Quantum Information and Computation, Vol. 16, No. 13&14 (2016) 1191–1211

THE STRUCTURE OF NEARLY-OPTIMAL QUANTUM STRATEGIES FOR THE CHSH(n) XOR GAMES

DIMITER OSTREV [a]
Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue Cambridge, Massachusetts 02139-4307, USA

Received June 4, 2015
Revised July 25, 2016

We consider the infinite family of non-local games CHSH(n). We consider nearly-optimal strategies for CHSH(n). We introduce a notion of approximate homomorphism for strategies and show that any nearly-optimal strategy for CHSH(n) is approximately homomorphic to the canonical optimal CHSH(n) strategy. This demonstrates that any nearly-optimal CHSH(n) strategy must approximately contain the algebraic structure of the canonical optimal strategy.

Keywords: non-local XOR games, anti-commuting observables, approximate homomorphism
Communicated by: R Cleve & R de Wolf

1 Introduction

Non-local games are a framework used to study the correlations that result from measuring two parts of an entangled quantum state using two spatially separated devices, each capable of performing one of several possible measurements.

In a non-local game two players, traditionally called Alice and Bob, play cooperatively but are separated in space and unable to communicate with each other. A third party, called a Referee or sometimes a Verifier, runs the game and decides whether Alice and Bob win or lose. The referee exchanges messages with Alice and Bob and decides whether they win or lose based on the transcript of the interaction.

To win, Alice and Bob must find a way to coordinate their actions without communicating. They may do so using classical or quantum resources. In a randomized classical strategy, Alice and Bob use a shared random string. In a quantum strategy, Alice and Bob use measurements of two parts of a shared entangled state.

For certain non-local games, a quantum strategy can achieve a higher probability of winning than any randomized classical strategy [1, 2]. This is interesting both from the point of view of foundations of physics, and from the point of view of applications. From the point of view of foundations of physics, the advantage of quantum strategies over classical ones has been central in the discussion about local realism [3]. From the point of view of applications, there have been many proposals for using quantum entanglement as a resource in information processing tasks, such as performing distributed computation with a lower communication

[a]Email address: ostrev@mit.edu

cost [4], teleportation of quantum states [5], a full scale computation by teleportation scheme [6], and quantum cryptography [7].

A recent application of quantum entanglement is in device-independent protocols. In these protocols, one or more parties attempt to perform an information processing task by interacting via classical inputs and outputs with quantum devices that cannot be trusted to perform according to specification. The devices may not be trusted for fear of malicious intent, as in quantum cryptography, or the manufacturing process used to make the devices may be unreliable and prone to errors.

Several examples of device independent protocols are already known. There are protocols for quantum key distribution [8, 9], [10], [2, pp. 34-35]. There are protocols for randomness generation with untrusted devices [2, p.33]. There is also a protocol in which a classical verifier commands two untrusted quantum provers to perform a full-scale quantum computation [11].

All of these protocols rely on mathematical results called self-testing. References [11, 12, 13] are three examples of self-testing results. These results show that nearly-optimal quantum strategies for certain non-local games must be close (in an appropriate sense) to the ideal quantum strategy for the game.

In this paper, we present a new self-testing result. We study the infinite family of non-local XOR games CHSH(n), $n \in \mathbb{N}, n \geq 2$ introduced in [14]. We show that nearly-optimal quantum strategies for the CHSH(n) games must approximately contain the same algebraic structure as the optimal quantum strategies. We introduce tools and concepts that allow us to understand and perform the comparison of algebraic structure; these tools are the approximate homomorphism of strategies, the concept of core space of a strategy, and a modified version of the group averaging technique. All of these we obtain by adapting tools and concepts from algebra and representation theory to the context of non-local games.

The rest of this paper is organized as follows: in Section 2, we present notation, concepts and known facts that are used later on. In Section 3, we give the precise statement of the result proved in this paper, as well as the intuition and motivation for the result. Section 4 contains the proof of the main result. In Section 5 we discuss open problems and possible future work.

2 Preliminaries

2.1 Non-local XOR games and their quantum strategies

Formally, a non-local XOR game consists of two finite sets S and T, a probability distribution π on $S \times T$, and a function $V : S \times T \to \{-1, 1\}$. The game proceeds as follows: first, the referee selects a pair $(s, t) \in S \times T$ according to the probability distribution π. Then, the referee sends s as a question to Alice and t as a question to Bob. Then, Alice replies to the referee with $a \in \{-1, 1\}$ and Bob replies to the referee with $b \in \{-1, 1\}$. Finally, the referee looks at $V(s, t)ab$. If $V(s, t)ab = 1$, then Alice and Bob win, and if $V(s, t)ab = -1$ then Alice and Bob lose. Notice that $V(s, t) = 1$ means that Alice and Bob must give matching answers to win and $V(s, t) = -1$ means Alice and Bob must give opposite answers to win.[b]

It is convenient to summarize all the information for an XOR game into a $|S| \times |T|$ matrix

[b]The name "XOR game" is related to the following: if we write $a = (-1)^{a'}$, $b = (-1)^{b'}$ for $a', b' \in \{0, 1\}$, then $V(s, t)ab = V(s, t)(-1)^{a' \oplus b'}$ so that whether Alice and Bob win or lose depends only on the XOR of the bits a' and b'. XOR games are a sub-class of non-local games.

G such that $G_{st} = \pi(s,t)V(s,t)$. This gives a bijective correspondence between XOR games and matrices G normalized so that $\sum_{st} |G_{st}| = 1$. From now on, we will identify non-local XOR games with their associated matrices.

A quantum strategy $\mathcal{S}$ for an XOR game consists of a state space $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, a state $|\psi\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, and ± 1 observables $\{A_s : s \in S\}$ on $\mathbb{C}^{d_A}$ and $\{B_t : t \in T\}$ on $\mathbb{C}^{d_B}$. The interpretation of this strategy is the following: Alice and Bob share a bipartite quantum system with state space $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. Prior to the beginning of the game, the system has been prepared in state $|\psi\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. On receiving question s, Alice measures observable A_s and uses the outcome, 1 or -1, as her answer to the referee. Similarly, on receiving question t, Bob measures observable B_t and uses the outcome, 1 or -1, as his answer to the referee.

We would like to have a way to evaluate how well a given strategy $\mathcal{S}$ does for a given XOR game G. We do so using the success bias $\beta(G, \mathcal{S})$ defined by:

$$\beta(G, \mathcal{S}) = \sum_{s \in S} \sum_{t \in T} G_{st} \langle \psi | A_s \otimes B_t | \psi \rangle$$

The success bias is linearly related to the probability $\omega(G, \mathcal{S})$ of winning G using strategy $\mathcal{S}$:

$$\beta(G, \mathcal{S}) = 2\omega(G, \mathcal{S}) - 1$$

We define the quantum success bias $\beta(G)$ for an XOR game G to be the supremum of the success bias over all quantum strategies:

$$\beta(G) = \sup_{\mathcal{S}} \beta(G, \mathcal{S}) \tag{1}$$

We define an optimal strategy for the XOR game G to be a strategy $\mathcal{S}$ such that

$$\beta(G, \mathcal{S}) = \beta(G)$$

and we define an ϵ-optimal strategy to be a strategy $\mathcal{S}$ such that

$$(1 - \epsilon)\beta(G) \leq \beta(G, \mathcal{S}) \leq \beta(G)$$

2.2 Non-local XOR games and semi-definite programs

Let G be a $n \times m$ XOR game matrix. To the maximization problem (1), we associate a semi-definite program:[c]

$$\sup_{Z \succeq 0,\, Tr(E_{ii}Z)=1,\, i=1,\dots(n+m)} Tr(G_{sym}Z) \tag{2}$$

Here, E_{ii} is the $(n+m) \times (n+m)$ matrix with 1 in the i-th diagonal entry and 0 everywhere else, and G_{sym} is the $(n+m) \times (n+m)$ matrix with block form

$$G_{sym} = \frac{1}{2} \begin{bmatrix} 0 & G \\ G^T & 0 \end{bmatrix}$$

The two maximization problems (1) and (2) are related as follows: for each feasible solution of one of them, there is a feasible solution of the other that achieves the same objective value. This follows from the results in [17] and [18].

[c]For an introduction to semi-definite programming, see reference [15], or the lecture notes [16].

Having established the relation between the optimization problem (1) and the semi-definite program (2), we now turn attention to the dual semidefinite program. The dual to (2) is:

$$\inf_{\sum_{i=1}^{m+n} y_i E_{ii} \succeq G_{sym}} \sum_{i=1}^{m+n} y_i \tag{3}$$

Both the primal and the dual semi-definite programs have strictly feasible solutions; therefore, the primal supremum is attained, the dual infimum is attained, and the two are equal.

Combining all our observations, we get that the three optimization problems (1), (2), and (3) have equal optimal objective values, and that all three optimal objective values are attained.

2.3 The CHSH(n) XOR games

In this paper, we study the infinite family of XOR games CHSH(n) , $n \in \mathbb{N}$, $n \geq 2$ introduced in [14]. This is a natural case to study due to the regular structure of the game and its optimal strategies.

For the CHSH(n) game, the set S of possible questions for Alice is $\{1, \ldots, n\}$ and the set T of possible questions for Bob is the set of ordered pairs $\{ij \,:\, i, j \in \{1, \ldots, n\},\, i \neq j\}$. The referee selects a pair i, j uniformly at random among all $\binom{n}{2}$ pairs such that $1 \leq i < j \leq n$. The referee then selects either i or j as question for Alice, and either ij or ji as question for Bob; the four possibilities are equally likely. In order to win, Alice and Bob must give matching answers on questions (i, ij), (i, ji) and (j, ij), and give opposite answers on questions (j, ji).

Note that the first element of the family, CHSH(2), is the usual CHSH game, based on reference [19]. Thus, the family CHSH(n) is a generalization of the CHSH game.

It is known [14] that the quantum success bias for all the CHSH(n) games is $1/\sqrt{2}$ and that the optimal quantum strategies for CHSH(n) are related to representations of the Clifford algebra with n anti-commuting generators.

2.4 The canonical optimal quantum strategy for CHSH(n)

Here we introduce the canonical optimal strategy for CHSH(n). The canonical strategy is defined so that Alice's observables generate an algebra that is isomorphic to the Clifford algebra with n anti-commuting generators. We will state some known facts about the Clifford algebra and then return to constructing the canonical optimal strategies for CHSH(n).

By a Clifford algebra with n anti-commuting generators we mean the algebra of linear combinations of products of generators $x_1, \ldots, x_n$ satisfying the relations $x_i x_j + x_j x_i = 2\delta_{ij}\mathbf{1}$. We will denote this algebra by $\mathbf{Cl_n}$.

It is known [20, Lemmas 20.9, 20.16] that $\mathbf{Cl_{2k}}$ is isomorphic to the algebra of $2^k \times 2^k$ matrices, and $\mathbf{Cl_{2k+1}}$ is isomorphic to the direct sum of two copies of the algebra of $2^k \times 2^k$ matrices. Consequently [21, Thm 2.6], $\mathbf{Cl_{2k}}$ has a unique (up to equivalence) irreducible representation on $\mathbb{C}^{2^k}$, $\mathbf{Cl_{2k+1}}$ has two irreducible representations on $\mathbb{C}^{2^k}$, and in both the even and the odd case, any finite dimensional representation is equivalent to a direct sum of irreducible representations.

The irreducible representations of $\mathbf{Cl_{2k}}$ and $\mathbf{Cl_{2k+1}}$ can be constructed explicitly. Consider the following $2k+1$ operators on $\mathbb{C}^{2^k} \cong \underbrace{\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{k\,terms}$:

$$\begin{aligned} \sigma_{k,2l-1} &= \underbrace{\sigma_y \otimes \cdots \otimes \sigma_y}_{l-1} \otimes \sigma_x \otimes \underbrace{I \otimes I \otimes I \otimes \cdots \otimes I \otimes I}_{k-l} \quad \text{for } l = 1, \ldots k \\ \sigma_{k,2l} &= \underbrace{\sigma_y \otimes \cdots \otimes \sigma_y}_{l-1} \otimes \sigma_z \otimes \underbrace{I \otimes I \otimes I \otimes \cdots \otimes I \otimes I}_{k-l} \quad \text{for } l = 1, \ldots k \\ \sigma_{k,2k+1} &= \sigma_y \otimes \sigma_y \otimes \sigma_y \otimes \sigma_y \otimes \cdots \otimes \sigma_y \otimes \sigma_y \end{aligned} \tag{4}$$

These operators are self-adjoint, unitary, and anti-commute. Mapping the generators of $\mathbf{Cl_{2k}}$ to $\sigma_{k,1}, \ldots \sigma_{k,2k}$ gives the irreducible representation of $\mathbf{Cl_{2k}}$. Mapping the generators of $\mathbf{Cl_{2k+1}}$ to either $\sigma_{k,1}, \ldots \sigma_{k,2k}, \sigma_{k,2k+1}$ or $\sigma_{k,1}, \ldots \sigma_{k,2k}, -\sigma_{k,2k+1}$ [d] gives the two irreducible representations of $\mathbf{Cl_{2k+1}}$.

Now, we can return to constructing the canonical optimal quantum strategy for the CHSH(n) game. The state space is $\mathbb{C}^{2^{\lceil n/2 \rceil}} \otimes \mathbb{C}^{2^{\lceil n/2 \rceil}}$. The state is the maximally entangled state

$$|\tilde{\psi}\rangle = \frac{1}{\sqrt{2^{\lceil n/2 \rceil}}} \sum_{i=1}^{2^{\lceil n/2 \rceil}} |i\rangle \otimes |i\rangle$$

When $n = 2k$, Alice's observables are defined by

$$\tilde{A}_i = \sigma_{k,i}, \quad i = 1, \ldots 2k$$

When $n = 2k+1$, Alice's observables are defined by

$$\tilde{A}_i = \begin{bmatrix} \sigma_{k,i} & 0 \\ 0 & \sigma_{k,i} \end{bmatrix}, \quad i = 1, \ldots 2k, \quad \tilde{A}_{2k+1} = \begin{bmatrix} \sigma_{k,2k+1} & 0 \\ 0 & -\sigma_{k,2k+1} \end{bmatrix}$$

In all cases, Bob's observables are defined by

$$\tilde{B}_{jl} = \frac{1}{\sqrt{2}}(\tilde{A}_j^T + \tilde{A}_l^T), \quad \tilde{B}_{lj} = \frac{1}{\sqrt{2}}(\tilde{A}_j^T - \tilde{A}_l^T), \quad 1 \le j < l \le n$$

In the rest of the paper, we will always use tildes to denote the observables and state of the canonical strategy.

3 The main result

We start with a CHSH(n) strategy $A_i, B_{jk}, |\psi\rangle$ on $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ that is ϵ-optimal. We wish to say that this ϵ-optimal strategy must approximately have certain structure. To do that, we compare it to the canonical optimal strategy $\tilde{A}_i, \tilde{B}_{jk}, |\tilde{\psi}\rangle$ on $\mathbb{C}^{2^{\lceil n/2 \rceil}} \otimes \mathbb{C}^{2^{\lceil n/2 \rceil}}$. Formally, we prove the following:

Theorem 1 *Let $A_i, B_{jk}, |\psi\rangle$ be an ϵ-optimal CHSH(n) strategy on $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. Let $\tilde{A}_i, \tilde{B}_{jk}, |\tilde{\psi}\rangle$ be the canonical optimal strategy on $\mathbb{C}^{2^{\lceil n/2 \rceil}} \otimes \mathbb{C}^{2^{\lceil n/2 \rceil}}$. Consider the linear operator*

$$T : \mathbb{C}^{2^{\lceil n/2 \rceil}} \otimes \mathbb{C}^{2^{\lceil n/2 \rceil}} \longrightarrow \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$$

[d] The two options are not equivalent because the product of the observables in the first collection is $(-\mathrm{i})^k I$ and the product in the second collection is $-(-\mathrm{i})^k I$.

given by

$$T = \frac{1}{\sqrt{2^n}} \sum_{(j_1 \dots j_n) \in \{0,1\}^n} A_1^{j_1} \dots A_n^{j_n} \otimes I |\psi\rangle\langle\tilde{\psi}| \left(\tilde{A}_1^{j_1} \dots \tilde{A}_n^{j_n} \otimes I \right)^\dagger$$

Then, T is non-zero and has the properties

$$\begin{aligned} \forall i \quad & \|(A_i \otimes I)T - T(\tilde{A}_i \otimes I)\| < 12n^2\sqrt{\epsilon}\|T\| \\ \forall j \neq k \quad & \|(I \otimes B_{jk})T - T(I \otimes \tilde{B}_{jk})\| < 17n^2\sqrt{\epsilon}\|T\| \end{aligned} \tag{5}$$

A remark about notation: when we apply $\|\cdot\|$ to a matrix we mean the Frobenius norm $\|M\| = \sqrt{Tr(MM^\dagger)}$ unless explicitly stated otherwise.

Next, we discuss the motivation and intuition behind Theorem 1. We look at it from three different points of view: the point of view of the concept of homomorphism in algebra, the point of view of extending the concept of core space of an optimal strategy (the core space will be defined below), and the point of view of proving approximate anti-commutation for the observables A_i.

3.1 Approximate homomorphism

When we talk of a homomorphism, we have two sets with certain operations on each, and the homomorphism is a map from one set to the other that preserves all the operations. In the context of Theorem 1, the two sets are $\mathbb{C}^{2^{\lceil n/2 \rceil}} \otimes \mathbb{C}^{2^{\lceil n/2 \rceil}}$ and $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. The operations on $\mathbb{C}^{2^{\lceil n/2 \rceil}} \otimes \mathbb{C}^{2^{\lceil n/2 \rceil}}$ are the action of the operators $\tilde{A}_i \otimes I$, $I \otimes \tilde{B}_{jk}$. The operations on $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ are the action of the operators $A_i \otimes I$, $I \otimes B_{jk}$. The operator T that we construct in Theorem 1 satisfies properties (5), so it approximately maps the action of the canonical strategy operators $\tilde{A}_i \otimes I$, $I \otimes \tilde{B}_{jk}$ to the action of the ϵ-optimal strategy operators $A_i \otimes I$, $I \otimes B_{jk}$.

3.2 The core space of a strategy

For the canonical optimal CHSH(n) strategy $\tilde{A}_i, \tilde{B}_{jk}, |\tilde{\psi}\rangle$, we define the core space to be

$$span \left\{ \tilde{A}_1^{j_1} \dots \tilde{A}_n^{j_n} \otimes I |\tilde{\psi}\rangle \ : \ (j_1 \dots j_n) \in \{0,1\}^n \right\}$$

and denote it by $\mathcal{CS}(\tilde{A}_i, \tilde{B}_{jk}, |\tilde{\psi}\rangle)$. This is the smallest vector space that contains $|\tilde{\psi}\rangle$ and is invariant under the canonical strategy operators $\tilde{A}_i \otimes I$, $I \otimes \tilde{B}_{jk}$. [e]

To get a feel for the core space, consider two small examples: the core space of the canonical optimal CHSH(2) strategy is $\mathbb{C}^2 \otimes \mathbb{C}^2$; the core space of the canonical optimal CHSH(3) strategy is an eight dimensional subspace of $\mathbb{C}^4 \otimes \mathbb{C}^4$. It is interesting to note that the core space of the canonical optimal CHSH(3) strategy is not a tensor product space. If this core space were of the form $V \otimes W$, then V and W would have to have the same dimension by symmetry. This is a contradiction, because 8 is not a perfect square.

Theorem 1 allows us to extend the concept of core space to nearly-optimal CHSH(n) strategies. First, we observe that besides the standard definition of invariant space, there is the following equivalent definition: a vector space V is invariant under a linear operator C if there exist linear operators S and D such that $Im(S) = V$ and $CS = SD$. Next, we see

[e] A vector space V is invariant under an operator C if $|v\rangle \in V \Rightarrow (C|v\rangle) \in V$.

that properties (5) are approximate versions of the relation $CS - SD = 0$. This allows us to regard

$$Im(T) = span\left\{A_1^{j_1}\dots A_n^{j_n}\otimes I|\psi\rangle\ :\ (j_1\dots j_n)\in\{0,1\}^n\right\}$$

as approximately invariant under the operators $A_i\otimes I$, $I\otimes B_{jk}$. By analogy with the canonical optimal strategies, we call this space the core space of the strategy $A_i, B_{jk}, |\psi\rangle$ and denote it by $\mathcal{CS}(A_i, B_{jk}, |\psi\rangle)$.

3.3 Approximate anti-commutation for the observables A_i of an ϵ-optimal strategy

We know that Alice's observables in the canonical optimal strategy anti-commute. Our intuition tells us that Alice's observables in an ϵ-optimal strategy must approximately anti-commute. Theorem 1 allows us to make this intuition rigorous.

When we first start thinking about approximate anti-commutation for the observables A_i, we encounter a conceptual problem: we would like a way to measure the size of the anti-commutators $A_iA_j + A_jA_i$ and claim that they are small, but none of the usual matrix norms is suitable. One can construct examples of ϵ-optimal CHSH(n) strategies such that $\|A_iA_j + A_jA_i\|$ remains bounded away from zero as ϵ goes to zero. This is because the restriction that the strategy is ϵ-optimal does not constrain the action of A_i outside the support of the strategy state $|\psi\rangle$.

One solution to this conceptual problem is to measure $A_iA_j + A_jA_i$ in an appropriately chosen seminorm instead of in one of the usual matrix norms. We introduce the concept of a seminorm with respect to a linear operator as follows:

Definition 1 *Let V, U, W be (finite dimensional) Hilbert spaces. Let $L : V \to U$ and $M : U \to W$ be linear operators. We define the seminorm of M with respect to L by*

$$\|M\|_L = \frac{\|ML\|}{\|L\|}$$

It can be checked that $\|\cdot\|_L$ satisfies all the defining properties of a seminorm.

A special case of a seminorm with respect to a linear operator has already been used implicitly in the self-testing literature (for example in [11, 12]). The state $|\psi\rangle \in \mathbb{C}^{d_A}\otimes\mathbb{C}^{d_B}$ can be thought of as a linear operator from $\mathbb{C}$ to $\mathbb{C}^{d_A}\otimes\mathbb{C}^{d_B}$. The seminorm with respect to $|\psi\rangle$ is

$$\|M\|_\psi = \frac{\|M|\psi\rangle\|}{\||\psi\rangle\|} = \|M|\psi\rangle\|$$

In this paper we prove approximate anti-commutation for the observables A_i of an ϵ-optimal strategy using both the seminorm with respect to $|\psi\rangle$ and the seminorm with respect to T. The statement using the seminorm with respect to $|\psi\rangle$ is one step of the proof of Theorem 1. In Lemma 5 below we will show that

$$\sum_{1\le i<j\le n}\|(A_iA_j + A_jA_i)\otimes I\|_\psi^2 \le 4(1+\sqrt{2})^2 n(n-1)\epsilon$$

The statement using the seminorm with respect to T is a corollary of Theorem 1. The triangle inequality and four applications of inequalites (5) give

$$\forall i\ne j\quad \|(A_iA_j + A_jA_i)\otimes I\|_T < 48n^2\sqrt{\epsilon}$$

4 Approximate homomorphism construction for CHSH(n) nearly-optimal strategies

Having given the motivation and intuition for Theorem 1, we now proceed with the proof. The argument has eight steps contained in the following propositions:

Proposition 1 *The following three statements are equivalent:*

- $A_i, B_{jk}, |\psi\rangle$ *is an* ϵ*-optimal CHSH*(n) *strategy.*
- *The observables and state satisfy*

$$\sum_{1\le i<j\le n}\left(\left\|\frac{A_i+A_j}{\sqrt{2}}\otimes I|\psi\rangle - I\otimes B_{ij}|\psi\rangle\right\|^2 + \left\|\frac{A_i-A_j}{\sqrt{2}}\otimes I|\psi\rangle - I\otimes B_{ji}|\psi\rangle\right\|^2\right)\le 2n(n-1)\epsilon \quad (6)$$

- *The observables and state satisfy*

$$\sum_{1\le i<j\le n}\left(\left\|A_i\otimes I|\psi\rangle - I\otimes \frac{B_{ij}+B_{ji}}{\sqrt{2}}|\psi\rangle\right\|^2 + \left\|A_j\otimes I|\psi\rangle - I\otimes \frac{B_{ij}-B_{ji}}{\sqrt{2}}|\psi\rangle\right\|^2\right)\le 2n(n-1)\epsilon \quad (7)$$

Proposition 2 *The vectors*

$$\left\{\tilde{A}_1^{j_1}\dots\tilde{A}_n^{j_n}\otimes I|\tilde{\psi}\rangle\ :\ (j_1\dots j_n)\in\{0,1\}^n\right\}$$

coming from the canonical strategy are orthonormal.

Proposition 3 $\|T\|=1$*, and so also* $T\neq 0$.

Proposition 4 *The following identity holds:*

$$(A_i\otimes I)T - T(\tilde{A}_i\otimes I) = \frac{1}{\sqrt{2^n}}\sum_{(j_1\dots j_n)\in\{0,1\}^n}\Big(A_iA_1^{j_1}\dots A_n^{j_n}\otimes I|\psi\rangle - sign(i,j_1,\dots j_n)A_1^{j_1}\dots A_i^{j_i\oplus 1}\dots A_n^{j_n}\otimes I|\psi\rangle\Big)\langle\tilde{\psi}|\left(\tilde{A}_1^{j_1}\dots\tilde{A}_n^{j_n}\otimes I\right)^\dagger \quad (8)$$

Here $sign(i,j_1,\dots j_n)$ *denotes the sign resulting from changing the order in a product of anti-commuting observables and will be defined in detail later. The* $j_i\oplus 1$ *in the exponent of* A_i *denotes addition mod 2.*

Proposition 5 *The following identity holds:*

$$(I\otimes B_{kl})T - T(I\otimes\tilde{B}_{kl}) = \frac{1}{\sqrt{2^n}}\sum_{(j_1\dots j_n)\in\{0,1\}^n}\left(A_1^{j_1}\dots A_n^{j_n}\otimes B_{kl}|\psi\rangle - \frac{1}{\sqrt{2}}\Big(\pm sign(j_1,\dots j_n,k)A_1^{j_1}\dots A_k^{j_k\oplus 1}\dots A_n^{j_n}\otimes I|\psi\rangle + sign(j_1,\dots j_n,l)A_1^{j_1}\dots A_l^{j_l\oplus 1}\dots A_n^{j_n}\otimes I|\psi\rangle\Big)\right)\langle\tilde{\psi}|\left(\tilde{A}_1^{j_1}\dots\tilde{A}_n^{j_n}\otimes I\right)^\dagger \quad (9)$$

In the place where there is $\pm$, we take $+$ if $k < l$ and we take $-$ if $k > l$.

Proposition 6 *For all $i \in \{1, \dots n\}$, for all $(j_1 \dots j_n) \in \{0,1\}^n$,*

$$\left\| A_i A_1^{j_1} \dots A_n^{j_n} \otimes I|\psi\rangle - sign(i, j_1, \dots j_n) A_1^{j_1} \dots A_i^{j_i \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle \right\| < 12n^2\sqrt{\epsilon} \quad (10)$$

Proposition 7 *For all $k \neq l \in \{1, \dots n\}$, for all $(j_1 \dots j_n) \in \{0,1\}^n$,*

$$\Bigg\| A_1^{j_1} \dots A_n^{j_n} \otimes B_{kl}|\psi\rangle - \frac{1}{\sqrt{2}}\Big(\pm sign(j_1, \dots j_n, k) A_1^{j_1} \dots A_k^{j_k \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle$$
$$+ sign(j_1, \dots j_n, l) A_1^{j_1} \dots A_l^{j_l \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle \Big) \Bigg\| < 17n^2\sqrt{\epsilon} \quad (11)$$

Proposition 8 *The seven propositions above taken together imply Theorem 1.*

The eight subsections below contain the proofs of the eight propositions above.

4.1 Relations for CHSH(n) nearly-optimal strategies

In this section we prove Proposition 1.

We take the $n \times n(n-1)$ matrix G that summarizes the information for the CHSH(n) game. Let $|1\rangle, \dots |n\rangle$ be an orthonormal basis of $\mathbb{R}^n$, and let $|ij\rangle$, $i \neq j \in \{1, \dots n\}$ be an orthonormal basis of $\mathbb{R}^{n(n-1)}$. Then, we can write:

$$G = \frac{1}{4\binom{n}{2}} \sum_{1 \le i < j \le n} (|i\rangle\langle ij| + |j\rangle\langle ij| + |i\rangle\langle ji| - |j\rangle\langle ji|)$$

Next, we form the $n^2 \times n^2$ matrix G_{sym} which has block form:

$$G_{sym} = \frac{1}{2}\begin{bmatrix} 0 & G \\ G^T & 0 \end{bmatrix}$$

In this context, it is convenient to think of $\mathbb{R}^{n^2}$ as having an orthonormal basis formed by concatenating the basis $|1\rangle, \dots |n\rangle$ of $\mathbb{R}^n$ and the basis $|ij\rangle$, $i \neq j \in \{1, \dots n\}$ of $\mathbb{R}^{n(n-1)}$. So, we can write

$$G_{sym} = \frac{1}{8\binom{n}{2}} \sum_{1 \le i < j \le n} \Big(|i\rangle\langle ij| + |j\rangle\langle ij| + |i\rangle\langle ji| - |j\rangle\langle ji| + |ij\rangle\langle i| + |ij\rangle\langle j| + |ji\rangle\langle i| - |ji\rangle\langle j| \Big)$$

Next, we form the dual semi-definite program (3) corresponding to the CHSH(n) game. We know that the optimal value is $1/\sqrt{2}$; this follows from the result in reference [14] about the quantum success bias of the CHSH(n) game, and the discussion in Section 2.2.

Next, we claim that $y_1 = \dots = y_n = \frac{1}{2\sqrt{2}n}$, $y_{n+1} = \dots = y_{n^2} = \frac{1}{2\sqrt{2}n(n-1)}$ is a dual optimal solution. We can see that $\sum_{i=1}^{n^2} y_i = \frac{1}{\sqrt{2}}$, the dual optimum, so all that is left to prove is that $y_1, \dots y_{n^2}$ is dual feasible.

We define the following vectors for $1 \le i < j \le n$

$$u_{ij} = |i\rangle \qquad v_{ij} = \frac{|ij\rangle + |ji\rangle}{\sqrt{2}}$$
$$u_{ji} = |j\rangle \qquad v_{ji} = \frac{|ij\rangle - |ji\rangle}{\sqrt{2}}$$

and observe that the following decomposition holds:

$$\sum_{i=1}^{n^2} y_i E_{ii} - G_{sym} = \frac{1}{2\sqrt{2}n(n-1)} \sum_{i\neq j} (u_{ij} - v_{ij})(u_{ij} - v_{ij})^T \tag{12}$$

It follows that the matrix $\sum_{i=1}^{n^2} y_i E_{ii} - G_{sym}$ is positive semi-definite, and therefore, the given $y_1, \ldots y_{n^2}$ are a dual optimal solution as claimed.

Next, we need a lemma:

Lemma 1 *Let $A_1 \ldots A_n$, $B_1, \ldots B_m$, $|\psi\rangle$ be a quantum strategy for an XOR game given by a $n \times m$ matrix Γ. Let $\alpha_1, \ldots \alpha_r \in \mathbb{R}^n$, $\beta_1, \ldots \beta_r \in \mathbb{R}^m$ be vectors with the properties:*

$$\sum_{i=1}^{r} \alpha_i \alpha_i^T = \sum_{i=1}^{n} w_i E_{ii} \tag{13}$$

$$\sum_{i=1}^{r} \beta_i \beta_i^T = \sum_{i=1}^{m} w_{n+i} E_{ii} \tag{14}$$

$$\sum_{i=1}^{r} \alpha_i \beta_i^T = \frac{1}{2}\Gamma \tag{15}$$

Then, the following identity holds:

$$\sum_{k=1}^{r} \left\| \alpha_k \cdot \vec{A} \otimes I|\psi\rangle - I \otimes \beta_k \cdot \vec{B}|\psi\rangle \right\|^2 = \sum_{i=1}^{m+n} w_i - \sum_{i=1}^{n}\sum_{j=1}^{m} \Gamma_{ij} \langle\psi|A_i \otimes B_j|\psi\rangle \tag{16}$$

Here, $\alpha_k \cdot \vec{A}$ denotes a linear combination of the observables $A_1, \ldots A_n$ with coefficients taken from the vector α_k.

Proof. We open the squares on the left-hand side:

$$\sum_{k=1}^{r} \left\| \alpha_k \cdot \vec{A} \otimes I|\psi\rangle - I \otimes \beta_k \cdot \vec{B}|\psi\rangle \right\|^2$$

$$= \sum_{i=1}^{r} \langle\psi| \left(\alpha_i \cdot \vec{A}\right)^2 \otimes I|\psi\rangle + \sum_{i=1}^{r} \langle\psi|I \otimes \left(\beta_i \cdot \vec{B}\right)^2 |\psi\rangle - 2\sum_{i=1}^{r} \langle\psi| \left(\alpha_i \cdot \vec{A}\right) \otimes \left(\beta_i \cdot \vec{B}\right) |\psi\rangle$$

Now, from property (13) we obtain

$$\sum_{i=1}^{r} \left(\alpha_i \cdot \vec{A}\right)^2 = \sum_{i=1}^{n} w_i A_i^2 + \sum_{i\neq j} 0 A_i A_j = \left(\sum_{i=1}^{n} w_i\right) I$$

Similarly, from property (14) we obtain

$$\sum_{i=1}^{r} \left(\beta_i \cdot \vec{B}\right)^2 = \sum_{i=1}^{m} w_{n+i} B_i^2 + \sum_{i\neq j} 0 B_i B_j = \left(\sum_{i=1}^{m} w_{n+i}\right) I$$

Finally, from property (15) we obtain

$$2\sum_{i=1}^{r} \left(\alpha_i \cdot \vec{A}\right) \otimes \left(\beta_i \cdot \vec{B}\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \Gamma_{ij} A_i \otimes B_j$$

The identity (16) follows. $\square$.

Now, using decomposition (12) and Lemma 1 we conclude that the following two statements are equivalent:

- $A_i, B_{jk}, |\psi\rangle$ is an ϵ-optimal CHSH(n) strategy.
- The observables and state satisfy inequality (7)

Next, we define the following vectors for $1 \le i < j \le N$

$$u'_{ij} = \frac{|i\rangle + |j\rangle}{\sqrt{2}} \qquad v'_{ij} = |ij\rangle$$
$$u'_{ji} = \frac{|i\rangle - |j\rangle}{\sqrt{2}} \qquad v'_{ji} = |ji\rangle$$

and observe that the following decomposition holds:

$$\sum_{i=1}^{n^2} y_i E_{ii} - G_{sym} = \frac{1}{2\sqrt{2}n(n-1)} \sum_{i \neq j} \left(u'_{ij} - v'_{ij}\right)\left(u'_{ij} - v'_{ij}\right)^T$$

From this, using Lemma 1 again, we conclude that the following two statements are equivalent:

- $A_i, B_{jk}, |\psi\rangle$ is an ϵ-optimal CHSH(n) strategy.
- The observables and state satisfy inequality (6)

This completes the proof of Proposition 1.

4.2 Orthonormal vectors

In this section we prove Proposition 2.

Given $(k_1, \ldots k_n)$, $(l_1, \ldots l_n) \in \{0,1\}^n$, take $(j_1, \ldots j_n) = (k_1 \oplus l_1, \ldots k_n \oplus l_n)$ and use the anti-commutation relations for the $\tilde{A}_i$, $i = 1, \ldots n$ to get

$$\langle\tilde{\psi}| \left(\tilde{A}_1^{k_1} \ldots \tilde{A}_n^{k_n} \otimes I\right)^\dagger \tilde{A}_1^{l_1} \ldots \tilde{A}_n^{l_n} \otimes I|\tilde{\psi}\rangle = \langle\tilde{\psi}| \left(\pm\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n} \otimes I|\tilde{\psi}\rangle\right)$$

Therefore, it suffices to prove that $|\tilde{\psi}\rangle$ is orthogonal to $\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n} \otimes I|\tilde{\psi}\rangle$ for each nonzero $(j_1 \ldots j_n) \in \{0,1\}^n$. We need two lemmas.

Lemma 2 *If $|\phi\rangle$ is a maximally entangled state in $\mathbb{C}^d \otimes \mathbb{C}^d$ and M is a $d \times d$ matrix, then $\langle\phi|M \otimes I|\phi\rangle = Tr(M)/d$.*

Proof. $d\langle\phi|M \otimes I|\phi\rangle = \left(\sum_i \langle ii|\right)\left(\sum_{jk} M_{jk}|j\rangle\langle k| \otimes I\right)\left(\sum_l |ll\rangle\right) = \sum_i M_{ii} = Tr(M)$ $\square$.

Lemma 3 *$Tr(\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n}) = 0$ for each nonzero $(j_1 \ldots j_n) \in \{0,1\}^n$.*

Proof. There are two cases: one case is if n is odd and $(j_1, \ldots j_n) = (1, \ldots 1)$ and the second case is all other situations.

Consider the first case. For n odd, we have

$$\prod_{i=1}^{n} \tilde{A}_i = (-\mathrm{i})^{\lfloor n/2 \rfloor} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$$

and this has zero trace.

Next, consider the second case. We want to show that

$$\exists i \text{ such that } \tilde{A}_i \tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n} \tilde{A}_i = -\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n} \tag{17}$$

If there are an even number of terms in the product $\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n}$, choose $\tilde{A}_i$ to be one of the observables that appears in the product. If there are an odd number of terms, choose $\tilde{A}_i$ to be one of the observables that does not appear in the product. This proves relation (17).

From relation (17) we get that $Tr(\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n}) = 0$ in the second case as well. This completes the proof of Lemma 3. $\Box$.

Now, combine Lemma 2 and Lemma 3 to get $\langle\tilde{\psi}| \left(\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n} \otimes I|\tilde{\psi}\rangle\right) = 0$. This completes the proof of Proposition 2.

4.3 The Frobenius norm of T

In this section we prove Proposition 3. We use the following lemma:

Lemma 4 *Let*

$$S = \frac{1}{\sqrt{r}} \sum_{i=1}^{r} |u_i\rangle\langle v_i|$$

where the vectors $|v_i\rangle$, $i = 1, \ldots, r$ *are orthnormal. Then,*

$$\|S\| = \sqrt{\frac{\sum_{i=1}^r \|u_i\|^2}{r}}$$

Proof. $\|S\| = \sqrt{Tr(SS^\dagger)} = \sqrt{\frac{1}{r}\sum_{i=1}^r Tr\,|u_i\rangle\langle u_i|} = \sqrt{\frac{1}{r}\sum_{i=1}^r \|u_i\|^2}$ $\Box$.

Apply Lemma 4 to the operator T to get $\|T\| = 1$. This concludes the proof of Proposition 3.

4.4 The expression for $(A_i \otimes I)T - T(\tilde{A}_i \otimes I)$

In this section we prove Proposition 4.

Consider $T(\tilde{A}_i \otimes I)$:

$$T(\tilde{A}_i \otimes I) = \frac{1}{\sqrt{2^n}} \sum_{(j_1 \ldots j_n) \in \{0,1\}^n} A_1^{j_1} \ldots A_n^{j_n} \otimes I|\psi\rangle\langle\tilde{\psi}| \left(\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n} \otimes I\right)^\dagger (\tilde{A}_i \otimes I)$$
$$= \frac{1}{\sqrt{2^n}} \sum_{(j_1 \ldots j_n) \in \{0,1\}^n} A_1^{j_1} \ldots A_n^{j_n} \otimes I|\psi\rangle\Big((\tilde{A}_i \otimes I)(\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n} \otimes I|\tilde{\psi}\rangle)\Big)^\dagger$$

Now we use the anti-commutation relations to insert $\tilde{A}_i$ into the product $\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n}$. This possibly incurs a minus sign, depending on the particular i and the particular $(j_1 \ldots j_n) \in \{0,1\}^n$. We define $sign(i, j_1, \ldots j_n)$ to be such that

$$(\tilde{A}_i)(\tilde{A}_1^{j_1} \ldots \tilde{A}_n^{j_n}) = sign(i, j_1, \ldots j_n) \tilde{A}_1^{j_1} \ldots \tilde{A}_i^{j_i \oplus 1} \ldots \tilde{A}_n^{j_n}$$

Using this, we get

$$T(\tilde{A}_i \otimes I)$$
$$= \frac{1}{\sqrt{2^n}} \sum_{(j_1 \ldots j_n) \in \{0,1\}^n} A_1^{j_1} \ldots A_n^{j_n} \otimes I|\psi\rangle\Big(sign(i, j_1, \ldots j_n) \tilde{A}_1^{j_1} \ldots \tilde{A}_i^{j_i \oplus 1} \ldots \tilde{A}_n^{j_n} \otimes I|\tilde{\psi}\rangle\Big)^\dagger$$

Now we change the index of summation, and use $sign(i, j_1, \dots j_i \dots j_n) = sign(i, j_1, \dots j_i \oplus 1 \dots j_n)$ to get

$$T(\tilde{A}_i \otimes I) = \frac{1}{\sqrt{2^n}} \sum_{(j_1 \dots j_n) \in \{0,1\}^n} sign(i, j_1, \dots j_n) A_1^{j_1} \dots A_i^{j_i \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle\langle\tilde{\psi}| \left(\tilde{A}_1^{j_1} \dots \tilde{A}_n^{j_n} \otimes I\right)^\dagger$$

From here, identity (8) follows. This completes the proof of Proposition 4.

4.5 **The expression for** $(I \otimes B_{kl})T - T(I \otimes \tilde{B}_{kl})$

In this section we prove Proposition 5. The argument is similar to the previous section. Consider $T(I \otimes \tilde{B}_{kl})$:

$$T(I \otimes \tilde{B}_{kl}) = \frac{1}{\sqrt{2^n}} \sum_{(j_1 \dots j_n) \in \{0,1\}^n} A_1^{j_1} \dots A_n^{j_n} \otimes I|\psi\rangle \left((\tilde{A}_1^{j_1} \dots \tilde{A}_n^{j_n} \otimes I)(I \otimes \tilde{B}_{kl})|\tilde{\psi}\rangle\right)^\dagger$$
$$= \frac{1}{\sqrt{2^n}} \sum_{(j_1 \dots j_n) \in \{0,1\}^n} A_1^{j_1} \dots A_n^{j_n} \otimes I|\psi\rangle \left((\tilde{A}_1^{j_1} \dots \tilde{A}_n^{j_n} \otimes I)(\frac{\pm\tilde{A}_k + \tilde{A}_l}{\sqrt{2}} \otimes I)|\tilde{\psi}\rangle\right)^\dagger$$

where $+\tilde{A}_k$ is taken if $k < l$ and $-\tilde{A}_k$ is taken if $k > l$.

Next, we use the anti-commutation relations to insert $\tilde{A}_k$ and $\tilde{A}_l$ into the product $\tilde{A}_1^{j_1} \dots \tilde{A}_n^{j_n}$. We get

$$T(I \otimes \tilde{B}_{kl}) = \frac{1}{\sqrt{2^n}} \sum_{(j_1 \dots j_n) \in \{0,1\}^n} A_1^{j_1} \dots A_n^{j_n} \otimes I|\psi\rangle$$
$$\frac{1}{\sqrt{2}} \left(\pm \left(sign(j_1, \dots j_n, k) \tilde{A}_1^{j_1} \dots \tilde{A}_k^{j_k \oplus 1} \dots \tilde{A}_n^{j_n} \otimes I|\tilde{\psi}\rangle\right)^\dagger + \left(sign(j_1, \dots j_n, l) \tilde{A}_1^{j_1} \dots \tilde{A}_l^{j_l \oplus 1} \dots \tilde{A}_n^{j_n} \otimes I|\tilde{\psi}\rangle\right)^\dagger \right)$$

We separate into two sums and change the index of summation in each and we get

$$T(I \otimes \tilde{B}_{kl}) = \frac{1}{\sqrt{2^n}} \sum_{(j_1 \dots j_n) \in \{0,1\}^n} \frac{1}{\sqrt{2}} \left(\pm sign(j_1, \dots j_n, k) A_1^{j_1} \dots A_k^{j_k \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle + sign(j_1, \dots j_n, l) A_1^{j_1} \dots A_l^{j_l \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle \right) \langle\tilde{\psi}| \left(\tilde{A}_1^{j_1} \dots \tilde{A}_n^{j_n} \otimes I\right)^\dagger$$

From here, identity (9) follows. This completes the proof of Proposition 5.

4.6 **The first error bound**

In this section we prove Proposition 6.

We would like to insert A_i into the product $A_1^{j_1} \dots A_n^{j_n}$ as if the A_i, $i = 1, \dots n$ were anti-commuting. However, we don't know that A_i, $i = 1, \dots n$ are anti-commuting; all we know about the A_i, $i = 1, \dots n$ is that they are part of an ϵ-optimal CHSH(n) strategy.

Even though A_i, $i = 1, \dots n$ may not be anti-commuting as operators, they nearly anti-commute in their action on the strategy state $|\psi\rangle$. We prove the following:

Lemma 5 *Let $A_i, B_{jk}, |\psi\rangle$ be an ϵ-optimal CHSH(n) strategy. Then,*

$$\sum_{1 \le i < j \le n} \left\| \frac{A_i A_j + A_j A_i}{2} \otimes I |\psi\rangle \right\|^2 \le (1+\sqrt{2})^2 n(n-1)\epsilon$$

Proof. The operators

$$\frac{A_i + A_j}{\sqrt{2}} \otimes I + I \otimes B_{ij}$$

and

$$\frac{A_i - A_j}{\sqrt{2}} \otimes I + I \otimes B_{ji}$$

each have operator norm at most $(1+\sqrt{2})$, by the triangle inequality.

Next, we see that

$$\begin{aligned} \left\| \frac{A_i A_j + A_j A_i}{2} \otimes I |\psi\rangle \right\| &= \left\| \left(\frac{A_i + A_j}{\sqrt{2}} \otimes I + I \otimes B_{ij} \right) \left(\frac{A_i + A_j}{\sqrt{2}} \otimes I - I \otimes B_{ij} \right) |\psi\rangle \right\| \\ &\le (1+\sqrt{2}) \left\| \left(\frac{A_i + A_j}{\sqrt{2}} \otimes I - I \otimes B_{ij} \right) |\psi\rangle \right\| \end{aligned}$$

and similarly,

$$\begin{aligned} \left\| \frac{A_i A_j + A_j A_i}{2} \otimes I |\psi\rangle \right\| &= \left\| \left(\frac{A_i - A_j}{\sqrt{2}} \otimes I + I \otimes B_{ji} \right) \left(\frac{A_i - A_j}{\sqrt{2}} \otimes I - I \otimes B_{ji} \right) |\psi\rangle \right\| \\ &\le (1+\sqrt{2}) \left\| \left(\frac{A_i - A_j}{\sqrt{2}} \otimes I - I \otimes B_{ji} \right) |\psi\rangle \right\| \end{aligned}$$

Now use inequality (6) from Proposition 1 to get

$$\begin{aligned} &\sum_{1 \le i < j \le n} 2 \left\| \frac{A_i A_j + A_j A_i}{2} \otimes I |\psi\rangle \right\|^2 \\ &\le (1+\sqrt{2})^2 \sum_{1 \le i < j \le n} \left(\left\| \frac{A_i + A_j}{\sqrt{2}} \otimes I |\psi\rangle - I \otimes B_{ij} |\psi\rangle \right\|^2 + \left\| \frac{A_i - A_j}{\sqrt{2}} \otimes I |\psi\rangle - I \otimes B_{ji} |\psi\rangle \right\|^2 \right) \\ &\le (1+\sqrt{2})^2 2n(n-1)\epsilon \end{aligned}$$

Lemma 5 is proved. □.

Now we know that A_i, $i = 1, \dots n$ almost anti-commute in their action on the strategy state $|\psi\rangle$. This is a step forward, but still not enough to prove inequality (10). To see why, consider a product like $A_i A_1 A_2 \otimes I |\psi\rangle$. We want to switch the order of A_i and A_1. We know that A_i and A_1 nearly anti-commute in their action on $|\psi\rangle$, but we don't yet know that they nearly anti-commute in their action on $A_2 \otimes I |\psi\rangle$.

Fortunately, this difficulty can be avoided: we know from Proposition 1 that

$$A_2 \otimes I |\psi\rangle \approx I \otimes \frac{1}{\sqrt{2}} (B_{12} - B_{21}) |\psi\rangle$$

This helps, because

$$(A_iA_1 \otimes I)(A_2 \otimes I)|\psi\rangle \approx (A_iA_1 \otimes I)(I \otimes \frac{1}{\sqrt{2}}(B_{12} - B_{21}))|\psi\rangle$$
$$= (I \otimes \frac{1}{\sqrt{2}}(B_{12} - B_{21}))(A_iA_1 \otimes I)|\psi\rangle$$

and now we can switch the order of A_i and A_1 in their action on $|\psi\rangle$.

Now we see that as long as we can "get some of the A_i's out of the way", we can use the anti-commutation on $|\psi\rangle$ to switch the order of a product of the A_i's acting on $|\psi\rangle$. To keep the errors of approximation under control, we want the operators on the B side that we use to have operator norm 1. The operators $\frac{1}{\sqrt{2}}(B_{ij} \pm B_{ji})$ do not necessarily have operator norm 1, but fortunately this difficulty can also be avoided.

The discussion in the previous paragraphs motivates us to prove the following lemma:

Lemma 6 *Fix k. Then, there exists an l such that*

$$\left\| A_k \otimes I|\psi\rangle - I \otimes \frac{\pm B_{kl} + B_{lk}}{|\pm B_{kl} + B_{lk}|}|\psi\rangle \right\| \leq (2\sqrt{2} + 2)\sqrt{n}\sqrt{\epsilon} \tag{18}$$

where $+B_{kl}$ is taken if $l > k$ and $-B_{kl}$ is taken if $l < k$. The notation

$$\frac{\pm B_{kl} + B_{lk}}{|\pm B_{kl} + B_{lk}|}$$

means that we take all eigenvalues of the operator $\pm B_{kl} + B_{lk}$ and normalize the positive ones to 1, the negative ones to -1, and, by convention, the eigenvalue 0 gets normalized to 1.

Proof. There are two steps:

$$\left\| A_k \otimes I|\psi\rangle - I \otimes \frac{\pm B_{kl} + B_{lk}}{\sqrt{2}}|\psi\rangle \right\| \leq \sqrt{2}\sqrt{n}\sqrt{\epsilon} \tag{19}$$

and

$$\left\| I \otimes \frac{\pm B_{kl} + B_{lk}}{\sqrt{2}}|\psi\rangle - I \otimes \frac{\pm B_{kl} + B_{lk}}{|\pm B_{kl} + B_{lk}|}|\psi\rangle \right\| \leq (2 + \sqrt{2})\sqrt{n}\sqrt{\epsilon} \tag{20}$$

We prove the first step. We take relation (7) from Proposition 1. We focus only on those terms of the sum that contain A_k and we get

$$\sum_{j=k+1}^{n} \left\| A_k \otimes I|\psi\rangle - I \otimes \frac{B_{kj} + B_{jk}}{\sqrt{2}}|\psi\rangle \right\|^2 + \sum_{j=1}^{k-1} \left\| A_k \otimes I|\psi\rangle - I \otimes \frac{-B_{kj} + B_{jk}}{\sqrt{2}}|\psi\rangle \right\|^2 \leq 4\binom{n}{2}\epsilon$$

Pick the smallest of these $(n-1)$ terms. It satisfies

$$\left\| A_k \otimes I|\psi\rangle - I \otimes \frac{\pm B_{kl} + B_{lk}}{\sqrt{2}}|\psi\rangle \right\|^2 \leq 2n\epsilon$$

This proves inequality (19).

Now we focus on the second step. By Lemma 7 which we will prove below,

$$\left\| I \otimes \frac{\pm B_{kj} + B_{jk}}{\sqrt{2}}|\psi\rangle - I \otimes \frac{\pm B_{kl} + B_{lk}}{|\pm B_{kl} + B_{lk}|}|\psi\rangle \right\| \leq \left\| I \otimes \frac{B_{kl}B_{lk} + B_{lk}B_{kl}}{2}|\psi\rangle \right\| \tag{21}$$

Next, observe that the operator

$$A_k \otimes I + I \otimes \frac{\pm B_{kl} + B_{lk}}{\sqrt{2}}$$

has operator norm at most $(1+\sqrt{2})$, and so

$$\left\| I \otimes \frac{B_{kl}B_{lk} + B_{lk}B_{kl}}{2}|\psi\rangle \right\| = \left\| \left(A_k \otimes I + I \otimes \frac{\pm B_{kl} + B_{lk}}{\sqrt{2}} \right) \left(A_k \otimes I - I \otimes \frac{\pm B_{kl} + B_{lk}}{\sqrt{2}} \right) |\psi\rangle \right\|$$
$$\leq (1+\sqrt{2}) \left\| A_k \otimes I|\psi\rangle - I \otimes \frac{\pm B_{kl} + B_{lk}}{\sqrt{2}}|\psi\rangle \right\| \leq (1+\sqrt{2})\sqrt{2n\epsilon}$$

This proves inequality (20).

Now, we combine inequalities (19) and (20), and get inequality (18). This completes the proof of Lemma 6. □.

Next, we prove a Lemma that implies inequality (21) which we used in the proof of Lemma 6. This has to do with operators of the form $\frac{R+S}{\sqrt{2}}$ and $\frac{R+S}{|R+S|}$ when R, S are ± 1 observables.

Lemma 7 *Let R, S be two ± 1 observables on $\mathbb{C}^d$. Then,*

1. *The following operator identity holds:*

$$\left(\frac{R+S}{\sqrt{2}} - \frac{R+S}{|R+S|} \right)^2$$
$$= \left(\frac{RS+SR}{2} \right) \left(2I + \frac{RS+SR}{2} + 2\sqrt{I + \frac{RS+SR}{2}} \right)^{-1} \left(\frac{RS+SR}{2} \right) \tag{22}$$

2. *The operator*

$$\left(\frac{RS+SR}{2} \right)^2 - \left(\frac{R+S}{\sqrt{2}} - \frac{R+S}{|R+S|} \right)^2$$

is positive semi-definite.

3. *For any vector $|v\rangle$,*

$$\left\| \frac{R+S}{\sqrt{2}}|v\rangle - \frac{R+S}{|R+S|}|v\rangle \right\| \leq \left\| \frac{RS+SR}{2}|v\rangle \right\|$$

Proof. First, we prove the operator identity. We will show that the two operators have the same eigenvalues and eigenspaces.

We break up $\mathbb{C}^d$ into eigenspaces for the self-adjoint operator $R+S$. Since $RS+SR = (R+S)^2 - 2I$, these are also eigenspaces for the operator $RS+SR$, and so also eigenspaces for the operator

$$\left(\frac{RS+SR}{2} \right) \left(2I + \frac{RS+SR}{2} + 2\sqrt{I + \frac{RS+SR}{2}} \right)^{-1} \left(\frac{RS+SR}{2} \right)$$

Consider an eigenspace where $R+S$ has eigenvalue λ.

On this eigenspace, the operator

$$\left(\frac{R+S}{\sqrt{2}} - \frac{R+S}{|R+S|}\right)^2$$

has eigenvalue $\left((sign\lambda)\lambda/\sqrt{2}-1\right)^2$; this holds in all the three cases $\lambda > 0,\ \lambda < 0,\ \lambda = 0$.

The eigenvalue of $(RS+SR)/2$ on this eigenspace is $(\lambda^2-2)/2$.

The eigenvalue of

$$\left(\frac{RS+SR}{2}\right)\left(2I + \frac{RS+SR}{2} + 2\sqrt{I + \frac{RS+SR}{2}}\right)^{-1}\left(\frac{RS+SR}{2}\right)$$

on this eigenspace is

$$\left(\frac{\lambda^2-2}{2}\right)^2 \frac{1}{2+\frac{\lambda^2-2}{2}+2\sqrt{1+\frac{\lambda^2-2}{2}}} = \left(\frac{(sign\lambda)\lambda}{\sqrt{2}} - 1\right)^2$$

Therefore the operators

$$\left(\frac{R+S}{\sqrt{2}} - \frac{R+S}{|R+S|}\right)^2$$

and

$$\left(\frac{RS+SR}{2}\right)\left(2I + \frac{RS+SR}{2} + 2\sqrt{I + \frac{RS+SR}{2}}\right)^{-1}\left(\frac{RS+SR}{2}\right)$$

have the same eigenvalue on this eigenspace.

The argument works for any eigenspace, so operator identity (22) holds.

Next we prove the second part. We can see from the argument above that the operator

$$\left(2I + \frac{RS+SR}{2} + 2\sqrt{I + \frac{RS+SR}{2}}\right)^{-1}$$

has eigenvalues of the form

$$\frac{1}{\left(\frac{(sign\lambda)\lambda}{\sqrt{2}}+1\right)^2}$$

and they are all in $(0,1]$. Therefore,

$$\left(\frac{R+S}{\sqrt{2}} - \frac{R+S}{|R+S|}\right)^2 = \left(\frac{RS+SR}{2}\right)\left(2I + \frac{RS+SR}{2} + 2\sqrt{I + \frac{RS+SR}{2}}\right)^{-1}\left(\frac{RS+SR}{2}\right) \preceq \left(\frac{RS+SR}{2}\right)^2$$

Finally, we observe that the third part follows directly from the second. Lemma 7 is proved. $\square$.

Recall that the goal of this section is to prove Proposition 6 and the overall strategy is to insert A_i into the product $A_1^{j_1} \dots A_n^{j_n}$ as if the A_i, $i = 1, \dots n$ were anti-commuting. We have prepared the tools necessary for this goal. Lemma 5 tells us that

$$A_k A_l \otimes I|\psi\rangle \approx -A_l A_k \otimes I|\psi\rangle$$

with an error of approximation at most $(2\sqrt{2}+2)n\sqrt{\epsilon}$. We call this apporoximation step an anticommutation switch. Lemma 6 tells us that

$$A_k \otimes I|\psi\rangle \approx I \otimes \frac{\pm B_{kl} + B_{lk}}{|\pm B_{kl} + B_{lk}|}|\psi\rangle$$

where $\frac{\pm B_{kl}+B_{lk}}{|\pm B_{kl}+B_{lk}|}$ is a suitable ± 1 observable acting on the B side, and the error of approximation is at most $(2\sqrt{2}+2)\sqrt{n}\sqrt{\epsilon}$. We call this approximation step an AB-switch.

We concatenate a number of these approximation steps to get inequality (10). We present a procedure that goes from $A_i A_1^{j_1} \dots A_n^{j_n} \otimes I|\psi\rangle$ to $sign(i, j_1, \dots j_n) A_1^{j_1} \dots A_i^{j_i \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle$ using at most n anti-commutator switches and $2n$ AB-switches:

1. Start with $A_i A_1^{j_1} \dots A_n^{j_n} \otimes I|\psi\rangle$.
2. Switch all elements of the product $A_1^{j_1} \dots A_n^{j_n}$ to the B side using the AB-switches.
3. Repeat
   (a) Switch the last observable on the B side back to the A side
   (b) Anti-commute A_i and the newly switched observable

   until A_i comes to its proper position.
4. Switch the observables still remaining on the B side back to the A side.

The total approximation error of this procedure is at most

$$n(2\sqrt{2}+2)n\sqrt{\epsilon} + (2n)(2\sqrt{2}+2)\sqrt{n}\sqrt{\epsilon} \leq (6+4\sqrt{2})n^2\sqrt{\epsilon} < 12n^2\sqrt{\epsilon}$$

Inequality (10) is proved. This completes the proof of Proposition 6.

4.7 The second error bound

In this section we prove Proposition 7. The argument is similar to the previous section.

By the triangle inequality, we have

$$\begin{aligned}
&\Bigg\| A_1^{j_1} \dots A_n^{j_n} \otimes B_{kl}|\psi\rangle - \frac{1}{\sqrt{2}}\Big(\pm\, sign(j_1, \dots j_n, k) A_1^{j_1} \dots A_k^{j_k \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle \\
&\qquad + sign(j_1, \dots j_n, l) A_1^{j_1} \dots A_l^{j_l \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle\Big)\Bigg\| \\
&\leq \left\| A_1^{j_1} \dots A_n^{j_n} \otimes B_{kl}|\psi\rangle - A_1^{j_1} \dots A_n^{j_n} \frac{\pm A_k + A_l}{\sqrt{2}} \otimes I|\psi\rangle \right\| \\
&+ \frac{1}{\sqrt{2}} \left\| A_1^{j_1} \dots A_n^{j_n} A_k \otimes I|\psi\rangle - sign(j_1, \dots j_n, k) A_1^{j_1} \dots A_k^{j_k \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle \right\| \\
&+ \frac{1}{\sqrt{2}} \left\| A_1^{j_1} \dots A_n^{j_n} A_l \otimes I|\psi\rangle - sign(j_1, \dots j_n, l) A_1^{j_1} \dots A_l^{j_l \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle \right\| \qquad (23)
\end{aligned}$$

For the first term we have:

$$\left\| A_1^{j_1} \dots A_n^{j_n} \otimes B_{kl}|\psi\rangle - A_1^{j_1} \dots A_n^{j_n} \frac{\pm A_k + A_l}{\sqrt{2}} \otimes I|\psi\rangle \right\| = \left\| I \otimes B_{kl}|\psi\rangle - \frac{\pm A_k + A_l}{\sqrt{2}} \otimes I|\psi\rangle \right\| \leq \sqrt{2n(n-1)\epsilon} \quad (24)$$

where we have used the inequalities in Proposition 1.

For the second term, we claim that

$$\left\| A_1^{j_1} \dots A_n^{j_n} A_k \otimes I|\psi\rangle - sign(j_1, \dots j_n, k) A_1^{j_1} \dots A_k^{j_k \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle \right\| \leq (6 + 4\sqrt{2}) n^2 \sqrt{\epsilon} \quad (25)$$

The argument is similar to the argument in the previous section: we present a procedure that goes from $A_1^{j_1} \dots A_n^{j_n} A_k \otimes I|\psi\rangle$ to $sign(j_1, \dots j_n, k) A_1^{j_1} \dots A_k^{j_k \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle$ using at most n anti-commutator switches and $2n$ AB-switches. The procedure is the following:

1. Start with $A_1^{j_1} \dots A_n^{j_n} A_k \otimes I|\psi\rangle$.
2. Repeat
   (a) Anti-commute A_k and the next to last observable on the A side
   (b) Move the newly switched observable to the B side

   until A_k comes to its proper position.
3. Switch the observables still remaining on the B side back to the A side.

The third term is analyzed in the same manner and we get

$$\left\| A_1^{j_1} \dots A_n^{j_n} A_l \otimes I|\psi\rangle - sign(j_1, \dots j_n, l) A_1^{j_1} \dots A_l^{j_l \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle \right\| \leq (6 + 4\sqrt{2}) n^2 \sqrt{\epsilon} \quad (26)$$

Finally, we combine inequalities (23), (24), (25) and (26) to get inequality (11). This completes the proof of Proposition 7.

4.8 Putting everything together

The aim of this subsection is to put all the previous steps together and prove Theorem 1.

We start with the first of inequalities (5). We know from Proposition 4 that

$$(A_i \otimes I)T - T(\tilde{A}_i \otimes I) = \frac{1}{\sqrt{2^n}} \sum_{(j_1 \dots j_n) \in \{0,1\}^n} \Big(A_i A_1^{j_1} \dots A_n^{j_n} \otimes I|\psi\rangle - sign(i, j_1, \dots j_n) A_1^{j_1} \dots A_i^{j_i \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle \Big) \langle\tilde{\psi}| \left(\tilde{A}_1^{j_1} \dots \tilde{A}_n^{j_n} \otimes I \right)^\dagger$$

We know from Proposition 2 that the vectors

$$\left\{ \tilde{A}_1^{j_1} \dots \tilde{A}_n^{j_n} \otimes I|\tilde{\psi}\rangle \ : \ (j_1 \dots j_n) \in \{0,1\}^n \right\}$$

are orthonormal.

We also know from Proposition 6 that for all i, for all $(j_1 \dots j_n) \in \{0,1\}^n$

$$\left\| A_i A_1^{j_1} \dots A_n^{j_n} \otimes I|\psi\rangle - sign(i, j_1, \dots j_n) A_1^{j_1} \dots A_i^{j_i \oplus 1} \dots A_n^{j_n} \otimes I|\psi\rangle \right\| < 12n^2\sqrt{\epsilon}$$

We combine these facts using Lemma 4 and we get that for all i,

$$\|(A_i \otimes I)T - T(\tilde{A}_i \otimes I)\| < 12n^2\sqrt{\epsilon} = 12n^2\sqrt{\epsilon}\|T\|$$

where in the last step we have used $\|T\| = 1$ (Proposition 3).

In a similar manner, we take the results of Propositions 2, 3, 5, and 7, apply Lemma 4, and get that for all $j \neq k \in \{1, \dots n\}$

$$\|(I \otimes B_{jk})T - T(I \otimes \tilde{B}_{jk})\| < 17n^2\sqrt{\epsilon}\|T\|$$

The proof of Theorem 1 is complete.

5 Conclusion and open problems

In this paper, we focused on the CHSH(n) XOR games, and derived the structure of their nearly-optimal quantum strategies.

One possible direction for future work is whether structure results like the one for CHSH(n) nearly-optimal quantum strategies can be proved for other non-local games. The optimal quantum strategies for the CHSH(n) games are related to representations of the Clifford algebra with n anti-commuting generators, and the arguments in this paper use this connection. However, it may be possible to construct an argument of this form, or another form altogether, for other XOR games with less regular structure.

Another possible direction for future work is whether the CHSH(n) games can be used in device-independent protocols. The CHSH game, the first member of the CHSH(n) family, has already been used in device-independent protocols. Whether all the CHSH(n) games can be used, and which of the CHSH(n) games gives protocols with the best parameters, are two questions that are still open.

Acknowledgements

The results of this paper first appear in my PhD thesis submitted to the Department of Mathematics at Massachusetts Institute of Technology. The material is used here with permission from MIT.

I would like to thank my thesis advisor Prof. Peter Shor for his unconditional support through the years. Prof. Shor gave me the freedom I needed to explore, and to find my own way. He was also generous with his time, and patiently listened to my mathematical arguments and ideas.

I would like to thank Thomas Vidick for bringing to my attention the problem of self-testing and entanglement rigidity. Thomas has always been friendly, enthusiastic, and open to discussion. The conversations with him have been a source of many great ideas.

I would like to thank an anonymous referee for helpful comments and suggestions.

References

1. R. F. Werner and M. M. Wolf (2001), *Bell inequalities and entanglement*, arXiv preprint quant-ph/0107093.
2. N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani and S. Wehner (2014), *Bell nonlocality*, Reviews of Modern Physics, vol. 86(2), p. 419.
3. J. F. Clauser and A. Shimony (1978), *Bell's theorem. experimental tests and implications*, Reports on Progress in Physics, vol. 41(12), p. 1881.
4. H. Buhrman, R. Cleve, S. Massar and R. de Wolf (2010), *Nonlocality and communication complexity*, Reviews of modern physics, vol. 82(1), p. 665.
5. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. K. Wootters (1993), *Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels*, Physical review letters, vol. 70(13), p. 1895.
6. D. Gottesman and I. L. Chuang (1999), *Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations*, Nature, vol. 402(6760), pp. 390–393.
7. N. Gisin, G. Ribordy, W. Tittel and H. Zbinden (2002), *Quantum cryptography*, Reviews of modern physics, vol. 74(1), p. 145.
8. D. Mayers and A. Yao (1998), *Quantum cryptography with imperfect apparatus*, in *Foundations of Computer Science, 1998. Proceedings. 39th Annual Symposium on*, pp. 503–509, IEEE.
9. D. Mayers and A. Yao (2003), *Self testing quantum apparatus*, arXiv preprint quant-ph/0307205.
10. U. Vazirani and T. Vidick (2012), *Fully device independent quantum key distribution.*, arXiv preprint arXiv:1210.1810, vol. 2(11).
11. B. W. Reichardt, F. Unger and U. Vazirani (2012), *A classical leash for a quantum system: Command of quantum systems via rigidity of chsh games*, arXiv preprint arXiv:1209.0448.
12. M. McKague, T. H. Yang and V. Scarani (2012), *Robust self-testing of the singlet*, Journal of Physics A: Mathematical and Theoretical, vol. 45(45), p. 455304.
13. C. A. Miller and Y. Shi (2012), *Optimal robust quantum self-testing by binary nonlocal xor games*, arXiv preprint arXiv:1207.1819.
14. W. Slofstra (2011), *Lower bounds on the entanglement needed to play xor non-local games*, Journal of Mathematical Physics, vol. 52(10), p. 102202.
15. L. Vandenberghe and S. Boyd (1996), *Semidefinite programming*, SIAM review, vol. 38(1), pp. 49–95.
16. L. Lovasz, *Semi-definite programs and combinatorial optimization*, Lecture Notes available online at http://www.cs.elte.hu/ lovasz/semidef.ps, URL `www.cs.elte.hu/~lovasz/semidef.ps`.
17. B. S. Tsirel'son (1987), *Quantum analogues of the bell inequalities. the case of two spatially separated domains*, Journal of Soviet Mathematics, vol. 36(4), pp. 557–570.
18. R. Cleve, P. Hoyer, B. Toner and J. Watrous (2004), *Consequences and limits of nonlocal strategies*, in *Computational Complexity, 2004. Proceedings. 19th IEEE Annual Conference on*, pp. 236–249, IEEE.
19. J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt (1969), *Proposed experiment to test local hidden-variable theories*, Physical review letters, vol. 23, pp. 880–884.
20. W. Fulton and J. Harris (1991), *Representation theory*, Springer-Verlag New York.
21. P. Etingof, *Introduction to representation theory*, Lecture Notes available online at http://ocw.mit.edu/courses/mathematics/18-712-introduction-to-representation-theory-fall-2010/lecture-notes/, URL `http://ocw.mit.edu/courses/mathematics/18-712-introduction-to-representation-theory-fall-2010/lecture-notes/`.