alpha_{11}^2 \leq \epsilon & \text{and} \quad \alpha_{01}^2 + \alpha_{10}^2 \geq 1 - \epsilon & \text{if } k = 1. \end{cases} \tag{255}$$

Suppose that, instead of measuring this state, Alice and Bob each apply Z gates to their first qubits. Then the state becomes $|\mu\rangle_{AB}$, where

$$|\mu\rangle_{AB} = \big((Z_A \otimes I) \otimes (Z_B \otimes I)\big)(V_A \otimes V_B)|\phi_k\rangle_{AB}|\psi\rangle_{AB} \tag{256}$$

$$= \alpha_{00}|00\rangle_{AB}|S_{00}\rangle_{AB} + \alpha_{01}|01\rangle_{AB}|S_{01}\rangle_{AB} - \alpha_{10}|10\rangle_{AB}|S_{10}\rangle_{AB} - \alpha_{11}|11\rangle_{AB}|S_{11}\rangle_{AB}. \tag{257}$$

Note that

$$\begin{cases} \langle\mu|V_A \otimes V_B\big(|\phi_k\rangle_{AB}|\psi\rangle_{AB}\big) = \alpha_{00}^2 + \alpha_{11}^2 - \alpha_{01}^2 - \alpha_{10}^2 \geq 1 - 2\epsilon & \text{if } k = 0 \\ \langle\mu|V_A \otimes V_B\big(-|\phi_k\rangle_{AB}|\psi\rangle_{AB}\big) = -\alpha_{00}^2 - \alpha_{11}^2 + \alpha_{01}^2 + \alpha_{10}^2 \geq 1 - 2\epsilon & \text{if } k = 1. \end{cases} \tag{258}$$

Therefore, applying $V_A^* \otimes V_B^*$ to $|\mu\rangle_{AB}$ yields

$$\begin{cases} |\phi_k\rangle_{AB}|\psi\rangle_{AB} \text{ within fidelity } 1 - 2\epsilon & \text{if } k = 0 \\ -|\phi_k\rangle_{AB}|\psi\rangle_{AB} \text{ within fidelity } 1 - 2\epsilon & \text{if } k = 1. \end{cases} \tag{259}$$

To summarize, for $k \in \{0,1\}$,

$$\big(V_A^* \otimes V_B^*\big)\big((Z_A \otimes I) \otimes (Z_B \otimes I)\big)(V_A \otimes V_B)|\phi_k\rangle_{AB}|\psi\rangle_{AB} = (-1)^k|\phi_k\rangle_{AB}|\psi\rangle_{AB} \tag{260}$$

within fidelity $1 - 2\epsilon$.

It follows that if we apply the local operations $\big(V_A^* \otimes V_B^*)((Z_A \otimes I) \otimes (Z_B \otimes I)\big)(V_A \otimes V_B)$ to the superpostion

$$\Big(\tfrac{1}{\sqrt{2}}|\phi_0\rangle_{AB} + \tfrac{1}{\sqrt{2}}|\phi_1\rangle_{AB}\Big)|\psi\rangle_{AB} \tag{261}$$

then we obtain as output

$$\Big(\tfrac{1}{\sqrt{2}}|\phi_0\rangle_{AB} - \tfrac{1}{\sqrt{2}}|\phi_1\rangle_{AB}\Big)|\psi\rangle_{AB} \tag{262}$$

within fidelity $1 - \sqrt{2}\,2\epsilon = 1 - \sqrt{8}\epsilon$.

But, since

$$\tfrac{1}{\sqrt{2}}|\phi_0\rangle_{AB} + \tfrac{1}{\sqrt{2}}|\phi_1\rangle_{AB} = |00\rangle_{AB}|00\rangle_{AB} \tag{263}$$

$$\tfrac{1}{\sqrt{2}}|\phi_0\rangle_{AB} - \tfrac{1}{\sqrt{2}}|\phi_1\rangle_{AB} = |11\rangle_{AB}\big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big), \tag{264}$$

the local operations are mapping $|00\rangle_{AB}|00\rangle_{AB}|\psi\rangle_{AB}$ to $|11\rangle_{AB}\big(\frac{1}{\sqrt{2}}|00\rangle_{AB} + \frac{1}{\sqrt{2}}|11\rangle_{AB}\big)|\psi\rangle_{AB}$ within fidelity $1 - \sqrt{8}\epsilon$. Therefore, we have a $\sqrt{8}\epsilon$-approximate embezzling[10] strategy. Therefore, by the lower bound in section 12.3.2, the local dimension of the resource entanglement $|\psi\rangle_{AB}$ must satisfy $d \geq \exp\big(\Omega(1/\sqrt{\epsilon})\big)$. This completes the proof of Theorem 12.2.

12.6 Note about related results for nonlocal games with classical input

The above sections showed that there is a quantum-input nonlocal game, based on the idea of quantum embezzlement, which has the property that its maximum success probability only occurs in the limit of infinite entanglement. Does this have any impact on the *standard* nonlocal game framework (where the inputs are classical)? In Ref. [10], such a classical game (based on embezzlement) is obtained, albeit a game with three players, rather than two

13 States in tensor products of infinite dimensional Hilbert spaces

The idea that nonlocal games can exist that attain their maximum success probability only in the limit of infinite entanglement (i.e., only in the limit of strategies with ever increasing finite entanglement) suggests that the notion of "infinite entanglement" merits exploration. For example, with infinite entanglement, can "perfect embezzlement" (definition 12.3 with $\epsilon = 0$) be performed? Can the quantum-input nonlocal game defined in section 12.2 be won with success probability 1? It turns out that the answer depends on how infinite entanglement is defined. In this section, and the next, we explore two models of infinite entanglement.

First, suppose that we extend the framework that has been considered so far by allowing our Hilbert spaces to be infinite dimensional. Every Hilbert space $\mathcal{H}$ has an *orthonormal basis*, which is a set of the form $\{|e_\gamma\rangle \in \mathcal{H} : \gamma \in \Gamma\}$ such that:

- For all $\gamma_1, \gamma_2 \in \Gamma$,

$$\langle e_{\gamma_1}|e_{\gamma_2}\rangle = \begin{cases} 1 & \text{if } \gamma_1 = \gamma_2 \\ 0 & \text{if } \gamma_1 \neq \gamma_2. \end{cases} \tag{265}$$

[10] Technically, to meet the definition of embezzlement, we should also apply local X gates to convert the $|11\rangle_{AB}$ in the output to $|00\rangle_{AB}$, which is easy to do.

- For all $v \in \mathcal{H}$, there exist unique coefficients $\alpha_\gamma \in \mathbb{C}$ (for all $\gamma \in \Gamma$) such that

$$v = \sum_{\gamma \in \Gamma} \alpha_\gamma |e_\gamma\rangle, \tag{266}$$

  where the infinite sum is defined as the limit of the sums over finite subsets of Γ (where convergence is defined in terms of the Hilbert space norm). Note that Eq. (266) implies that

$$\sum_{\gamma \in \Gamma} |\alpha_\gamma|^2 = \|v\|^2. \tag{267}$$

When a Hilbert space has an orthonormal basis of finite or countably infinite size, it is called *separable*. It makes sense to have Hilbert spaces that are not separable, i.e., where Γ is uncountably infinite; however, for such spaces, every particular element is a linear combination of some countably infinite subset of the orthonormal basis elements. This holds because the left side of Eq. (267) cannot converge if more than a countably infinite number of the coefficients α_γ are nonzero (we no not prove this here).

Suppose that we have two separate systems, one with Hilbert space $\mathcal{H}_A$ and orthonormal basis $\{|e_\gamma\rangle : \gamma \in \Gamma\}$ and the other with Hilbert space $\mathcal{H}_B$ and orthonormal basis $\{|e'_{\gamma'}\rangle : \gamma' \in \Gamma'\}$. Then the compound system has Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ associated with it, whose orthonormal basis is $\{|e_\gamma\rangle \otimes |e'_{\gamma'}\rangle : \gamma \in \Gamma \text{ and } \gamma' \in \Gamma'\}$.

Every state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ has a *Schmidt decomposition* of the form

$$|\psi\rangle = \sum_{k \in \mathbb{N}} \alpha_k |\phi_k\rangle \otimes |\mu_k\rangle, \tag{268}$$

where $|\phi_1\rangle, |\phi_2\rangle, \cdots \in \mathcal{H}_A$ are orthonormal, $|\mu_1\rangle, |\mu_2\rangle, \cdots \in \mathcal{H}_B$ are orthonormal, and

$$\sum_{k \in \mathbb{N}} |\alpha_k|^2 = 1. \tag{269}$$

Note that the number of Schmidt coefficients is countable, even if $\mathcal{H}_A$ and $\mathcal{H}_B$ have uncountably infinite dimension; a proof of this is in Appendix A of [5].

This can be deemed as "infinitely entangled" if it has infinite Schmidt rank (i.e., infinitely many of the α_k are nonzero), or if its *entanglement entropy*, defined as

$$\sum_{k \in \mathbb{N}} |\alpha_k|^2 \log |\alpha_k|^2, \tag{270}$$

is infinite.

It turns out that perfect embezzlement is not possible with such a catalyst state. A rough sketch of the proof of this is that $|00\rangle_{AB}|\psi\rangle_{AB}$ has Schmidt coefficients $\lambda_1, \lambda_2, \lambda_3 \ldots$, whereas $\left(\frac{1}{\sqrt{2}}|00\rangle_{AB} + \frac{1}{\sqrt{2}}|11\rangle_{AB}\right)|\psi\rangle_{AB}$ has Schmidt coefficients $\frac{1}{\sqrt{2}}\lambda_1, \frac{1}{\sqrt{2}}\lambda_1, \frac{1}{\sqrt{2}}\lambda_2, \frac{1}{\sqrt{2}}\lambda_2, \ldots$, and local operations cannot affect the Schmidt coefficients. In a similar spirit, it can be proved that no entangled state in the tensor product of two Hilbert spaces (however large their dimension) can be used to play the game in section 12.2 with success probability 1.

So far, we have some negative results, indicating what this infinite entanglement *cannot* accomplish. It turns out that we can accomplish interesting things with states that can be informally thought of as

$$\left(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\right) \otimes \left(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\right) \otimes \left(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\right) \otimes \cdots, \quad (271)$$

by which we mean Alice and Bob each have a countably infinite number of qubits and, for each $k \in \mathbb{N}$, Alice's k-th qubit is entangled with Bob's k-th qubit. However, such as state does not have a Schmidt decomposition of the form of Eq. (268). In fact, such a state cannot be expressed as a vector in the tensor product of *any* two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ (even if they are uncountably infinite dimensional). We need a different formalism to capture states like the above (as well as some other interesting states), which we develop in the next section.

14 Abstract states on C*-algebras

Our goal here is to explain a notion of entanglement that goes beyond the scope of what can be expressed in the tensor product model. It is natural to express this notion using an object called a C*-algebra (but we do not need to use the theory of C*-algebras in a deep way).

For readers unfamiliar with C*-algebras (which will be formally defined in section 14.1), let us begin by considering what they are in broad terms and how one is naturally led to such an object by the mathematics of quantum information.

First, it is clear that much of the mathematics that arises in quantum information concerns Hilbert spaces and linear operators acting on them. And a Hilbert space is a vector space with a norm and additional structure arising from the norm (such as being topologically complete).

In the context of linear algebra, an object called an *algebra* (or sometimes a *linear algebra*) is an abstraction that captures the structure of linear operators on a vector space, but without the explicit presence of a vector space. A *C*-algebra* can be viewed as simply an analogue of an algebra, but for operators on a Hilbert space rather than for operators on a vector space. In other words, a C*-algebra is an abstraction that captures the structure of a set of operators acting on a Hilbert space, but without the explicit presence of a Hilbert space.

Information about the theory of C*-algebras can be found in [2, 8, 11]. The next few subsections is a brief primer on C*-algebras as they pertain to our notions of entanglement.

14.1 Definition of a C*-algebra

A *C*-algebra* $\mathcal{A}$ is a set with the following properties:

- $\mathcal{A}$ is an *algebra*. This means it is a vector space that also has a multiplication operation, where $a(bc) = (ab)c$, and $a(b + c) = ab + ac$, for all $a, b, c \in \mathcal{A}$.
- $\mathcal{A}$ also has a **-map*, such that $a^{**} = a$, $(a + \lambda b)^* = a^* + \overline{\lambda} b^*$, and $(ab)^* = b^* a^*$, for all $a, b \in \mathcal{A}$ and $\lambda \in \mathbb{C}$. An algebra with such a *-map is called a *$*$-algebra*.
- $\mathcal{A}$ also has a *norm* $\|\cdot\|$ satisfying $\|a\| \geq 0$ (with equality if and only if $a = 0$), $\|a + b\| \leq \|a\| + \|b\|$, and $\|ab\| \leq \|a\|\|b\|$, for all $a, b, c \in \mathcal{A}$; and $\mathcal{A}$ is complete with respect to the norm. An algebra with such a norm is called a *Banach algebra*.

- The norm and *-map satisfy $\|a^*a\| = \|a\|^2$, for all $a \in \mathcal{A}$.

All the C*-algebras that we are interested in are *unital*, which means they have a multiplicative identity, that we denote as I.

14.1.1 Example 1: the concrete C*-algebra of operators on a Hilbert space

For a Hilbert space $\mathcal{H}$, define $B(\mathcal{H})$ as the set of all bounded linear operators on $\mathcal{H}$, where a linear operator A is *bounded* if

$$\sup_{\substack{|\psi\rangle \in \mathcal{H} \\ \||\psi\rangle\| = 1}} \big\| A|\psi\rangle \big\| \text{ is finite.} \tag{272}$$

The set $B(\mathcal{H})$ is a C*-algebra. More generally, any subset $\mathcal{A} \subseteq B(\mathcal{H})$ that is closed with respect to the algebraic operations (i.e., linear combinations, product, and the *-map) and closed with respect to the norm is a C*-algebra.

14.1.2 Example 2: the CAR algebra

The so-called CAR algebra is named as an abbreviation of "canonical anti-commutation relations" (see [8] for the standard mathematical construction of this algebra).

Here, we show how to construct the CAR algebra in terms of infinite tensor products of Pauli operators of finite weight, where the Pauli operators are $I = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $X = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $Z = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $W = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) = XZ$, and the weight of such an infinite tensor product is the number of instances of X, Z, or W. For example, $I \otimes X \otimes W \otimes I \otimes Z \otimes I \otimes I \otimes \cdots$ has weight 3. We can denote each such operator as $X^a Z^b$, where $a, b \in \{0,1\}^*$, where it is understood that each string is padded on the right with an infinite sequence of 0s. Thus, $X^a Z^b = (X^{a_1} \otimes X^{a_2} \otimes \cdots)(Z^{b_1} \otimes Z^{b_2} \otimes \cdots)$. The above example is $X^a Z^b$, where $a = 011 \equiv 011000\ldots$ and $b = 00101 \equiv 00101000\ldots$. Define the set of generators $G = \{X^a Z^b : a, b \in \{0,1\}^*\}$ and $\mathbb{C}G$ to be the set of all (finite) linear combinations[11] of elements of G. $\mathbb{C}G$ is closed under multiplication and is a $*$-algebra. (Note that $\{\pm X^a Z^b : a, b \in \{0,1\}^*\} \subset \mathbb{C}G$ is a multiplicative group that we can think of as an infinite version of the Pauli group; however, G itself is not closed under multiplication.)

For each element $A \in \mathbb{C}G$, there is an $m \in \mathbb{N}$ and $M \in \mathbb{C}^{2^m \times 2^m}$ such that $A = M \otimes I \otimes I \otimes \cdots$. Define a norm on $\mathbb{C}G$ as $\|A\| = \|M\|$ (i.e., the spectral norm of M as an operator on $\mathbb{C}^{2^m}$). The CAR algebra is the completion of $\mathbb{C}G$ with respect to this norm.

Note that, in the aforementioned description of the elements of the generating set G as $X^a Z^b$, we have used $\mathbb{N}$ as the index set for the bits of $a = a_1 a_2 \ldots$ and $b = b_1 b_2 \ldots$; however, any countably infinite set may be used. It is sometimes convenient to use $\mathbb{Z}$ as the index set, which corresponds to thinking of the infinite tensor products of Paulis as *two-way infinite* strings.

14.1.3 Positive elements

We say that $a \in \mathcal{A}$ is *positive*, denoted $a \geq 0$, if there exists $b \in \mathcal{A}$ such that $a = b^*b$. Note that, for the case of concrete C*-algebra of operators on a Hilbert space, this is consistent with the standard definition of positive semidefinite.

[11] This is well-defined because there are finitely many terms, each of which has all but finitely many factors of I.

14.1.4 $*$-isomorphisms and $*$-automorphisms

Definition 14.1. *A* $*$-isomorphism from $\mathcal{A}_1$ to $\mathcal{A}_2$ *(C*-algebras) is a map* $\pi : \mathcal{A}_1 \to \mathcal{A}_2$ *that preserves the algebraic structure and the norm. That is,* $\pi(a+\lambda b) = \pi(a)+\lambda\pi(b)$, $\pi(ab) = \pi(a)\pi(b)$, $\pi(a^*) = \pi(a)^*$, *and* $\|\pi(a)\| = \|a\|$, *for all* $a, b \in \mathcal{A}_1$ *and* $\lambda \in \mathbb{C}$.

It can be shown that all $*$-isomorphisms are injective.

A remarkable result, commonly known as the *GNS Theorem*[12], is that every C*-algebra can be viewed as a subset of the bounded operators on some Hilbert space in the following sense.

Theorem 14.2 (Gelfand, Naimark [9]; Segal [16])**.** *For any C*-algebra* $\mathcal{A}$*, there exists a Hilbert space* $\mathcal{H}$ *and a* $*$*-isomorphism* $\pi : \mathcal{A} \to B(\mathcal{H})$.

However, it should be noted that, in general, the underlying Hilbert space can be more difficult to describe directly than the C*-algebra.

Definition 14.3. *A* *-automorphism *of a C*-algebra* $\mathcal{A}$ *is a* $*$*-isomorphism from* $\mathcal{A}$ *to itself.*

An example of a $*$*-automorphism* $\alpha : \mathcal{A} \to \mathcal{A}$ is conjugation by some unitary $u \in \mathcal{A}$ (where unitary means $uu^* = u^*u = I$). That is, $\alpha_u(a) = u^*au$. These are called *inner* automorphisms. Automorphisms that are not inner are called *outer* automorphisms. An example of an outer automorphism for the CAR algebra is the "bilateral-shift of the qubits" operation that maps $X^a Z^b$ (where $a, b : \mathbb{Z} \to \{0, 1\}$) to $X^{a'} Z^{b'}$, where $a'_j = a_{j+1}$ and $b'_j = b_{j+1}$. More generally *any* permutation of the index set corresponds to a $*$-automorphism.

14.2 Definition of a state

Let us first note that a desirable property of a "state" is to assign meaningful values to all potential POVM measurement elements; in other words, given a state, the notion of a POVM measurement on it should be well-defined.

Definition 14.4. *An* (abstract) state *on a C*-algebra* $\mathcal{A}$ *is a linear functional* $s : \mathcal{A} \to \mathbb{C}$ *with two additional properties:*

- s *is* positive*. By this we mean that, for* $a \in \mathcal{A}$*, if* $a \geq 0$ *then* $s(a) \geq 0$.
- s *is* unital*. By this we mean that* $s(I) = 1$.

Note that, in the case of a concrete C*-algebra $\mathcal{A} \subseteq B(\mathcal{H})$, for any unit vector $|\psi\rangle \in \mathcal{H}$, defining $s : \mathcal{A} \to \mathbb{C}$ as

$$s(A) = \langle\psi|A|\psi\rangle \tag{273}$$

results in a state in the sense of definition 14.4. Also, for a density operator $\rho \in B(\mathcal{H})$, defining

$$s(A) = \mathrm{Tr}(A\rho) \tag{274}$$

results in a state in the sense of definition 14.4.

[12] There are different statements of this theorem, which all imply the statement given here.

14.3 Definition of a POVM measurement

Given a C*-algebra $\mathcal{A}$, a POVM measurement for the system is $a_1, \ldots, a_m \in \mathcal{A}$ such that $a_1, \ldots, a_m \geq 0$ and $\sum_k a_k = I$. (It is also possible to define POVM measurements with an infinite number of outcomes, but we do not do so here.)

Given a state $s : \mathcal{A} \to \mathbb{C}$, *applying* such a POVM to this state results in an *outcome* $k \in \{1, \ldots, m\}$, where the respective outcome probabilities are $s(a_1), \ldots, s(a_m)$.

This measurement is destructive in the sense that the state no longer exists after the measurement is performed. In fact there is not even a well-defined "collapsed state".

14.4 Definition of a reversible operation

In order to have a full-fledged model of information processing, we also need to include *dynamics*, such as the unitary operations that arise in conventional quantum information[13].

14.4.1 Inner automorphisms (as reversible operations on states)

The simplest kind of reversible operation that we can define in the context of C*-algebraic registers are the *unitary operations*. What is the effect of *applying* a unitary $u \in \mathcal{A}$ to a state $s : \mathcal{A} \to \mathbb{C}$?

Definition 14.5. *For a unitary $u \in \mathcal{A}$,* applying *u to an input state $s : \mathcal{A} \to \mathbb{C}$ produces as output the state $s' : \mathcal{A} \to \mathbb{C}$, where*

$$s'(a) = s(u^* a u). \tag{275}$$

It is not difficult to show that, in the general case of an abstract C*-algebra, the output s' satisfies the condition of being a state.

Note that definition 14.5 makes sense in the concrete case, where $s(A) = \langle\psi|A|\psi\rangle$, and U is a unitary operator on $\mathcal{H}$. Intuitively, there are two equivalent ways of viewing

$$s'(A) = \langle\psi|U^* A U|\psi\rangle. \tag{276}$$

We can view the state vector as changing from $|\psi\rangle$ to $U|\psi\rangle$; or we can imagine that the state vector stays put but the measurement operator changes from A to $U^* A U$ (this latter perspective is sometimes referred to as the "Heisenberg picture").

14.4.2 Outer automorphisms (as reversible operations on states)

Not all reversible operations are of the form of definition 14.5. In general, we can take any $*$-automorphism $\pi : \mathcal{A} \to \mathcal{A}$ and obtain a reversible operation as follows.

Definition 14.6. *For a $*$-automorphism $\pi : \mathcal{A} \to \mathcal{A}$,* applying *$\pi$ to an input state $s : \mathcal{A} \to \mathbb{C}$ produces as output the state $s' : \mathcal{A} \to \mathbb{C}$, where*

$$s'(a) = s(\pi(a)). \tag{277}$$

[13] There are other kinds of operations, such as channels, or other notions of measurement that produce an output state in addition to classical information. It is interesting to consider these, though they can in principle be modelled in terms of unitary operations on larger systems and POVM measurements. For simplicity, we focus on only POVM measurements and reversible operations here.

A unitary operation is a special case of a $*$-automorphism defined as $\pi(a) = u^* a u$. Any $*$-automorphism that can be so expressed is called an *inner automorphism*; otherwise, it is called an *outer automorphism*.

Recall that, for the CAR algebra, the "bilateral left shift of the qubits" (explained in section 14.1.4) is an example of an outer automorphism. Intuitively, it can be thought of as shifting all the qubits left by one position.

14.5 Definition of a register

We now have the ingredients to define a *register* associated with a C*-algebra. In operational terms:

- The system can be *set* to any abstract state on the C*-algebra.
- The state can be *modified* by applying any $*$-automorphism.
- The state can be *measured* by applying a POVM measurement to it.

There is no well-defined residual state after a POVM measurement, so this particular operational framework requires a reset after a measurement in order to be meaningful. (There are various ways of enhancing the framework to include a well-defined "collapsed state" after a measurement, but we do not pursue this here.)

14.6 Definition of a compound register

If A and B are two C*-algebraic registers with associated C*-algebras $\mathcal{A}$ and $\mathcal{B}$ then we can consider the *compound register* (A, B), which is the joint system. The C*-algebra associated with (A, B) is the tensor product of $\mathcal{A}$ and $\mathcal{B}$, which requires some discussion.

In general, there are multiple ways of defining the tensor product of two C*-algebras. Each tensor product is defined as the completion of the $*$-algebra generated by the set of formal tensor products

$$\mathcal{A} \otimes \mathcal{B} = \{a \otimes b : a \in \mathcal{A} \text{ and } b \in \mathcal{B}\}, \tag{278}$$

according to some norm on $\mathcal{A} \otimes \mathcal{B}$ (note that $\mathcal{A} \otimes \mathcal{B}$ itself is not a C*-algebra). For general C*-algebras, there are different choices for this norm, each resulting in a different tensor product C*-algebra; however, for the CAR algebra, all of these norms coincide, so there is a unique tensor product for the CAR algebra (this follows from the CAR algebra being "hyperfinite", which is defined in [8]).

Here we define the *min-norm*, resulting in the *min tensor product*[14], denoted as $\mathcal{A} \otimes_{\min} \mathcal{B}$ (another tensor product, called the *max tensor product* is important in the context of other C*-algebras, and will be introduced later, when they become relevant to us).

Definition 14.7. *For C*-algebras $\mathcal{A}$ and $\mathcal{B}$, define the* min-norm *on $\mathcal{A} \otimes \mathcal{B}$ as, for any $x \in \mathcal{A} \otimes \mathcal{B}$,*

$$\|x\|_{\min} = \sup\big\{\|(\pi_A \otimes \pi_B)(x)\| : \pi_A : \mathcal{A} \to B(\mathcal{H}_A) \text{ and } \pi_B : \mathcal{A} \to B(\mathcal{H}_B)\big\} \tag{279}$$

where $\pi_A \otimes \pi_B : \mathcal{A} \otimes \mathcal{B} \to B(\mathcal{H}_A \otimes \mathcal{H}_B)$ is the $*$-homomorphism satisfying $(\pi_A \otimes \pi_B)(A \otimes B) = \pi_A(A) \otimes \pi_B(B)$, and the supremum is over all Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ and $*$-homomorphisms $\pi_A : \mathcal{A} \to B(\mathcal{H}_A)$ and $\pi_B : \mathcal{B} \to B(\mathcal{H}_B)$.

[14] Also known as the *spatial tensor product*.

Note that the supremum is over a non-empty set because, by the GNS Theorem (Theorem 14.2), there exist $*$-homomorphisms $\pi_A : \mathcal{A} \to B(\mathcal{H}_A)$ and $\pi_B : \mathcal{B} \to B(\mathcal{H}_B)$, for some Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$.

A final remark here is that compound registers consisting of more than two components can also be defined as a straightforward extension to the above. In particular, all of the tensor products on C*-algebras are associative, e.g., $(\mathcal{A} \otimes_{\min} \mathcal{B}) \otimes_{\min} \mathcal{C} = \mathcal{A} \otimes_{\min} (\mathcal{B} \otimes_{\min} \mathcal{C})$.

14.6.1 Product states

Given abstract states $s_A : \mathcal{A} \to \mathbb{C}$ (on register A) and $s_B : \mathcal{B} \to \mathbb{C}$ (on register B), define the *product state* $s_A \otimes s_B : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathbb{C}$ (on register (A, B)) as $(s_A \otimes s_B)(a \otimes b) = s_A(a) s_B(b)$ for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$. This extends to the full domain $\mathcal{A} \otimes_{\min} \mathcal{B}$ by the linearity and continuity of s'.

14.6.2 Localized reversible operations

Suppose that A and B are registers with respective C*-algebras $\mathcal{A}$ and $\mathcal{B}$. Let $s : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathbb{C}$ be a state on the compound register (A, B). Let $\pi : \mathcal{A} \to \mathcal{A}$ be a $*$-automorphism representing a reversible operation on A. Then applying π to the joint system is defined as follows. We can extend π to $\pi \otimes I : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathcal{A} \otimes_{\min} \mathcal{B}$ defined as $(\pi \otimes I)(a \otimes b) = \pi(a) \otimes b$, for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$. Then the local operation π changes s to $s' : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathbb{C}$, where $s'(x) = s\big((\pi \otimes I)(x)\big)$.

14.6.3 Partial trace

Given a state $s : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathbb{C}$, define the partial trace $\mathrm{Tr}_B[s] : \mathcal{A} \to \mathbb{C}$ as $\mathrm{Tr}_B[s](a) = s(a \otimes I)$ and $\mathrm{Tr}_A[s] : \mathcal{B} \to \mathbb{C}$ as $\mathrm{Tr}_A[s](b) = s(I \otimes b)$.

Note that $s : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathbb{C}$ is a product state if and only if $s(a \otimes b) = \mathrm{Tr}_B[s](a)\mathrm{Tr}_A[s](b)$ for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$.

14.6.4 Localized measurements

Here we address what happens if a POVM measurement on register A is performed in the context of the compound register (A, B). Let $a_1, \ldots, a_m \in \mathcal{A}$ be a POVM measurement for register A. Applying this measurement to the state $s : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathbb{C}$ produces two items: a classical outcome $k \in \{1, \ldots, m\}$, where each k arises with probability $s(a_k \otimes I)$; and a corresponding residual quantum state on B, which is $s_k : \mathcal{B} \to \mathbb{C}$ defined as $s_k(b) = s(a_k \otimes b)/s(a_k \otimes I)$, for all $b \in \mathcal{B}$.

14.7 Embezzlement in the C*-algebraic model

Here we show how to perform perfect embezzlement, as well as coherent embezzlement, in the C*-algebraic model. Intuitively, the idea is simple: informally take the state

$$\big(|00\rangle_{AB}\big)^{\otimes\infty}\big(|00\rangle_{AB}\big)\big(\tfrac{1}{\sqrt{2}}|00\rangle_{AB} + \tfrac{1}{\sqrt{2}}|11\rangle_{AB}\big)^{\otimes\infty} \tag{280}$$

and Alice and Bob each perform a left shift of their qubits. The middle pair of qubits changes from state $|00\rangle$ to state $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$, while all the other qubits remain in the same state. This cannot be formalized in terms of state vectors in tensor products of Hilbert spaces; however, it can be formalized as abstract states on tensor products of C*-algebras.

14.7.1 Expressing $\left(\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\right)^{\otimes\infty}$ as an abstract state

We first define a state that can be intuitively thought of as a countably infinite tensor product of states of the form $|\phi^+\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$. As noted in section 13, it is impossible to express such a state as a vector in the tensor product of two Hilbert spaces (even if the Hilbert spaces are allowed to have uncountably infinite dimension). However, as was essentially pointed out in [12], such a state *can* be defined as an abstract state of two CAR-algebra registers $s_{\phi^+} : \mathcal{A} \otimes_{\min} \mathcal{A} \to \mathbb{C}$. Set

$$s_{\phi^+}((X^a Z^b) \otimes (X^{a'} Z^{b'})) = \prod_{j=1}^{\infty} \langle \phi^+ | (X^{a_j} Z^{b_j}) \otimes (X^{a'_j} Z^{b'_j}) | \phi^+ \rangle = \prod_{j=1}^{\infty} \delta_{a_j, a'_j} \delta_{b_j, b'_j}, \tag{281}$$

where δ is the Kronecker delta function. In [12], such a state is described as an example of the notion of an "infinitely entangled state" and several of the properties of this state are explained.

14.7.2 Expressing $|00\rangle^{\otimes\infty}$ as an abstract state

We also define an abstract state $s_{00} : \mathcal{A} \otimes_{\min} \mathcal{A} \to \mathbb{C}$ that corresponds to an infinite tensor product of $|00\rangle = |0\rangle \otimes |0\rangle$ states as

$$s_{00}((X^a Z^b) \otimes (X^{a'} Z^{b'})) = \prod_{j=-1}^{-\infty} \langle 00 | (X^{a_j} Z^{b_j}) \otimes (X^{a'_j} Z^{b'_j}) | 00 \rangle = \prod_{j=-1}^{-\infty} (1 - a_j)(1 - a'_j). \tag{282}$$

Notice that are using the index set $-\mathbb{N} = \{-1, -2, \dots\}$ here (so as to be disjoint from the indices used in the definition of s_{ϕ^+} in section 14.7.1.

14.7.3 Embezzlement strategy

The resource-entanglement is $s_{\text{catalyst}} = s_{00} \otimes s_{\psi^+}$ (the combination of s_{00} and s_{ψ^+}). That is,

$$s_{\text{catalyst}}((X^a Z^b) \otimes (X^{a'} Z^{b'})) = \prod_{j=-1}^{-\infty} (1 - a_j)(1 - a'_j) \prod_{j=1}^{\infty} \delta_{a_j, a'_j} \delta_{b_j, b'_j}, \tag{283}$$

where $a, b, a', b' : (-\mathbb{N}) \cup \mathbb{N} \to \mathbb{C}$.

The initial state is $|00\rangle$ and the target state is $|\phi^+\rangle$.

If we express the combination of initial state and resource-entanglement by placing the $|00\rangle$ state "between" the s_{00} and s_{ψ^+} components of the resource state—using index value 0—then we obtain

$$s_{\text{initial}}((X^a Z^b) \otimes (X^{a'} Z^{b'})) = \prod_{j=0}^{-\infty} (1 - a_j)(1 - a'_j) \prod_{j=1}^{\infty} \delta_{a_j, a'_j} \delta_{b_j, b'_j}, \tag{284}$$

where $a, b, a', b' : \mathbb{Z} \to \mathbb{C}$.

In this same ordering, the final state should be

$$s_{\text{final}}((X^a Z^b) \otimes (X^{a'} Z^{b'})) = \prod_{j=-1}^{-\infty} (1 - a_j)(1 - a'_j) \prod_{j=0}^{\infty} \delta_{a_j, a'_j} \delta_{b_j, b'_j}, \tag{285}$$

which is accomplished by Alice and Bob each performing a bilateral left shift of the qubit positions. These are localized $*$-automorphisms, which are explained is section 14.4.2.

14.7.4 Coherent Embezzlement strategy

The proof of Theorem 12.1 is a reduction from an embezzlement strategy to a strategy for the quantum-input game in section 12.2. In our current C*-algebraic framework, we can perform embezzlement with success probability 1 (i.e., $1-\epsilon$, with $\epsilon = 0$). In this case, that reduction yields a strategy for coherent embezzlement with success probability $1 - O(\sqrt{\epsilon}) = 1$.

15 Binary linear system games with infinite dimensional operator solutions

We now return to BLS games. A *binary linear system game (BLS game)* consists of n $\{0,1\}$-valued variables $v_1, \ldots, v_n$ and m constraints, each of which specifies whether a mod-2 sum of a subset of the variables is 0 or 1. In multiplicative form, the variables are ± 1-valued and each constraint specifies whether or not a product of a subset of the variables is 1 or -1.

Recall that an *operator solution* to a BLS game is defined as follows.

Definition 15.1. *An* operator solution *to a BLS game is a sequence of Hermitian operators $A_1, \ldots, A_n$ on a Hilbert space $\mathcal{H}$ such that:*

- *For all j, $A_j^2 = I$ (that is, A_j is a binary observable).*
- *If variables v_j and v_k appear in the same constraint then A_j and A_k commute (we call this* local compatibility*).*
- *For each constraint of the form $v_{k_1} v_{k_2} \ldots v_{k_r} = (-1)^b$, the observables satisfy*

$$A_{k_1} A_{k_2} \cdots A_{k_r} = (-1)^b I \tag{286}$$

(we call this constraint satisfaction*).*

In section 1.4, we explained that a BLS game has a perfect strategy using entanglement in the tensor product of Hilbert spaces model if and only if it has a *finite dimensional operator solution* (i.e., where the Hilbert space in the above definition is finite dimensional).

Here we consider the case where a BLS game has an *infinite dimensional operator solution* (i.e., where the Hilbert space in the above definition is finite dimensional). We will first show that in this case there is a perfect strategy in a *commuting operator* model of entanglement (which will be defined), and then that this can be translated into a perfect strategy in the C*-algebraic model (where the joint state is in the *max-tensor product* of the two registers, which will also be defined). This work is based on [6].

15.1 Solution Group of a BLS game

The definition of an operator solution is similar to a presentation of a group (where the group operation is multiplication), namely a set of generators and a set equations that they satisfy. One respect in which an operator solution differs from a group presentation is in the presence of $-I$. Although I very naturally corresponds to the group identity (that we'll refer to as 1), $-I$ does not directly correspond to a meaningful element of an abstract group (there is no -1). It turns out that we can capture the properties of -1 that are relevant for our purposes by introducing a new

generator J with these properties: $J^2 = 1$, J commutes with all the other generators, and $J \neq 1$. A group presentation is defined in terms of equalities, not inequalities, so we need to set aside the $J \neq 1$ condition for now, in the following definition (but the $J \neq 1$ condition will come up later).

Definition 15.2. *The* solution group *of a BLS game is a sequence generators $g_1, \ldots, g_n, J$ such that:*

- *For all j, $g_j^2 = 1$.*
- *If variables v_j and v_k appear in the same constraint then $g_j g_k g_j g_k = 1$ (i.e., g_j and g_k commute).*
- *For each constraint of the form $v_{k_1} v_{k_2} \ldots v_{k_r} = (-1)^b I$, we have*

$$g_{k_1} g_{k_2} \cdots g_{k_r} = J^b. \tag{287}$$

- *$J^2 = 1$ and, for all j, $g_j J g_j J = 1$ (i.e., J and g_j commute).*

It turns out that solution group of the Magic Square game has finite size, but, in general, the solution group of a BLS game can be finite or countably infinite.

For the BLS game in part (c) of figure 3, there is no perfect strategy and this can be proven by showing that there is no operator solution. The proof that there is no operator solution is by deriving $I = -I$ from the equations, a contradiction. In the language of solution groups, this is a proof that the solution group of that game has the property that $J = 1$.

In fact, we show in the following lemma that, for any BLS game, if its solution group has the property $J = 1$ then it has no operator solution. In contrapositive form, the statement is: if there is an operator solution then $J \neq 1$.

Lemma 15.3 ([6])**.** *If a BLS game has a commuting operator solution then its solution group has the property that $J \neq 1$.*

Proof. Note that any operator solution to a BLS game is a representation of its solution group that maps J to $-I$. It follows that, if $J = 1$ in the solution group, then it's operator solution has the property that $-I = I$, a contradiction. □

15.2 Definition of a commuting operator strategy

Definition 15.4. *A* commuting operator scenario *is a triple $(\mathcal{H}, \mathcal{A}, \mathcal{B})$, where $\mathcal{H}$ is a Hilbert space and $\mathcal{A}, \mathcal{B} \subseteq B(\mathcal{H})$ are two C*-algebras such that every operator in $\mathcal{A}$ commutes with every operator in $\mathcal{B}$.*

Note that, whenever $\mathcal{H}$ has a tensor product structure $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, $\mathcal{A} = B(\mathcal{H}_A) \otimes I$ and $\mathcal{B} = I \otimes B(\mathcal{H}_B)$ are mutually commuting, so we have a commuting operator scenario.

Definition 15.5. *A* commuting operator strategy *for a nonlocal game exists in the context of a commuting operator scenario $(\mathcal{H}, \mathcal{A}, \mathcal{B})$, and consists of a state $|\psi\rangle \in \mathcal{H}$ and measurement operators for Alice in $\mathcal{A}$ and measurement operators for Bob in $\mathcal{B}$.*

15.3 Operator solution implies a perfect commuting operator strategy

Our main result here is the following.

Theorem 15.6 ([6]). *If a BLS game has an operator solution then there exists a perfect commuting-operator strategy for the game.*

Due to Lemma 15.3, it suffices to prove the following lemma.

Lemma 15.7 ([6]). *If a BLS game has the property that, in its solution group, $J \neq 1$ then there is a perfect commuting-operator strategy for the game.*

Prior to proving Lemma 15.7, we establish the following definitions. Let G denote any group whose size is finite or countably infinite.

Definition 15.8. *The* group algebra *of G is the $*$-algebra of formal linear combinations of group elements*

$$\mathbb{C}G = \left\{ \sum_{j=1}^{m} \alpha_j g_j : m \in \mathbb{N},\, g_1, \ldots, g_m \in G, \text{ and } \alpha_1, \ldots, \alpha_m \in \mathbb{C} \right\}, \tag{288}$$

where multiplication is in terms of the group multiplication, i.e.,

$$\left(\sum_{j=1}^{m} \alpha_j g_j \right) \left(\sum_{k=1}^{m} \alpha'_k g'_k \right) = \sum_{j=1}^{m} \sum_{k=1}^{m} \alpha_j \alpha'_k \, (g_j g'_k), \tag{289}$$

*and the *-operation is in terms of group inverses, i.e.,*

$$\left(\sum_{j=1}^{m} \alpha_j g_j \right)^* = \sum_{j=1}^{m} \overline{\alpha}_j (g_j)^{-1}. \tag{290}$$

Definition 15.9. *$\ell^2(G)$ denotes the Hilbert space*

$$\ell^2(G) = \left\{ \sum_{g \in G} \lambda_g |g\rangle : \sum_{g \in G} |\lambda_g|^2 \text{ is finite} \right\}. \tag{291}$$

Note that, $\mathbb{C}G$ is defined in terms of true linear combinations of group elements (so the sums are finite); whereas the "linear combinations" arising in $\ell^2(G)$ are actually square summable sequences (so the sums are possibly infinite).

The group algebra $\mathbb{C}G$ acts on $\ell^2(G)$ in the following two natural ways: as multiplication from the left; and as multiplication from the right. In each case, $\mathbb{C}G \subseteq B(\ell^2(G))$ can acquire the spectral norm from $\ell^2(G)$ and the resulting closure is a C*-algebra.

Definition 15.10. *The* left regular C*-algebra *of G, denoted as $C^*_\lambda(G)$, is the C*-algebra of operators acting on $\ell^2(G)$ that is the closure of $\mathbb{C}G$ acting on $\ell^2(G)$ by left multiplication (i.e., $g|h\rangle \mapsto |gh\rangle$).*

Definition 15.11. *The* right regular C*-algebra *of G, denoted as $C^*_\rho(G)$, is the C*-algebra of operators acting on $\ell^2(G)$ that is the closure of $\mathbb{C}G$ acting on $\ell^2(G)$ by right multiplication (i.e., $g|h\rangle \mapsto |hg\rangle$).*

Note that $C^*_\lambda(G)$, $C^*_\rho(G) \subseteq B(\ell^2(G))$ and they commute with each other, since $|(g_1 h) g_2\rangle = |g_1(h g_2)\rangle$ for all $g_1, g_2, h \in G$.

Now we are ready to prove Lemma 15.7.

Proof of Lemma 15.7. Let G be the solution group of the BLS game. Let the commuting operator scenario be $\big(\ell^2(G),\, C^*_\lambda(G),\, C^*_\rho(G)\big)$ (where $\ell^2(G)$, $C^*_\lambda(G)$, $C^*_\rho(G)$ are defined above). For clarity, denote the generators of $C^*_\lambda(G)$ as L_g ($g \in G$), where $L_g|h\rangle = |gh\rangle$; and denote the generators of $C^*_\rho(G)$ as R_g ($g \in G$), where $R_g|h\rangle = |hg\rangle$.

For the commuting operator strategy, set the resource state to

$$|\psi\rangle = \frac{|J\rangle - |1\rangle}{\sqrt{2}}. \tag{292}$$

If Alice receives the constraint $v_{k_1} \dots v_{k_m} = (-1)^b$, she measures according to the binary observables $L_{g_{k_1}}, \dots, L_{g_{k_m}}$ (which can be measured separately, since $g_{k_1}, \dots, g_{k_m}$ commute with each other). Since $g_{k_1} \cdots g_{k_m} = J^b$, the product of the outcomes is equal to the outcome of measuring $|\psi\rangle$ with respect to the binary observable L_{J^b}. Since $L_{J^b}|\psi\rangle = (-1)^b|\psi\rangle$, this outcome is $+1$ if $b = 0$ and -1 if $b = 1$. This implies that Alice's output values satisfy the constraint.

If Bob receives variable v_i, he measures with respect to observable R_{g_i}. The outcome value will be the same as Alice's measurement with respect to the observable L_{g_i}, since

$$R_{g_i} L_{g_i} |\psi\rangle = \frac{|g_i J g_i\rangle - |g_i g_i\rangle}{\sqrt{2}} \tag{293}$$

$$= \frac{|J\rangle - |1\rangle}{\sqrt{2}} \tag{294}$$

$$= |\psi\rangle. \tag{295}$$

This implies that Alice and Bob produce consistent values for the variable v_i. □

15.4 Commuting operator strategies vs. C*-algebraic strategies

In section 14.6, we defined compound registers in the C*-algebra framework in terms of the min-tensor product of C*-algebras. Here, we show that any commuting-operator strategy can be converted into a strategy with identical performance in the C*-algebra framework using the *max-tensor product*.

First, we define the max-tensor product of two C*-algebras, $\mathcal{A}$ and $\mathcal{B}$, as the completion of the $*$-algebra $\mathcal{A} \otimes \mathcal{B}$ with respect to the *max-norm*, which is defined as follows.

Definition 15.12. *For C*-algebras $\mathcal{A}$ and $\mathcal{B}$, define the* max-norm *on $\mathcal{A} \otimes \mathcal{B}$ as, for any $x \in \mathcal{A} \otimes \mathcal{B}$,*

$$\|x\|_{\max} = \sup\big\{ \|\pi(x)\| : \pi : \mathcal{A} \otimes \mathcal{B} \to B(\mathcal{H}) \big\}, \tag{296}$$

where the supremum is over all Hilbert spaces $\mathcal{H}$ and $*$-homomorphisms $\pi : \mathcal{A} \otimes \mathcal{B} \to B(\mathcal{H})$.

Next we define what it means for two strategies to be *equivalent*.

Definition 15.13. *Two strategies for nonlocal games with input sets S and T, and output sets A and B are* equivalent *if, for all $(s,t) \in S \times T$ and $(a,b) \in A \times B$, the probability of output (a,b) on inputs (s,t) are the same.*

Clearly, if two strategies for a nonlocal game are equivalent then their success probability is the same.

Now, we can formally state the result as follows.

Theorem 15.14. *Let* $\big\{\mathcal{H}, |\psi\rangle, (\mathbf{A}_s^a)_{s\in S, a\in A}, (\mathbf{B}_t^b)_{t\in T, b\in B}\big\}$ *be a commuting operator strategy. Then there exist C*-algebras* $\mathcal{A}$*,* $\mathcal{B}$ *and an equivalent C*-algebraic strategy on registers with C*-algebras* $\mathcal{A}$*,* $\mathcal{B}$*, and compound register with C*-algebra* $\mathcal{A} \otimes_{\max} \mathcal{B}$*.*

Prior to proving Theorem 15.14, we define a universal C*-algebra associated with the strategies any nonlocal game with (finite) input sets S and T, and (finite) output sets A and B.

A *universal C*-algebra* is specified by a *presentation*, which is a set of variables that generate it and a sequence of algebraic relationships among the generators. There is a natural $*$-algebra associated with any presentation, which consists of all algebraic expressions in the generators modulo the equivalence relation defined by the algebraic relationships. The $*$-algebra becomes a C*-algebra by defining a max-norm on it as

$$\|x\|_{\max} = \sup\big\{\|\pi(x)\| : \pi : \mathcal{A} \to B(\mathcal{H})\big\}, \tag{297}$$

where the supremum is over all Hilbert spaces $\mathcal{H}$ and all $*$-homomorphisms $\pi : \mathcal{A} \to B(\mathcal{H})$.

It should be noted that, for some presentations (e.g., $\langle a : a = a^*\rangle$) the max norm is not well defined (in the sense that the supremum is ∞). However, the presentations that we are interested in arise from a commuting operator strategy and are of the form[15]

$$\mathcal{A} = \Big\langle A_s^a \text{ (for } s \in S,\, a \in A) \,\Big|\, \text{where } \forall s, a,\, (A_s^a)^* = A_s^a,\, (A_s^a)^2 = A_s^a \text{ and, } \forall s,\, \textstyle\sum_a A_s^a = 1 \Big\rangle \tag{298}$$

$$\mathcal{B} = \Big\langle B_t^b \text{ (for } t \in T,\, b \in B) \,\Big|\, \text{where } \forall t, b,\, (B_t^b)^* = B_t^b,\, (B_t^b)^2 = B_t^b \text{ and, } \forall t,\, \textstyle\sum_b B_t^b = 1 \Big\rangle. \tag{299}$$

For these presentations it can de deduced that the norm of each generator is 1, and from this it follows that the max-norm is well-defined.

It turns out that $\mathcal{A} \otimes_{\max} \mathcal{B}$ is equivalent to the presentation that is the union of the above generators and relationships with these additional relationships: $\forall s, t, a, b,\ A_s^a B_t^b A_s^a B_t^b = 1$ (i.e., each A_s^a and B_t^b commute). The reason why it is appropriate to include these commutations is that, in $\mathcal{A} \otimes \mathcal{B}$, for each $A \in \mathcal{A}$ and $B \in \mathcal{B}$, $A \otimes I$ and $I \otimes B$ commute.

Proof of Theorem 15.14. Let $\big\{\mathcal{H}, |\psi\rangle, (\mathbf{A}_s^a)_{s\in S, a\in A}, (\mathbf{B}_t^b)_{t\in T, b\in B}\big\}$ be a given commuting-operator strategy. Without loss of generality (by the Stinespring dilation), it can be assumed that, for each $s \in S$ and $t \in T$, $(\mathbf{A}_s^a)_{a\in A}$ and $(\mathbf{B}_t^b)_{b\in B}$ are projective measurements.

Define $\mathcal{A}$ and $\mathcal{B}$ as above, in Eqns. (298)and (299). Then the unique bounded linear mapping

$$\pi : \mathcal{A} \otimes_{\max} \mathcal{B} \to B(\mathcal{H}) \tag{300}$$

that satisfies, for all $s \in S$, $t \in T$, $a \in A$, $b \in B$,

$$\pi(A_s^a \otimes B_t^b) = \mathbf{A}_s^a \mathbf{B}_t^b \tag{301}$$

is a *representation* of $\mathcal{A} \otimes_{\max} \mathcal{B}$.

[15] These presentations are for projective measurements which may appear more restrictive that POVM measurements; however, they are not more restrictive, by the Stinespring dilation.

Define the state s on $\mathcal{A} \otimes_{\max} \mathcal{B}$ as, for all $x \in \mathcal{A} \otimes_{\max} \mathcal{B}$,

$$s(x) = \langle \psi | \pi(x) | \psi \rangle. \tag{302}$$

This is a valid state because of the form of Eq. (302) (in particular, if $x \geq 0$ then $\pi(x) \geq 0$).

Consider the C*-algebraic strategy based on operators $A_s^a \in \mathcal{A}$ (for all $s \in S$, $a \in A$), $B_s^b \in \mathcal{B}$ (for all $t \in T$, $b \in B$), and state $s : \mathcal{A} \otimes_{\max} \mathcal{B} \to \mathbb{C}$ as defined in Eq. (302). This strategy is equivalent to the original commuting operator strategy because, for all $s \in S, t \in T, a \in A, b \in B$,

$$s(A_s^a \otimes B_t^b) = s(\pi(A_s^a \otimes B_t^b)) = \langle \psi | \mathbf{A}_s^a \mathbf{B}_t^b | \psi \rangle. \tag{303}$$

□

Acknowledgments

I gratefully acknowledge guidance from Vern Paulsen and William Slofstra with the material on C*-algebras in sections 14 and 15. In particular, William showed me how to prove Theorem 15.14. I gratefully acknowledge the feedback from the students in the course, Jacob Barnett, Ian Davis, Adina Goldberg, Junan Lin, Junqiao Lin, and Abel Molina, who pointed out errors and made helpful suggestions. The remaining errors and lack or clarity are my responsibility.

References

[1] A. Arkhipov, *Extending and characterizing quantum magic games*, 2012, Manuscript available at arXiv:1209.3819.

[2] W. Arveson, *An invitation to C*-algebras*, 1976.

[3] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Proposed experiment to test local hidden-variable theories*, Physical Review Letters **23** (1969), no. 15, 880–884.

[4] R. Cleve, P. Høyer, B. Toner, and J. Watrous, *Consequences and limits of nonlocal strategies*, Proceedings of the 19th Annual IEEE Conference on Computational Complexity, 2004, pp. 236–249.

[5] R. Cleve, L. Liu, and V. Paulsen, *Perfect embezzlement of entanglement*, Journal of Mathematical Physics **58** (2017), 012204.

[6] R. Cleve, L. Liu, and W. Slofstra, *Perfect commuting-operator strategies for linear system games*, Journal of Mathematical Physics **58** (2017), 012202.

[7] R. Cleve and R. Mittal, *Characterization of binary constraint system games*, Proceedings 41st International Colloquium on Automata, Languages, and Programming (ICALP 2014), 2014, pp. 320–331.

[8] K. R. Davidson, *C*-algebras by example*, 1983.

[9] I. M. Gelfand and M. A. Naimark, *On the embedding of normed rings into the ring of operators in Hilbert space*, Matematiceskij sbornik **12** (1943), 197–213.

[10] Z. Ji, D. Leung, and T. Vidick, *A three-player coherent state embezzlement game*, 2018, Manuscript available at arXiv:1802.04926.

[11] R. V. Kadison and J. R. Ringrose, *Fundamentals of the theory of operator algebras, volume I*, 1983.

[12] M. Keyl, D. Schlingemann, and R. Werner, *Infinitely entangled states*, Quantum Information and Computation **3** (2003), no. 4, 281–306.

[13] D. Leung, B. Toner, and J. Watrous, *Coherent state exchange in multi-prover quantum interactive proof systems*, Chicago Journal of Theoretical Computer Science **2013** (2013), article 11.

[14] M. McKague, T. H. Yang, and V. Scarani, *Robust self-testing of the singlet*, Journal of Physics A: Mathematical and Theoretical **45** (2012), no. 45, 455304.

[15] O. Regev and T. Vidick, *Quantum XOR games*, Proceedings of IEEE Conference on Computational Complexity (CCC 2013), 2013, pp. 144–155.

[16] I. E. Segal, *Irreducible representations of operator algebras*, Bulletin of the American Mathematical Society **53** (1947), 73–88.

[17] W. van Dam and P. Hayden, *Universal entanglement transformations without communication*, Physical Review A **67** (2003), no. 6, 060302.

[18] T. Vidick, *UCSD summer school notes quantum multiplayer games, testing and rigidity*, 2018, Manuscript available at http://users.cms.caltech.edu/∼vidick/ucsd_games.pdf.

[19] D. Voiculescu, *Asymptotically commuting finite rank unitary operators without commuting approximants*, Acta Universitatis Szegediensis **45** (1983), 429–431.