Bounding quantum-classical separations for classes of nonlocal games

Tom Bannink[*]
bannink@cwi.nl

Jop Briët[*†]
j.briet@cwi.nl

Harry Buhrman[*‡]
buhrman@cwi.nl

Farrokh Labib[*]
labib@cwi.nl

Troy Lee[§]
troyjlee@gmail.com

Abstract

We bound separations between the entangled and classical values for several classes of nonlocal t-player games. Our motivating question is whether there is a family of t-player XOR games for which the entangled bias is 1 but for which the classical bias goes down to 0, for fixed t. Answering this question would have important consequences in the study of multi-party communication complexity, as a positive answer would imply an unbounded separation between randomized communication complexity with and without entanglement. Our contribution to answering the question is identifying several general classes of games for which the classical bias can not go to zero when the entangled bias stays above a constant threshold. This rules out the possibility of using these games to answer our motivating question. A previously studied set of XOR games, known not to give a positive answer to the question, are those for which there is a quantum strategy that attains value 1 using a so-called Schmidt state. We generalize this class to mod-m games and show that their classical value is always at least $\frac{1}{m} + \frac{m-1}{m} t^{1-t}$. Secondly, for free XOR games, in which the input distribution is of product form, we show $\beta(G) \geq \beta^*(G)^{2^t}$ where $\beta(G)$ and $\beta^*(G)$ are the classical and entangled biases of the game respectively. We also introduce so-called line games, an example of which is a slight modification of the Magic Square game, and show that they can not give a positive answer to the question either. Finally we look at two-player unique games and show that if the entangled value is $1 - \varepsilon$ then the classical value is at least $1 - \mathcal{O}(\sqrt{\varepsilon \log k})$ where k is the number of outputs in the game. Our proofs use semidefinite-programming techniques, the Gowers inverse theorem and hypergraph norms.

1

1 Introduction

The study of multiplayer games has been extremely fruitful in theoretical computer science across diverse areas including the study of complexity classes [BOGKW88], hardness of approximation [Kho02], and communication complexity [KLL$^+$15]. They are also a great framework in which to study Bell inequalities [Bel64] and analyze the nonlocal properties of entanglement. A particularly simple kind of multiplayer game is an XOR game. An XOR game $G = (f, \pi)$ between t-players is defined by a function $f : X_1 \times X_2 \times \cdots \times X_t \to \{0, 1\}$ and a probability distribution π over $X_1 \times \cdots \times X_t$. An input $(x_1, \ldots, x_t) \in X_1 \times \cdots \times X_t$ is chosen by a referee according to π, who then gives x_i to player i. Without communicating, player i then outputs a bit $a_i \in \{0, 1\}$ with the collective goal of the players being that $a_1 \oplus \cdots \oplus a_t = f(x_1, \ldots, x_t)$. In a classical XOR game, the players' strategies are deterministic. In an XOR game with entanglement, players are allowed to share a quantum state and make measurements on this state to inform their outputs.

As players can always win an XOR game with probability at least $\frac{1}{2}$, it is common to study the *bias* of an XOR game, the probability of winning minus the probability of losing. We use $\beta(G)$ to denote the largest bias achievable by a classical protocol for the game G, and $\beta^*(G)$ to denote the best bias achievable by a protocol using shared entanglement for the game G.

Our motivating question in this paper is:

Question 1.1. Is there a family of t-player XOR games $(G_n)_{n \in \mathbb{N}}$ such that $\beta^*(G_n) = 1$ and $\beta(G_n) \to 0$ as $n \to \infty$?

This question has important implications for multi-party communication complexity. For a function $f : X_1 \times \cdots X_t \to \{0, 1\}$, let $R(f)$ denote the t-party randomized communication complexity of f (in the number-in-the-hand model), and let $R^*(f)$ denote the t-party randomized communication complexity of f where the parties are allowed to share entanglement. A positive answer to Question 1.1 gives a family of functions $(f_n)_{n \in \mathbb{N}}$ with $R^*(f_n) = O(1)$ and $R(f_n) = \omega(1)$, i.e. an unbounded separation between these two communication models.

In the reverse direction, a family of functions $(f_n)_{n \in \mathbb{N}}$ with $R^*(f_n) = O(1)$ and $R(f_n) = \omega(1)$ gives a family of games $G_n = (f_n, \pi_n)$ with $\beta^*(G_n) \geq c$ for some constant c and $\beta(G_n) \to 0$ as $n \to \infty$. Thus there is a very close connection between Question 1.1 and the existence of an unbounded separation between randomized communication complexity with and without entanglement.

For the two-player case, it is known that the answer to Question 1.1 is negative. It was observed by Tsirelson [Tsi87] that Grothendieck's inequality [Gro53], a fundamental result from Banach space theory, is equivalent to the assertion that $\beta^*(G) \leq K_G \cdot \beta(G)$, where $K_G \leq 1.78\ldots$ [Kri77, BMMN11] is Grothendieck's constant.

Linial and Shraibman [LS09] and Shi and Zhu [SZ08] realized that the XOR bias of a game (f, π) can be used to lower bound the communication complexity of f, both in the randomized setting and the setting with entanglement. Together with Grothendieck's inequality they used this to show that $R(f) = O(2^{2R^*(f)})$ for any partial two-party function f. Thus in the two-party case an unbounded communication separation is not possible between the randomized model with and without entanglement. Raz has given an example of a partial function f with $R(f) = 2^{\Omega(R^*(f))}$ [Raz99], thus the upper bound of Linial-Shraibman and Shi-Zhu is essentially optimal.

In the case of three or more parties, Question 1.1 and the corresponding question of an unbounded separation between the entangled and non-entangled communication complexity models remain open. A striking result of Peréz-García et al. [PGWP$^+$08] shows that there is no analogue of Grothendieck's inequality in the three-player setting. In particular, they showed that there

exists an infinite family of three-player XOR games $(G_n)_{n\in\mathbb{N}}$ with the property that the ratio of the entangled and classical biases of G_n goes to infinity with n. This result was later quantitatively improved by Briët and Vidick [BV13]. Both results rely crucially on non-constructive (probabilistic) methods, and in both separating examples the entangled bias $\beta^*(G_n)$ also goes to zero with increasing n. These works leave open the question, posed explicitly in [BV13], of whether there is such a family of games in which the entangled bias does not vanish with n, but instead stays above a fixed positive threshold while the classical bias decays to zero. Crucially, having a separation in XOR bias where $\beta^*(G_n)$ remains constant is what is needed to also obtain an unbounded separation between randomized communication complexity with and without entanglement.

Our contribution to answering Question 1.1 One approach to Question 1.1 is to look at different classes of games and identify which ones could possibly lead to a positive answer.

Peréz-García et al. [PGWP$^+$08] show that in any XOR game where the entangled strategy uses a GHZ state, there is a bounded gap between the classical and entangled bias: namely, the bias with a GHZ state in a t-player XOR game G is at most $K_G(2\sqrt{2})^{t-1}\beta(G)$. This bound is essentially tight as there are examples of t-player XOR games achieving a ratio between the GHZ state bias and classical bias of $\frac{\pi}{2}^t$ [Zuk93]. Briët et al. [BBLV13] later extended the Grothendieck- type inequality of Peréz-García et al. to a larger class of entangled states called Schmidt states (see Equation 1). Thus any game where there is a perfect strategy where the players share a Schmidt state cannot give a positive answer to Question 1.1.

Watts et al. [WHKN18] recently investigated Question 1.1 and found that a t-player XOR game G that is symmetric, i.e. invariant under the renaming of players, and where $\beta^*(G) = 1$ always has a perfect entangled strategy where the players share a GHZ state. Thus symmetric games also cannot give a positive answer to Question 1.1.

We further study games that have a perfect strategy where players share a GHZ or Schmidt state. We do this for a generalization of XOR games called mod m games. In a mod m game the players output an integer between 0 and $m - 1$ and the goal is for the sum of the outputs mod m to equal a target value determined by their inputs. We show that the classical advantage over random guessing is at least $\frac{m-1}{m}t^{1-t}$ in any t player mod m game that can be won perfectly by sharing a Schmidt state (see Theorem 1.2).

We show this by introducing *angle games*, a class of games that can be won perfectly sharing a GHZ state and are the *hardest* of all such games. Thus a classical strategy in an angle game can be used to lower bound the winning probability of any mod m game that has a perfect Schmidt state strategy.

For small values of t we can directly analyze angle games to give bounds that are sometimes tight. One interesting consequence of our result is the following. The Mermin game G is a three-party XOR game where by sharing a GHZ state players can play perfectly, $\beta^*(G) = 1$, while classically $\beta(G) = \frac{1}{2}$. We show that this is the maximal possible separation of any 3-party XOR game where $\beta^*(G) = 1$ via a GHZ state. In particular, this means that when one looks at the XOR repetition of the Mermin game the classical bias *does not go down at all*.

We rule out other types of games that could positively answer Question 1.1 as well. A t-player *free* XOR game $G = (f, \pi)$ is a game where π is a product distribution. For such games we show that $\beta(G) \geq \beta^*(G)^{2^t}$, and thus they cannot be used for a positive answer to Question 1.1.

Another class of XOR games we consider are *line games*, where the questions asked to the players are related by a geometric property. An example of a line game is a slight modification

of the Magic Square game [IKP+08]. We show that line games cannot give a positive answer to Question 1.1 either.

Finally, we look at extensions of Question 1.1 beyond XOR games to more general classes of games like unique games [Kho02], which have been deeply studied because of their application in hardness of approximation. For unique games we show that in fact if there is strategy with entanglement that can win a unique game perfectly, then there is a perfect classical strategy as well. This can be compared with the result of Cleve et al. [CHTW04] that if a two-player game with binary outputs has a perfect strategy with entanglement then it also has a perfect classical strategy. More generally, we show that if the winning probability with entanglement is $1 - \varepsilon$ in a unique game with k outputs, then there is a classical strategy that wins with probability $1 - C\sqrt{\varepsilon \log k}$.

In the next subsections, we discuss our results in more detail.

1.1 Perfect Schmidt strategies for MOD games

A MOD-m game is a generalization of XOR games to non-binary outputs. A nonlocal game is a MOD-m game if the players are required to answer with integers from 0 to $m - 1$, and win if and only if the sum of their answers modulo m equals the target value determined by their inputs. We denote the optimal winning probability using classical strategies by $\omega(G)$, and we write $\omega^*(G)$ for the entangled winning probability. Random play in such a game ensures that the players can always win with probability at least $\frac{1}{m}$. As with XOR games, in a MOD-m game one often considers the *bias* given by the maximum amount by which the value can exceed $\frac{1}{m}$, scaled to be in the $[0,1]$ range. The bias is $\beta(G) = \frac{m}{m-1}(\omega(G) - \frac{1}{m})$, and similar for the entangled version. This generalizes the definition given for XOR games above.

Define a t-partite *Schmidt state* as a t-partite quantum state that can be written in the form

$$|\psi\rangle = \sum_{i=0}^{d-1} c_i |e_i^{(1)}\rangle |e_i^{(2)}\rangle \cdots |e_i^{(t)}\rangle, \tag{1}$$

where $c_i > 0$ and where the $|e_i^{(j)}\rangle$ ($i = 0, 1, ..., d - 1$) are orthogonal vectors in the j-th system. For $t = 2$ any state can be written this way, something commonly known as the Schmidt decomposition. Note that the well-known GHZ state is a Schmidt state where all the c_i are equal to $1/\sqrt{d}$. In the context of nonlocal games, define a *Schmidt strategy* as a quantum strategy that uses (only) a Schmidt state. We say a strategy is *perfect* if it achieves winning probability 1.

We consider t-player MOD-m games for which there is a perfect Schmidt strategy ("perfect Schmidt games") and for such games we give lower bounds on the classical winning probabilities. One particular set of games with this property is described by Boyer [Boy04]. Their entangled value is 1 but their classical value goes to 0 as the number of players goes to infinity. The authors of [WHKN18] define a closely-related class of games called *noPREF games*. This set of games is equal to the set of perfect Schmidt games when $m = 2$ and the distribution on the inputs is uniform. In [WHKN18] it is shown that checking whether a game is in this class can be done in polynomial time. Furthermore, for *symmetric* t-player XOR games they show that a game has entangled value 1 if and only if it falls in this class of perfect Schmidt games. They also provide an explicit non-symmetric XOR game with entangled value 1 that is not in this class. We introduce a t-player MOD-m game called the *uniform angle game*, denoted $\text{UAG}_{t,m}$ (defined in Section 3.1, Definition 3.6) for which there is a perfect Schmidt strategy and show a lower bound on the classical winning probability.

4

Theorem 1.2. *Any t-player MOD-m game G with perfect Schmidt strategy satisfies $\omega(G) \geq \omega(\text{UAG}_{t,m})$. Furthermore we have $\beta(\text{UAG}_{t,m}) \geq t^{1-t}$.*

For $t = 3, m = 2$ (3-player XOR games) we have $\omega(\text{UAG}_{3,2}) = 3/4$. In Section 3 we provide bounds on $\omega(\text{UAG}_{t,m})$ for other values of t, m.

Let the inputs to a game come from a set $X = X_1 \times X_2 \times ... \times X_t$ where X_i is the set of inputs for the i-th player. We say a game is *total* when all elements of X have a non-zero probability of being asked (sometimes also called having *full support*), similar to total functions in the setting of communication complexity. On the other hand, we say that a game has a *promise* on the inputs when it is not total. For the class of perfect Schmidt games we show that total games are trivial.

Lemma 1.3. *When a t-player MOD-m game G with perfect Schmidt strategy is total then $\omega(G) = 1$.*

1.2 Free XOR games

In this subsection we identify two types of games, namely *free games* and *line games*, for which either the ratio of the entangled and classical biases is small, or the entangled bias itself is small. Thus these games will not be able to give a positive answer to Question 1.1. Free games are a general and natural class of games in which the players' questions are independently distributed. Line games appear to be less studied (see below for their definition), but turn out to be relevant in the context of parallel repetition (also see below). The main idea behind these results is that a large entangled bias implies that the games are in a sense far from random. This is quantified by the magnitude of certain norms of the game tensors. The particular norms of interest here are related to norms used in Gowers' celebrated hypergraph- and Fourier-analytic proofs of Szemerédi's Theorem. A crucial fact of these norms is that they are large if and only if there is "correlation with structure", the opposite of what one would expect from randomness. We show that this structure can be turned into good classical strategies, thus establishing a relationship between the entangled and classical biases.

Theorem 1.4 (Polynomial bias relation for free XOR games). *For integer $t \geq 2$ and any free t-player XOR game with entangled bias β, the classical bias is at least β^{2^t}.*

This result may be considered as an analogue of a well-known result on quantum query algorithms for total functions. It is shown in [BBC+01] that the bounded-error quantum and classical query complexities of total functions are polynomially related.

1.3 Line games

Line games are not free, but have a simple geometric structure. For a finite field $\mathbb{F}$ of characteristic at least t and positive integer n, a t-player line game is given by a map $\tau : \mathbb{F}^n \to \{0,1\}$. In the game, the referee independently samples two uniformly random points $x, y \in \mathbb{F}^n$ and sends the point $x + (i-1)y$ to the ith player. The players win the game if and only if the XOR of their answers equals $\tau(y)$. In other words, the players' questions correspond to consecutive points (or an arithmetic progression) on a random affine line through $\mathbb{F}^n$ and the winning criterion depends only on the direction of the line. Refer to this as a line games *over $\mathbb{F}^n$*.

A small example of a line game can be obtained from a slight modification of the three-player Magic Square game, which was analyzed in [IKP+08]. The line game is played over the plane $\mathbb{F}_3^2$

and the predicate is zero only on the horizontal lines (with $y \in \{(1,0),(2,0)\}$. In the Magic Square game, the referee restricts only to horizontal and vertical lines.[1]

Theorem 1.5. *For any $\varepsilon \in (0,1]$, integer $t \geq 2$ and finite field $\mathbb{F}$ of characteristic at least t, there exists a $\delta(\varepsilon,t,\mathbb{F}) \in (0,1]$ such that the following holds. For any positive integer n and any t-player line game over $\mathbb{F}^n$ with entangled bias ε, the classical bias is at least $\delta(\varepsilon,t,\mathbb{F})$.*

Note that in the above result, the value of the classical bias is independent of the dimension n of the vector space determining the players' question sets.

While it is not relevant to Question 1.1, the proof techniques used for Theorem 1.5 allow us to prove a parallel repetition theorem for a class of games that include line games. It is known that the value of free games and so-called anchored games decays exponentially under parallel repetition. Dinur et al. [DHVY16] identified a general criterion of multi-player games to behave like this, encompassing free and anchored games. They showed that it is sufficient for a certain graph that can be obtained from a game to be expanding, a well-known pseudorandom property that gives a measure of graph connectivity. Line games do not belong to this class, as their graphs are not even connected. However, we show that if a map $\tau : \mathbb{F}^n \to \{0,1\}^n$ is pseudorandom in a different sense, then a line game defined by τ has exponential decaying value under parallel repetition. More generally, we show that this is the case for a family of XOR games over an arbitrary finite abelian group Γ. These games are given by a positive integer m, a family of affine linear maps $\psi_0,\ldots,\psi_t : \Gamma^m \to \Gamma$ such that no two are multiples of each other, and a "game map" $\rho : \Gamma \to \{0,1\}$. In the game, the referee samples a uniform random element x from Γ^m and sends the group element $\psi_i(x)$ to the ith player. The winning criterion is given by $\rho(\psi_0(x))$. The relevant notion of pseudoranomness is quantified by the Gowers t-uniformity norm of the map $(-1)^\rho : x \mapsto (-1)^{\rho(x)}$, denoted $\|(-1)^\rho\|_{U^t}$.

Lemma 1.6. *Let m,t be positive integers and let Γ be a finite abelian group. Let $\psi_0,\ldots,\psi_t : \Gamma^m \to \Gamma$ be affine linear maps such that no two are multiples of each other and let $\rho : \Gamma \to \{0,1\}$. Let G be the t-player XOR game given by the system $\{\psi_0,\ldots,\psi_t,\rho\}$. Then, for every positive integer k,*

$$\omega(G^k) \leq \left(\frac{1 + \|(-1)^\rho\|_{U^t}}{2}\right)^k.$$

1.4 Unique games

We know that the answer to Question 1.1 is negative in the two-player case, but we can generalize the question by dropping the XOR restriction. The set of XOR games is part of a larger class of games called unique games for which we investigate the relation between classical and entangled values. A two-player nonlocal game is a unique game if for every pair of questions, for every possible answer of the first player there is exactly one answer of the second player that lets them win, and vice versa. Stated differently, for every question there is a matching between the answers of the two players such that only the matching pairs of answers let the player win.

The Unique Games Conjecture (UGC) of Khot [Kho02] states that for any $\varepsilon,\delta > 0$, for any $k > k(\varepsilon,\delta)$, it is NP-hard to distinguish instances of unique games with winning probability at least $1 - \varepsilon$ from those with winning probability at most δ, where k is the number of possible answers. This conjecture has important consequences because it implies several hardness of approximation

---

[1]Though this is not the typical description of the game, it is easily seen to be equivalent.

results. For example, for the Max-Cut problem, Khot et al. [KKMO07] showed that the UGC implies that obtaining an approximation ratio better than ≈ 0.878 is NP-hard. Other results include inapproximability for Vertex Cover [KR08] and graph coloring problems [DMR09].

Our results relate the quantum and classical winning probabilities in the regime of near-perfect play and are based on a result in [CMM06].

Theorem 1.7. *Let $\varepsilon \geq 0$. There is an efficient algorithm that, given any two-player unique game with entangled value $1 - \varepsilon$, outputs a classical strategy with winning probability at least $1 - C\sqrt{\varepsilon \log k}$, where C is a constant independent of the game.*

Note that for $\varepsilon = 0$ this means a perfect quantum strategy implies a perfect classical strategy. Furthermore, the above result only beats a trivial strategy when $\varepsilon = \mathcal{O}(1/\log k)$.

Work in a similar direction includes [KRT08]. They show that entangled version of the UGC is false, by providing an efficient algorithm that gives an explicit quantum strategy with winning probability at least $1 - 6\varepsilon$ when the true entangled value is $1 - \varepsilon$. In the classical case, [CMM06] gives an algorithm that outputs a classical strategy with winning probability $1 - \mathcal{O}(\sqrt{\varepsilon \log k})$ when the true *classical value* is $1 - \varepsilon$. We extend this result by showing that this classical strategy also does the job when, not the classical, but the entangled value is $1 - \varepsilon$.

2 Techniques

This section provides an overview of the proof techniques that we employed. We give sketches of the main ideas which are worked out in full detail in later sections.

2.1 Reduction to angle games.

To prove Theorem 1.2 we introduce a new set of t-player MOD-m games that we call *angle games*. We define a particular angle game called the *uniform angle game*, denoted by $\mathrm{UAG}_{t,m}$ and show that it is the hardest of these games. In an angle game, players receive complex phases $e^{i\phi}$ (angles) satisfying a promise, and the winning answer depends only on the product of the inputs $e^{i\phi_1} \cdot e^{i\phi_2} \cdots e^{i\phi_t}$. We prove the theorem by extracting from any perfect Schmidt strategy a set of complex phases that satisfy such a promise, and thereby reducing any such game to the $\mathrm{UAG}_{t,m}$ game. Let us sketch how this is accomplished. Assume that a perfect Schmidt strategy exists, and let $\{P_1^{(j,x_j)}, ..., P_m^{(j,x_j)}\}$ be the projective measurement done by player j on input x_j so that $P_i^{(j,x_j)}$ corresponds to output i. Now define unitaries $U^{(j,x_j)} = \sum_i \omega_m^i P_i^{(j,x_j)}$, where $\omega_m = e^{2\pi i/m}$ is an m-th root of unity. Since the strategy is perfect we have for every input $(x_1, ..., x_t)$ that

$$\omega_m^{M(x_1, ..., x_t)} = \langle \psi | U^{(1,x_1)} \otimes U^{(2,x_2)} \otimes ... \otimes U^{(t,x_t)} | \psi \rangle.$$

Using the definition of a Schmidt state, we show that this equality implies that these unitaries must be of a simple form and their entries satisfy the promise of an angle game. We prove Theorem 1.2 and Lemma 1.3 in Section 3, where we also provide classical strategies for the uniform angle game and show that these are tight in the case of 3-player XOR games.

2.2 Norming hypergraphs and quasirandomness.

Our main tool for proving Theorem 1.4 is a relation between the entangled and classical biases and a norm on the set of game tensors. For t-tensors, this norm is given in terms of a certain t-partite t-uniform hypergraph H. Recall that such a hypergraph consists of t finite and pairwise disjoint vertex sets $V_1, \ldots, V_t$ and a collection of t-tuples $E(H) \subseteq V_1 \times \cdots \times V_t$, referred to as the edge set of H. For a t-tensor $T \in \mathbb{R}^{n_1 \times \cdots \times n_t}$, the norm has the following form:

$$\|T\|_H = \left(\mathbf{E}_{\phi_i : V_i \to [n_i]} \left[\prod_{(v_1, \ldots, v_t) \in E(H)} T\big(\phi_1(v_1), \ldots, \phi_t(v_t)\big) \right] \right)^{\frac{1}{|E(H)|}}, \tag{2}$$

where the expectation taken with respect to the uniform distribution over all t-tuples of mappings ϕ_i from V_i to $[n_i]$. Expressions such as (2) play an important role in the context of graph homomorphisms [BCL$^+$06]. If T is the adjacency matrix of a bipartite graph with left and right node sets $[n_1]$ and $[n_2]$ respectively, then each product in (2) is 1 if and only if the maps ϕ_1 and ϕ_2 preserve edges.

Criteria for H under which (2) defines a norm or a semi-norm were determined by Hatami [Hat10, Hat09] and Conlon and Lee [CL17]. Famous examples of graph norms include the Schatten-p norms for even $p \geq 4$ (in which case H is a p-cycle) and a well-known family of hypergraph norms are the Gowers octahedral norms. The latter were introduced for the purpose of quantifying a notion of quasirandomness of hypergraphs as an important part of Gowers' graph-theoretic proof of Szemerédi's theorem on arithmetic progressions. Having large Gowers norm turns out to imply *correlation with structure*, as opposed to quasirandomness. This is true also for the norm relevant for our setting. In particular, it turns out that the structure with which a game tensor correlates can be turned into a classical strategy for the game. As such, a large norm of the game tensor implies a large classical bias of the game itself. At the same time, we show that the entangled bias is bounded from above by the norm of the game tensor, provided the game is free. Putting these observations together gives the proof of Theorem 1.4, which we give in Section 4.

The particular hypergraph norm relevant in our setting was introduced in [CHPS12] and can be obtained recursively as follows. Starting with a t-partite t-uniform hypergraph H with vertex set $V_1 \cup \cdots \cup V_t$, write $\mathrm{db}_i(H)$ for the t-partite t-uniform hypergraph obtained by making two vertex-disjoint copies of H and gluing them together so that the vertices in the two copies of V_i are identified. We obtain our hypergraph by starting with a single edge $e = (v_1, \ldots, v_t)$ (and vertex sets of size 1), and applying this operation to all parts, forming the hypergraph $\mathrm{db}_1(\mathrm{db}_2(\ldots \mathrm{db}_t(e)))$ with vertex sets of size 2^{t-1} and 2^t edges. The fact that this hypergraph defines a norm via (2) was proved in [CL17].

2.3 Line games and Gowers uniformity norms.

The proof of Theorem 1.5 is based on two fundamental results from additive combinatorics: the generalized von Neuman inequality and the Gowers Inverse Theorem. The former easily shows that the classical bias of a line game is bounded from above by the Gowers t-uniformity norm of the game map. We show that in fact the same upper bound holds for the entangled bias as well. A large entangled bias thus implies a large uniformity norm for the game map. Analogous to the above-mentioned octahedral norms for tensors, uniformity norms were introduced to quantify a notion of pseudorandomness for bounded maps over abelian groups as an important step in Gowers' other proof of Szemerédi's Theorem, based on higher-order Fourier analysis. The

highly non-trivial Gowers Inverse Theorem of Tao and Ziegler [TZ12] establishes that high uniformity norm again implies correlation with structure. Although structure in this context means something quite different than for tensors, we show that it still implies a lower bound on the classical bias. The above observations together prove Theorem 1.5, details of which can be found in Section 5.

2.4 Semidefinite programming relaxation.

The proof of Theorem 1.7 is a small modification of a proof in [CMM06]. They consider a semidefinite programming (SDP) relaxation of the optimization problem for the classical value and then give two algorithms for rounding the result of the SDP to a classical strategy. In the SDP relaxation the objective is to optimize $\mathbb{E}_{x,y} \sum_{i=1}^{k} \langle u_i^{(x)} \mid v_{\pi_{xy}(i)}^{(y)} \rangle$ where $u_i^{(x)}, v_j^{(y)} \in \mathbb{R}^d$ are vectors corresponding to questions x, y and answers i, j. Furthermore, π_{xy} is the matching of correct answers on questions x, y. A classical strategy would correspond to the case where the vectors are integers instead, such that for each x exactly one $u_i^{(x)}$ is equal to 1 and all other $u_i^{(x)}$ are equal to zero and similar for the $v_j^{(y)}$. A quantum strategy also gives rise to a set of vectors, but satisfying different constraints [KRT08]. One of the constraints of the SDP considered in [CMM06] is $0 \leq \langle u_i \mid v_{\pi_{xy}(i)} \rangle \leq |u_i|^2$ which is valid for classical strategies, but in general not for quantum strategies. For our proof, we consider the same SDP but with this constraint dropped. In that case it is also a relaxation for the entangled case and with a few changes one of the rounding algorithms in [CMM06] is also valid when the constraint is dropped. Note that the result only beats a trivial strategy when $\varepsilon = \mathcal{O}(1/\log k)$ whereas the other rounding algorithm in [CMM06] is non-trivial for any ε. However this other algorithm is more dependent on the extra constraint and it is not clear if it can be dropped there as well.

To get some intuition for the rounding algorithm, we sketch a solution for $\varepsilon = 0$ here. In this case one can show that for each question pair x, y the set of vectors $|u_i^{(x)}\rangle$ ($i = 1, ..., k$) known by the first player is the same set of vectors as the set $|v_i^{(y)}\rangle$ ($i = 1, ..., k$) known to the second player. In particular, the vector $|u_i^{(x)}\rangle$ is the same as the matching vector $|v_{\pi_{xy}(i)}^{(y)}\rangle$ of the other player. Using shared randomness they can sample a random vector $|g\rangle$ and compute the overlaps $\xi_i^{(x)} = \langle g|u_i^{(x)}\rangle$ and $\xi_i^{(y)} = \langle g|v_i^{(y)}\rangle$ respectively. As they have the same vectors, the players will have the same values for answers in the matching: $\xi_i^{(x)} = \xi_{\pi_{xy}(i)}^{(y)}$. Now both players simply output the answer i for which $|\xi_i^{(x)}|$ (and $|\xi_i^{(y)}|$ for the other player) has the largest value. With probability one this will yield correct answers. For $\varepsilon > 0$ the sets of vectors will not be exactly equal and therefore the values $\xi_i^{(x)}, \xi_{\pi_{xy}(i)}^{(y)}$ will be close but not exactly equal. The discrepancy in these values will be bigger for vectors $|u_i^{(x)}\rangle$ with a small norm. In Section 6 we provide the rounding algorithm in full detail and show how this issue is solved.

3 Perfect quantum Schmidt strategies for MOD games

We start by defining a set of games that turn out to characterize the games we are interested in.

Definition 3.1 (Angle game). *Define an* angle game *as a t-player MOD-m game where player j gets an angle $e^{i\phi_j}$ as input, with the promise that $e^{i\phi_1} \cdot ... \cdot e^{i\phi_t} = \omega_m^{M(\phi_1,...,\phi_t)}$ where $M(\phi_1,...,\phi_t) \in \{0,1,...,m-1\}$ and $\omega_m = e^{i2\pi/m}$. The players win if and only if the sum of their outputs modulo m is equal to $M(\phi_1,...,\phi_t)$.*

Note that an angle game is completely defined by t, m and a probability distribution over angle tuples. Furthermore, t-player Boyer games [Boy04] with parameters (D,M) are angle games where the (discrete) probability distribution is uniform over all angles of the form $e^{i2\pi x/(MD)}$ with $x = 0,1,...,D-1$ whose product is an M-th root of unity. The promise $\sum x_j \equiv 0 \mod D$ as stated in the Boyer games translates to $\prod e^{i2\pi x_j/(MD)} = \omega_M^l$ in the angle game, where $l = \sum x_j/D$.

Lemma 3.2. *Any angle game has entangled value 1 which can be obtained using a shared GHZ state.*

Proof. Consider the following quantum strategy using a GHZ state of dimension m. Every player applies the local diagonal unitary $U_{jj} = e^{ij\cdot\phi}$ on input $e^{i\phi}$. Then every player applies an inverse Fourier transform $F_{ij}^\dagger = \frac{1}{\sqrt{m}}\omega_m^{-i\cdot j}$ and then measures in the computational basis and outputs the result. With probability 1 the sum of their outputs is equal to l. $\qquad\square$

Lemma 3.3. *Any t-player MOD-m game with a perfect Schmidt strategy can be reduced to an angle game.*

Proof. Let $\{P_1^{(j,x_j)},...,P_m^{(j,x_j)}\}$ be the projective measurement done by player j on input x_j so that $P_i^{(j,x_j)}$ corresponds to output i. This set of projectors is pairwise orthogonal and sums to identity. Now define unitaries $U^{(j,x_j)} = \sum_i \omega_m^i P_i^{(j,x_j)}$. Since the strategy is perfect we have for every input $(x_1,...,x_t)$ that

$$
\begin{aligned}
\omega_m^{M(x_1,...,x_t)} &= \langle\psi|U^{(1,x_1)} \otimes U^{(2,x_2)} \otimes ... \otimes U^{(t,x_t)}|\psi\rangle \\
&= \sum_{i,j} c_i c_j \langle e_i^{(1)}|U^{(1,x_1)}|e_j^{(1)}\rangle \langle e_i^{(2)}|U^{(2,x_2)}|e_j^{(2)}\rangle \cdots \langle e_i^{(t)}|U^{(t,x_t)}|e_j^{(t)}\rangle \\
&= \sum_{i,j} c_i c_j U_{ij}^{(1,x_1)} U_{ij}^{(2,x_2)} \cdots U_{ij}^{(t,x_t)}.
\end{aligned}
\tag{3}
$$

where we entered the definition of a Schmidt state as given in Section 1.1 and we shortened $U_{ij}^{(t,x_t)} := \langle e_i^{(t)}|U^{(t,x_t)}|e_j^{(t)}\rangle$. Now apply Cauchy-Schwarz to obtain

$$
\left|\sum_{i,j} c_i c_j U_{ij}^{(1,x_1)} U_{ij}^{(2,x_2)} \cdots U_{ij}^{(t,x_t)}\right| \le \left(\sum_{i,j} c_i^2 \left|U_{ij}^{(1,x_1)}\right|^2\right)^{\frac{1}{2}} \left(\sum_{i,j} c_j^2 \left|U_{ij}^{(2,x_2)} \cdots U_{ij}^{(t,x_t)}\right|^2\right)^{\frac{1}{2}}
$$

$$
\le 1 \left(\sum_{i,j} c_j^2 \left|U_{ij}^{(2,x_2)}\right|^2\right)^{\frac{1}{2}} = 1.
$$

Here we used that the $U^{(j,x_j)}$ are unitary and therefore their rows and columns are unit vectors. When $|\langle a,b\rangle| = \|a\| \cdot \|b\|$ then we have $|a\rangle = \lambda|b\rangle$ for some $\lambda \in \mathbb{C}$. Keeping in mind the complex

conjugation in the inner product, there is a λ such that

$$\lambda c_i \overline{U_{ij}^{(1,x_1)}} = c_j U_{ij}^{(2,x_2)} \cdots U_{ij}^{(t,x_t)}$$

where $\bar{z}$ denote the complex conjugate of z. Plugging this into (3) gives $\lambda = \omega_m^{M(x_1,\dots,x_t)}$. From the above equation it follows that when $U_{ij}^{(t,x_t)}$ is non-zero for $t = 1$ then it is non-zero for every t. Instead of the first player we could have used any other player in the above derivation, so if any $U_{ij}^{(t,x_t)}$ is non-zero for some t then it is non-zero for all t. Let i,j be such that $U_{ij}^{(t,x_t)} \neq 0$, then we can take the argument of the above equation to find

$$\frac{2\pi}{m} M(x_1,\dots,x_t) = \arg(U_{ij}^{(1,x_1)}) + \arg(U_{ij}^{(2,x_2)}) + \cdots + \arg(U_{ij}^{(t,x_t)}).$$

On any input $(x_1,\dots,x_t)$, the players simply looks at the first non-zero element of their matrix $U^{(t,x_t)}$ and look at the argument $\phi_t := \arg(U_{ij}^{(t,x_t)})$. These angles have the property that $e^{i\phi_1} \cdot \dots \cdot e^{i\phi_t} = \omega_m^{M(x_1,\dots,x_t)}$. This reduces the game to an angle game. $\qquad \square$

Definition 3.4 (Connected inputs). *For any game, define a graph where every input (a t-tuple) with non-zero probability of being asked is a vertex. Two inputs are connected via an edge if they differ on only one player and agree on the other $t-1$ coordinates. We say the game has* connected inputs *if this graph is connected.*

Dinur et al. [DHVY16] consider the same graph which they call the $(t-1)$-connection graph of the game. Total games and free games have connected inputs. Games that do *not* have connected imputs typically have a *promise* on the inputs.

Lemma 3.5. *Any angle game with connected inputs has classical value 1.*

Proof. Fix an input $(e^{i\alpha_1},\dots,e^{i\alpha_t})$. Now define $\beta_1(e^{i\phi_1}) = e^{i\phi_1}e^{i\alpha_2}e^{i\alpha_3}\cdots e^{i\alpha_t}$ and $\beta_j(e^{i\phi_j}) = e^{i\phi_j}e^{-i\alpha_j}$ for $j \geq 2$. The product of the angles is left unchanged under these maps, $\beta_1(e^{i\phi_1})\beta_2(e^{i\phi_2})\cdots\beta_t(e^{i\phi_t}) = e^{i\phi_1}\cdots e^{i\phi_t}$. We claim that every input is mapped to an m-th root of unity, i.e. $\beta_j(e^{i\phi_j}) = \omega_m^{l_j}$. Therefore player j can output l_j and $\sum_j l_j = M(e^{i\phi_1},\dots,e^{i\phi_t})$ thus winning the game with probability 1.

First note that on the fixed input we have $\beta_1(e^{i\alpha_1}) = \omega_m^{M(e^{i\alpha_1},\dots,e^{i\alpha_t})}$ and $\beta_j(e^{i\alpha_j}) = 1$ for $j \geq 2$. so the claim holds on the fixed input. By the strong connectivity we can obtain another input to the game by changing only the input for a single player. We now show that when the claim holds for one input then it also holds when only one player's value is changed. Using these single-player edits we can eventually reach all inputs. Let $(e^{i\phi_1},\dots,e^{i\phi_j},\dots,e^{i\phi_t})$ and $(e^{i\phi_1},\dots,e^{i\phi_j'},\dots,e^{i\phi_t})$ be two inputs that only differ for player j. Now assume that the claim holds for the first input. We then have

$$\omega_m^{M(e^{i\phi_1},\dots,e^{i\phi_j'},\dots,e^{i\phi_t})} = \beta_1(e^{i\phi_1})\cdots\beta_j(e^{i\phi_j'})\cdots\beta_t(e^{i\phi_t})$$

$$= \beta_1(e^{i\phi_1})\cdots\beta_j(e^{i\phi_j})\cdots\beta_t(e^{i\phi_t}) \cdot \beta_j(e^{i\phi_j})^{-1}\beta_j(e^{i\phi_j'})$$

$$= \omega_m^{M(e^{i\phi_1},\dots,e^{i\phi_j},\dots,e^{i\phi_t})}\omega_m^{-l_j}\beta_j(e^{i\phi_j'})$$

from which it follows that $\beta_j(e^{i\phi_j'}) = \omega_m^{l_j'}$ for some l_j'. $\qquad \square$

Lemma 1.3 follows directly from Lemma 3.3 and Lemma 3.5.

3.1 Classical strategies for angle games

Having characterized our class of games as angle games we proceed by presenting classical strategies for these games. Our aim is to provide strategies that work for any probability distribution on the set of inputs. In this section it will be convenient to write the angles as $e^{i\frac{2\pi}{m}\phi}$ so that ϕ runs from 0 to m instead of 0 to 2π.

Definition 3.6. *Define the probability distribution π_{U_t} on the set*

$$U_t = \left\{ (e^{i\frac{2\pi}{m}\phi_1},...,e^{i\frac{2\pi}{m}\phi_t}) \mid \phi_j \in [0,1) \ , \ e^{i\frac{2\pi}{m}\phi_1}\cdots e^{i\frac{2\pi}{m}\phi_t} = \omega_m^l, \ l \in \{0,1,...,m-1\} \right\}$$

as follows: for $1 \le j \le t-1$, draw ϕ_j independently uniformly at random from $[0,1)$. Then define ϕ_t as the unique number in $[0,1)$ that makes the product $e^{i\frac{2\pi}{m}\phi_1}\cdots e^{i\frac{2\pi}{m}\phi_t}$ an m-th root of unity. We define the uniform angle game, denoted $\mathrm{UAG}_{t,m}$, as a t-player MOD-m angle game (Definition 3.1) where the input distribution is π_{U_t}.

As stated before, in a t-player Boyer game [Boy04] with parameters (D,M) the (discrete) probability distrubtion is uniform over all angles of the form $e^{i2\pi x/(MD)}$ with $x = 0,1,...,D-1$ whose product is an M-th root of unity. This is similar to the π_{U_t} distribution but where the angles $\phi_j \in [0,1)$ are now discrete $\phi_j \in \{0,\frac{1}{D},\frac{2}{D},...,\frac{D-1}{D}\}$.

The distribution π_{U_t} is the hardest distribution as captured by the following claim.

Claim 3.7. *Let G be a t-player MOD-m angle game with input distribution π_G. Then $\omega(G) \ge \omega(\mathrm{UAG}_{t,m})$.*

Proof. Assume the players get an input $(e^{i\frac{2\pi}{m}\phi_1},...,e^{i\frac{2\pi}{m}\phi_t})$ with $e^{i\frac{2\pi}{m}\phi_1}\cdots e^{i\frac{2\pi}{m}\phi_t} = \omega_m^a$ from π_G. Using shared randomness, draw $t-1$ independent random angles $e^{i\alpha_1},...,e^{i\alpha_{t-1}}$ where each α_i is uniform on $[0,2\pi)$. Multiply the input $e^{i\frac{2\pi}{m}\phi_j}$ of player j by $e^{i\alpha_j}$ for $j \le t-1$ and multiply $e^{i\frac{2\pi}{m}\phi_t}$ by $e^{-i(\alpha_1+...+\alpha_{t-1})}$ to preserve the product. The resulting distribution is uniform on the set

$$\left\{ (e^{i\frac{2\pi}{m}\phi_1},...,e^{i\frac{2\pi}{m}\phi_t}) \mid \phi_j \in [0,m) \ , \ e^{i\frac{2\pi}{m}\phi_1}\cdots e^{i\frac{2\pi}{m}\phi_t} = \omega_m^a \right\}$$

One can always write $\phi_j = l_j + \varphi_j$ with $l_j \in \{0,1,...,m-1\}$ and $0 \le \varphi_j < 1$. Note that $(e^{i\frac{2\pi}{m}\varphi_1},...,e^{i\frac{2\pi}{m}\varphi_t})$ is distributed according to π_{U_t} so the players can play a strategy for $\mathrm{UAG}_{t,m}$ to obtain answers $(a_1,...,a_t)$. On input $e^{i\frac{2\pi}{m}(l_j+\varphi_j)}$, player j outputs $a_j + l_j$. They are correct if and only if the answers a_j are correct for $\mathrm{UAG}_{t,m}$ on input $(e^{i\frac{2\pi}{m}\varphi_1},...,e^{i\frac{2\pi}{m}\varphi_t})$. This proves the claim. $\square$

Claim 3.8. *The uniform angle game satisfies $\omega(\mathrm{UAG}_{t,m}) \ge \frac{1}{m} + \frac{m-1}{m}t^{1-t}$.*

Proof. The players get inputs from U_t as in Definition 3.6. On input ϕ_i, player i computes $x_i = \lfloor t\,\phi_i \rfloor$, so that $x_i \in \{0,...,t-1\}$. We then have $\frac{1}{t}\sum_{i=1}^t x_i \le \sum_{i=1}^t \phi_i < 1 + \frac{1}{t}\sum_{i=1}^t x_i$ and since the correct answer l is given by $l = \sum_{i=1}^t \phi_i$, we see that l is uniquely determined by the sum of x_i. Now using shared randomness the players sample $t-1$ random numbers from $\{0,...,t-1\}$. With probability $\frac{1}{t^{t-1}}$ these numbers are exactly equal to $x_1,...,x_{t-1}$. The last player assumes that the random numbers are indeed $x_1,...,x_{t-1}$ and outputs the correct l. The other players output 0 if their x_i matches the random sample and output a random number otherwise. This yields a bias of $\frac{1}{t^{t-1}}$, or winning probability of $\frac{1}{m} + \frac{m-1}{m}t^{1-t}$. $\square$
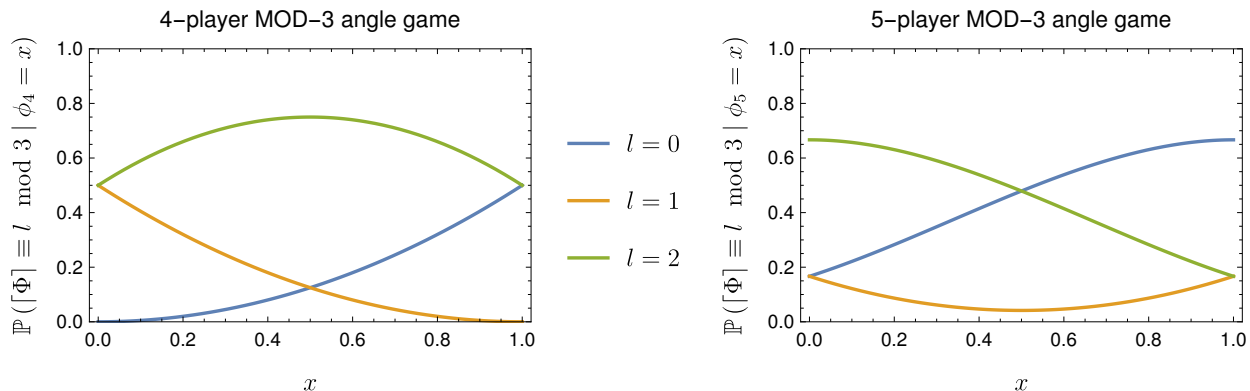
Figure 1: Probability that the correct answer of the angle game is l, conditioned on the last player receiving input x, as defined in Equation (4).

Theorem 1.2 follows from Lemma 3.3, Claim 3.7 and Claim 3.8.

We proceed by describing a strategy for $\text{UAG}_{t,m}$ that improves on the bound of Claim 3.8. Let $(\phi_1, ..., \phi_t)$ be drawn from π_{U_t}. Define $\Phi = \phi_1 + \phi_2 + ... + \phi_{t-1}$ then by definition of π_{U_t}, Φ is the sum of $t - 1$ independent uniform $[0,1)$ variables. Furthermore, the last input is $\phi_t = \lceil \Phi \rceil - \Phi$ and the correct answer l is defined by $l \equiv \lceil \Phi \rceil \mod m$. The distribution of Φ is known as the Irwin-Hall distribution [Irw27, Hal27]:

$$\mathbb{P}(\Phi \leq x) = \frac{1}{(t-1)!} \sum_{j=0}^{\lfloor x \rfloor} (-1)^j \binom{t-1}{j} (x - j)^{t-1}.$$

We consider a set of strategies that we call the *semi-trivial strategies*, in which the first $t - 1$ players always output 0. The last player then plays optimally when given the input x. We conjecture that this strategy is optimal for the uniform angle game. In the semi-trivial strategy, the last player chooses the l that maximizes

$$\mathbb{P}(\text{correct answer is } l \mid \text{last input is } x) = \mathbb{P}(\lceil \Phi \rceil \equiv l \mod m \mid \phi_t = x). \tag{4}$$

This probability is plotted as a function of x in Figure 1. The figure shows that for a 4-player MOD-3 game (left plot) the optimal choice for the last player is to ignore the input x and always output 2. Interestingly, this is the pattern we observe for any value of m for any even number of players (checked up to $m = 10, t = 10$): the value of l for which this probability is maximal is independent of x. This means that these strategies (for t even) are locally optimal in the sense that changing any single player's strategy will not improve the winning probability. The right plot of the figure shows that for 5 players the optimal answer depends on whether $x \leq \frac{1}{2}$ or $x > \frac{1}{2}$. This, too, is seems to be a general pattern for any m and any odd number of players (checked up to $m = 10, t = 11$). The winning probabilities provided by these semi-trivial strategies are given in Table 1 for $m = 2$ and $m = 3$. The semi-trivial strategies are lower bounds for the winning probability of *every* t-player angle game. We can find upper bounds by finding upper bounds for particular angle games. The table provides some upper bounds obtained from brute-force searching through all strategies for Boyer games.

13

t (# players)	2	3	4	5	6	7	8	9
MOD 2								
lower bound	1	3/4	2/3	29/48	17/30	781/1440	166/315	8341/16128
	1	0.75	0.6667	0.6042	0.5667	0.5424	0.5270	0.5172
upper bound	1	3/4	43/64	155/256	583/1024	35/64	273/512	1056/1048
	1	0.75	0.6719	0.6055	0.5693	0.5469	0.5332	0.5200
MOD 3								
lower bound	1	3/4	2/3	115/192	11/20	785/1536	403/840	260451/573440
	1	0.75	0.6667	0.5990	0.5500	0.5111	0.4798	0.4542
upper bound	1	61/81	163/243	17/27	47/81	131/243	41/81	349/729
	1	0.7531	0.6708	0.6296	0.5802	0.5391	0.5062	0.4787

Table 1: Lower and upper bounds for the winning probabilities of the $\text{UAG}_{t,m}$ games (Definition 3.6). The lower bounds are the semi-trivial strategies described in the text. The upper bounds were found by iterating through all strategies for t-player Boyer games with values of D up to $D = 9$ when computation time allowed it.

For the case of 3-player XOR games, the upper bound $\omega(\text{UAG}_{3,2}) = 3/4$ is tight. For the 4-player case we have $\omega(\text{UAG}_{4,2}) \geq 2/3$ and it seems that searching through Boyer games gives increasingly better bounds, approaching $2/3$ as the input size D is increased. However, one can show that for any finite D the lower bound of $2/3$ will not be reached, because when $D = 2^m$ there is a strategy that achieves a winning probability of $\frac{2}{3} + \frac{4^{-m}}{3}$. The corresponding strategy is that the first 3 players output 0 and the fourth player outputs 1 when their input is 0 or 1 and outputs 0 when their input is $2, ..., D - 1$. This strategy is optimal for $D = 2, 4, 8$. One can show that for values of D that are not a power of 2, i.e. $D = D'2^m$ the game reduces to one with $D = 2^m$.

4 Free XOR games

In this section we will define free XOR games and give the definition of hypergraph norms (only for real-valued functions on discrete domains). For more details on hypergraph norms we refer to [Hat09]. We will then relate the hypergraph norm, with respect to a certain hypergraph, of the game tensor to the quantum bias of free XOR games. Our main tool is a Cauchy-Schwarz type of inequality for operators, that is why we will state it here.

Proposition 4.1. *Let* $A_i, B_i \in \text{End}(\mathbb{C}^n)$ *for* $i = 1, \ldots, k$. *Then*

$$\| \sum_{i \in [k]} A_i B_i \| \leq \| \sum_{i \in [k]} A_i A_i^* \|^{1/2} \| \sum_{i \in [k]} B_i^* B_i \|^{1/2},$$

where all the norms are operator norms.

Proof. Write

$$A = \begin{bmatrix} A_1 & A_2 & \cdots & A_k \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} B_1 & 0 & \cdots & 0 \\ B_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ B_k & 0 & \cdots & 0 \end{bmatrix}.$$

14

Then we use the fact that

$$AB = \begin{bmatrix} \sum_{i \in [k]} A_i B_i & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

together with the properties

$$\|C \otimes E_{ij}\| = \|C\|,$$
$$\|C\|^2 = \|C^*C\| = \|CC^*\|,$$

to conclude

$$\| \sum_{i \in [k]} A_i B_i \| = \| \sum_{i \in [k]} A_i B_i \otimes E_{11} \| = \|AB\| \le \|A\| \|B\|$$
$$= \|AA^*\|^{1/2} \|B^*B\|^{1/2} = \| \sum_{i \in [k]} A_i A_i^* \|^{1/2} \| \sum_{i \in [k]} B_i^* B_i \|^{1/2}.$$

$\square$

A t-player free XOR game G is given by finite non-empty sets $X_1, \ldots, X_t$, a product distribution over $X := X_1 \times \cdots \times X_t$ and a game tensor

$$T \colon X \to \{\pm 1\}. \tag{5}$$

The classical bias of the free XOR game G, which we denote by $\beta(G)$ is given by

$$\beta(G) := \max_{a_i \colon X_i \to \{\pm 1\}} \big| \mathop{\mathbb{E}}_{(x_1, \ldots, x_t) \in X} T(x_1, \ldots, x_t) \prod_{i=1}^t a_i(x_i) \big|.$$

The quantum bias of the free XOR game G, which we denote by $\beta^*(G)$ is given by the expression

$$\beta^*(G) := \max_{N \in \mathbb{N}, A_i \colon X_i \to \mathrm{Obs}^{\pm}(\mathbb{C}^N)} \big\| \mathop{\mathbb{E}}_{(x_1, \ldots, x_t) \in X} T(x_1, \ldots, x_t) \prod_{i=1}^t A_i(x_i) \big\|_{\mathrm{op}}, \tag{6}$$

the maximization is taken over $\pm$-observable valued functions A_i such that $[A_i, A_j] = 0$ for $i \ne j$, this corresponds to a quantum strategy of the players. The expectation is taken over the given distribution.

Before we go into the detail of the proof of Theorem 1.4 for any number of players, we first sketch the core idea of the proof for *two* players, for which we do not yet need to resort to hypergraphs. For a two-player game G with game tensor T, the commuting-operator strategies A, B yield a bias of

$$\eta = \big\| \mathop{\mathbb{E}}_{(x,y) \in X \times Y} T(x,y) A(x) B(y) \big\|.$$

where the norm is the operator norm. Using Proposition 4.1 we peel off the operator $B(y)$

$$\eta = \left\| \mathop{\mathbb{E}}_{y\in Y} \left(\mathop{\mathbb{E}}_{x\in X} T(x,y)A(x) \right) B(y) \right\|. \qquad \text{(independent questions)}$$

$$\leq \left\| \mathop{\mathbb{E}}_{y\in Y} \left(\mathop{\mathbb{E}}_{x\in X} T(x,y)A(x) \right) \left(\mathop{\mathbb{E}}_{x'\in X} T(x',y)A(x') \right)^* \right\|^{1/2} \left\| \mathop{\mathbb{E}}_{y\in Y} B(y)B(y)^* \right\|^{1/2}$$

$$\leq \left\| \mathop{\mathbb{E}}_{y\in Y} \mathop{\mathbb{E}}_{x,x'\in X} T(x,y)T(x',y)A(x)A(x')^* \right\|^{1/2}. \qquad \text{(using } \|B(y)\|\leq 1\text{)}$$

Now we apply the inequality again on the sum over (x,x') to get rid of the A operator.

$$\eta \leq \left\| \mathop{\mathbb{E}}_{x,x'\in X} \left(\mathop{\mathbb{E}}_{y\in Y} T(x,y)T(x',y) \right) A(x)A(x')^* \right\|^{1/2}$$

$$\leq \left| \mathop{\mathbb{E}}_{x,x'} \left(\mathop{\mathbb{E}}_{y} T(x,y)T(x',y) \right) \left(\mathop{\mathbb{E}}_{y'} T(x,y')T(x',y') \right) \right|^{\frac{1}{4}} \left\| \mathop{\mathbb{E}}_{x,x'} \left(A(x)A(x')^* \right) \left(A(x)A(x')^* \right)^* \right\|^{\frac{1}{4}}$$

$$\leq \left| \mathop{\mathbb{E}}_{x,x'\in X} \mathop{\mathbb{E}}_{y,y'\in Y} T(x,y)T(x',y)T(x,y')T(x',y') \right|^{1/4}. \qquad \text{(using } \|A(x)\|\leq 1\text{)}$$

We proceed by rewriting the last expression

$$\eta^4 \leq \left| \mathop{\mathbb{E}}_{(x',y')\in X\times Y} T(x',y') \mathop{\mathbb{E}}_{(x,y)\in X\times Y} T(x,y)T(x',y)T(x,y') \right|$$

$$\leq \mathop{\mathbb{E}}_{(x',y')\in X\times Y} \left| \mathop{\mathbb{E}}_{(x,y)\in X\times Y} T(x,y)T(x',y)T(x,y') \right|. \qquad \text{(triangle inequality)}$$

By the pigeonhole principle there must be choices of x', y' such that

$$\eta^4 \leq \left| \mathop{\mathbb{E}}_{(x,y)\in X\times Y} T(x,y)T(x',y)T(x,y') \right|,$$

which is the expression for the bias of the classical strategies $a(x) = T(x,y')$ and $b(y) = T(x',y)$, proving Theorem 1.4 for $t=2$ players. For $t\geq 3$ we can apply the same idea, peeling off the operators one by one, but the final expression is more involved. We will now develop the techniques to deal with this. In particular, we need the notion of hypergraph norms. For our purposes, we only consider t-uniform hypergraphs which are also t-partite.

Definition 4.2. *For $t\geq 2$, let $V_1,\ldots,V_t$ be finite non-empty sets and $V := V_1 \times \cdots \times V_t$. Given a subset $E \subset V$, we say that the pair $H = (V_1 \cup \cdots \cup V_t, E)$ is a t-partite t-uniform hypergraph with vertex set $V_1 \cup \cdots \cup V_t$ and edge set E.*

Definition 4.3. *Let $t\geq 2$ and $X_1,\ldots,X_t$ be finite non-empty sets and suppose a product distribution on $X := X_1 \times \cdots \times X_t$ is given to us. Let $T\colon X \to \mathbb{R}$ be a function and $H = (V_1 \cup \cdots \cup V_t, E)$ be a t-partite t-uniform hypergraph. We define a non-negative function $\|\cdot\|_H$ on the function T by*

$$\|T\|_H := \left| \mathop{\mathbb{E}}_{\phi_i\colon V_i\to X_i} \prod_{(v_1,\ldots,v_t)\in E} T(\phi_1(v_1),\ldots,\phi_t(v_t)) \right|^{\frac{1}{|E|}}. \qquad (7)$$

The expectation is taken with respect to the following distribution: a particular map $\phi_i\colon V_i \to X_i$ occurs with probability $\prod_{v\in V_i} p_i(\phi_i(v))$ where p_i is the probability distribution on X_i.

The particular hypergraph which arises naturally when we study the quantum bias of free XOR games is constructed as follows. Starting with a t-partite t-uniform hypergraph H, write $db_i(H)$ for the t-partite t-uniform hypergraph obtained by making two vertex-disjoint copies of H and gluing them together so that the vertices in the two copies of V_i are identified. To construct our hypergraph, we start with the hypergraph given by a single edge $e = (v_1, \ldots, v_t)$ and vertex sets of size 1 and apply the doubling operation to all parts, i.e. $db_1(db_2(\ldots db_t(e)))$. We denote this hypergraph by $H(t)$. A more useful way to define $H(t)$ is as follows. We will do this first for $t = 2$ and explain how to do it for any t afterwards. We use 2-bit strings to identify vertices. We start with the hypergraph with a single edge $(x_{00}, y_{00}) \in V_1 \times V_2$. As we will start using the doubling operator, we make copies of the vertex sets. We can use a table to visualize it.

	V_1	V_2
starting position	x_{00}	y_{00}
db_2	x_{01}	y_{00}
db_1	x_{00}	y_{10}
	x_{01}	y_{10}

The table may be read as follows; the rows are the edges of the hypergraph and columns are the vertex sets. In this example we have that $V_1 = \{x_{00}, x_{01}\}$ and $V_2 = \{y_{00}, y_{10}\}$ and the edge set consists of $\{(x_{00}, y_{00}), (x_{01}, y_{00}), (x_{00}, y_{10}), (x_{01}, y_{10})\}$. The algorithm for constructing the table is as follows: we start with the starting position row, which corresponds to the (hyper)graph with a single edge (x_{00}, y_{00}), and as we apply the doubling operator db_2, we add a new row (which corresponds to making a vertex-disjoint copy) where we increase the 2nd bit in the subscript of x but leave y alone (so we have a new copy of V_1 but not of V_2). After this first step we have a graph with vertex sets $V_1 = \{x_{00}, x_{01}\}$ and $V_2 = \{y_{00}\}$ and edge set $\{(x_{00}, y_{00}), (x_{01}, y_{00})\}$. Next we apply db_1 and we get a new copy of V_2, but leave V_1 alone.

For arbitrary $t \geq 2$; let v_ω^i be a formal variable with $i \in [t]$ and ω is a t-bit string. We define for $j \in [t]$ an operation on the formal variable by

$$\Delta_j(v_\omega^i) := v_{\omega_1, \ldots, \omega_j + 1, \ldots, \omega_t}^i \text{ for } j \neq i$$
$$\Delta_i(v_\omega^i) := v_\omega^i.$$

where we add modulo 2. The table then looks like

	V_1	V_2	$\ldots$	V_t
starting position	$v_{0^t}^1$	$v_{0^t}^2$	$\ldots$	$v_{0^t}^t$
db_t	$\Delta_t(v_{0^t}^1)$	$\Delta_t(v_{0^t}^2)$	$\ldots$	$\Delta_t(v_{0^t}^t)$
db_{t-1}	$\Delta_{t-1}(v_{0^t}^1)$	$\Delta_{t-1}(v_{0^t}^2)$	$\ldots$	$\Delta_{t-1}(v_{0^t}^t)$
	$\Delta_{t-1}(\Delta_t(v_{0^t}^1))$	$\Delta_{t-1}(\Delta_t(v_{0^t}^2))$	$\ldots$	$\Delta_{t-1}(\Delta_t(v_{0^t}^t))$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$

At step k, the algorithm takes all the rows of the previous steps together and applies Δ_{t-k+1} on each of the formal variables in the rows. We also write $db_i(e)$ for the row where we apply Δ_i on each variable of the row e. We see in this way that, for example, the edge set of $H(t)$ has cardinality 2^t and the number of vertices in each V_i is 2^{t-1}. In the following proposition we list some properties of $H(t)$ which we prove using this description. We will be using the terms row and edge interchangeably as they mean the same in this context.

Proposition 4.4. *The hypergraph $H(t)$ has the following properties: (1) it is t-partite and t-uniform, (2) it is 2-regular and (3) for all vertices v the following holds: let e, e' be the unique edges such that $v \in e$, $v \in e'$ and $e \neq e'$. For $w \in e \setminus \{v\}$, denote by e, e'' the unique edges such that $w \in e$, $w \in e''$ and $e \neq e''$. Then $e' \cap e'' = \emptyset$.*

Proof. (1) follows directly from the algorithm described above using the table. We can prove (2) as follows. Suppose in column V_i we have a vertex in some row/edge which we denote by v^i_ω, here ω is a t-bit string. First we note that applying db_j with $j \neq i$ will change ω as it will flip the j-th bit. There are two cases; either we have already applied db_i in which case v^i_ω appears in exactly one more row above the current row, or we have not applied db_i yet in which case there is no v^i_ω in an earlier row. It will appear exactly once in a later row since applying db_i will not change ω. For (3), choose again some vertex v^i_ω in V_i and denote by e the row which appears first in the table containing v^i_ω. The other row/edge which contains v^i_ω is $e' := \mathrm{db}_i(e)$. Now, let v^j_τ be a vertex in V_j with $j \neq i$ and $v^j_\tau \in e$, i.e. it is in the same row as v^i_ω. There are two cases; either $j > i$ in which case $e = \mathrm{db}_j(e'')$ where e'' is the other (unique) edge containing v^j_τ. Or $j < i$ and the other edge which contains v^j_τ is $e'' := \mathrm{db}_j(e)$. In any case, a moments thought shows that $e' \cap e'' = \emptyset$. $\qquad\square$

The next ingredient is the following lemma.

Lemma 4.5. *For a t-player free XOR game G with game tensor T, we have that*

$$\beta^*(G) \leq \|T\|_{H(t)}.$$

Proof. For convenience, we write the hypergraph in a slightly different way. Write $\phi_i : V_i \to X_i$ and we define an operation Δ_j on such maps in the same way as above, i.e.

$$(\Delta_j\phi_i)(v^i_\omega) = \phi_i(\Delta_j v^i_\omega) \text{ for } j \neq i$$
$$(\Delta_i\phi_i)(v^i_\omega) = \phi_i(v^i_\omega).$$

Also, using the same symbol, we define on functions $T : X_1 \times \cdots \times X_t \to \mathbb{C}$

$$\Delta_j T(\phi_1(v^1_{\omega^1}), \ldots, \phi_t(v^t_{\omega^t})) := T(\phi_1(v^1_{\omega^1}), \ldots, \phi_t(v^t_{\omega^t})) T^*((\Delta_j\phi_1)(v^1_{\omega^1}), \ldots, (\Delta_j\phi_t)(v^t_{\omega^t})),$$

one could think of this operation as a kind of multiplicative derivative. If T were an operator-valued map, we still define it in this way. It is then not hard to see that

$$\Delta_1 \ldots \Delta_t T(\phi_1(v^1_{0^t}), \ldots, \phi_t(v^t_{0^t})) = \prod_{(v^1_{\omega^1}, \ldots, v^t_{\omega^t}) \in E(H(t))} T(\phi_1(\omega^1), \ldots, \phi_t(\omega^t)),$$

using the table as a description of $H(t)$. So we can write

$$\|T\|_{H(t)} = |\mathbb{E}\, \Delta_1 \ldots \Delta_t T(\phi_1(v^1_{0^t}), \ldots, \phi_t(v^t_{0^t}))|^{1/|E|},$$

where the expectation is taken over all maps $\phi_i : V_i \to X_i$ with the particular distribution given in Definition 4.3.

Now let us look at the bias of a two player game G with game tensor T and strategies A, B

$$\eta = \|\mathop{\mathbb{E}}_{(x,y) \in X \times Y} T(x,y)A(x)B(y)\|_{\mathrm{op}}.$$

We will do the example of two players to clarify the idea and will prove it in general afterwards. First we use Proposition 4.1, the Cauchy-Schwarz inequality, to peel off strategy B

$$\eta = \|\mathop{\mathbb{E}}_{y \in Y}\left(\mathop{\mathbb{E}}_{x \in X} T(x,y)A(x)\right)B(y)\|_{\mathrm{op}}$$

$$\leq \|\mathop{\mathbb{E}}_{y}\left(\mathop{\mathbb{E}}_{x} T(x,y)A(x)\right)\left(\mathop{\mathbb{E}}_{x} T(x,y)A(x)\right)^*\|_{\mathrm{op}}^{1/2} \|\mathop{\mathbb{E}}_{y} B(y)^*B(y)\|_{\mathrm{op}}^{1/2}$$

$$\leq \|\mathop{\mathbb{E}}_{y,x,x'} T(x,y)T(x',y)A(x)A(x')^*\|_{\mathrm{op}}^{1/2}.$$

In the second inequality we used that operator norm of strategies are smaller or equal to 1. We will use the Cauchy-Schwarz inequality one more time, now to peel off strategy A

$$\eta \leq \|\mathop{\mathbb{E}}_{x,x'}(\mathop{\mathbb{E}}_{y} T(x,y)T(x',y))A(x)A(x')^*\|_{\mathrm{op}}^{1/2}$$

$$\leq \|\mathop{\mathbb{E}}_{x,x'}(\mathop{\mathbb{E}}_{y} T(x,y)T(x',y))(\mathop{\mathbb{E}}_{y} T(x,y)T(x',y))^*\|_{\mathrm{op}}^{1/4} \|\mathop{\mathbb{E}}_{x,x'}(A(x)A(x')^*)^*A(x)A(x')^*\|_{\mathrm{op}}^{1/4}$$

$$\leq |\mathop{\mathbb{E}}_{x,x',y,y'} T(x,y)T(x',y)T(x,y')T(x',y')|^{1/4}.$$

Now, to see that this last expression is equal to $\|T\|_{H(2)}$, we write the expectation in a a different way. Instead of writing $\mathbb{E}_{x,x'}$ we write $\mathbb{E}_{\phi: V \to X}$ where $V = \{v_0, v_1\}$ is a vertex set, so that $x = \phi(v_0)$ and $x' = \phi(v_1)$. Similarly, instead of $\mathbb{E}_{y,y'}$ we write $\mathbb{E}_{\psi: W \to Y}$ where $W = \{w_0, w_1\}$ and we view $H(2)$ to be on this vertex sets. Then, we evaluate T on the edges of $H(2)$, so

$$|\mathop{\mathbb{E}}_{x,x',y,y'} T(x,y)T(x',y)T(x,y')T(x',y')|^{1/4} = |\mathop{\mathbb{E}}_{\phi: V \to X, \psi: W \to Y} \prod_{(v,w) \in E(H(2))} T(\phi(v), \psi(w))|^{1/4}.$$

In general, for t players, the proof is as follows

$$\eta := \|\mathbb{E}\, T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t))\left(\prod_{i \in [t-1]} A_i(\phi_i(v_{0^t}^i))\right)A_t(\phi_t(v_{0^t}^t))\|_{\mathrm{op}}$$

$$\leq \|\mathbb{E}\, T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t))T(\phi_1(v_{0^{t-1}1}^1), \ldots, \phi_t(v_{0^t}^t)) \prod_{i \in [t-1]} A_i(\phi_i(v_{0^t}^i))A_i(\phi_i(v_{0^{t-1}1}^i))^*\|_{\mathrm{op}}^{1/2}$$

$$= \|\mathbb{E}\, \Delta_t T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t)) \prod_{i \in [t-1]} \Delta_t A_i(\phi_i(v_{0^t}^i))\|_{\mathrm{op}}^{1/2}.$$

Now assume that we have applied the Cauchy-Schwarz inequality $1 < n < t$ times to peel off the last n operators and we have obtained the expression

$$\eta \leq \|\mathbb{E}\, \Delta_{t-n+1} \cdots \Delta_t T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t)) \prod_{i \in [t-n]} \Delta_{t-n+1} \cdots \Delta_t A_i(\phi_i(v_{0^t}^i))\|_{\mathrm{op}}^{1/2^n}.$$

Now apply Cauchy-Schwarz inequality to remove the operator $\Delta_{t-n+1} \cdots \Delta_t A_{t-n}(\phi_i(v_{0^t}^i))$ so that we obtain

$$\eta \leq \|\mathbb{E}\, \Delta_{t-n} \cdots \Delta_t T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t)) \prod_{i \in [t-n-1]} \Delta_{t-n} \cdots \Delta_t A_i(\phi_i(v_{0^t}^i))\|_{\mathrm{op}}^{1/2^{n+1}}.$$

This completes the induction. Putting $n = t - 1$ we have the inequality

$$\eta \leq |\mathbb{E}\, \Delta_1 \cdots \Delta_t T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t))|^{1/2^t}.$$

$\square$

We are now ready to give a proof of Theorem 1.4

Proof of Theorem 1.4. We assume $\beta^*(G) > \eta$. Lemma 4.5 immediately implies $\|T\|_{H(t)} > \eta$. To construct a classical strategy, we choose an edge $e^* = (v_1^*, \ldots, v_t^*) \in E(H(t))$. Any choice of edge works for our argument. $H(t)$ is 2-regular (by Proposition 4.4), so denote by e_i^* the unique edge different from e^* such that $v_i^* \in e_i^*$. Write $e_i^* = (v_1^{(i)}, \ldots, v_i^*, \ldots, v_t^{(i)})$ and $V_i' := V_i \setminus \{v_i^*\}$. Using Proposition 4.4 we see that $v_j^* \notin e_i^*$ whenever $i \neq j$. Then

$$
\eta^{2^t} < \left| \underset{\phi_i \colon V_i \to X_i}{\mathbb{E}} \prod_{(v_1, \ldots, v_t) \in E} T(\phi_1(v_1), \ldots, \phi_t(v_t)) \right|
$$

$$
= \Big| \underset{\phi_i \colon V_i' \to X_i}{\mathbb{E}} \Big[\prod_{(v_1, \ldots, v_t) \in E \setminus \{e^*, e_1^*, \ldots, e_t^*\}} T(\phi_1(v_1), \ldots, \phi_t(v_t))
$$

$$
\underset{\phi_i^* \colon \{v_i^*\} \to X_i}{\mathbb{E}} T(\phi_1^*(v_1^*), \ldots, \phi_t(v_t^*)) T(\phi_1^*(v_1^*), \ldots, \phi_t(v_t^{(1)})) \cdots T(\phi_1(v_1^{(t)}), \ldots, \phi_t^*(v_t^*)) \Big] \Big|
$$

$$
\leq \underset{\phi_i \colon V_i' \to X_i}{\mathbb{E}} \Big| \underset{\phi_i^* \colon \{v_i^*\} \to X_i}{\mathbb{E}} T(\phi_1^*(v_1^*), \ldots, \phi_t^*(v_t^*)) T(\phi_1^*(v_1^*), \ldots, \phi_t(v_t^{(1)}))
$$

$$
\cdots T(\phi_1(v_1^{(t)}), \ldots, \phi_t^*(v_t^*)) \Big|.
$$

Let us explain the second and third line in detail. Write $V_i = V_i' \cup \{v_i^*\}$. Any map $\phi_i \colon V_i \to X_i$ can be given by two maps $\phi_i' \colon V_i' \to X_i$ and $\phi_i^* \colon \{v_i^*\} \to X_i$ by defining $\phi_i(v)$ to be $\phi_i'(v)$ when $v \in V_i'$ and otherwise equal to $\phi_i^*(v)$. It can then be seen that

$$
\underset{\phi_i \colon V_i \to X_i}{\mathbb{E}} (\text{some expression}) = \underset{\phi_i' \colon V_i' \to X_i}{\mathbb{E}} \Big[\underset{\phi_i^* \colon \{v_i^*\} \to X_i}{\mathbb{E}} (\text{some expression}) \Big].
$$

After this we use the triangle inequality. Using the pigeonhole principle we see that there exists specific choices of maps $\phi_i \colon V_i' \to X_i$ such that

$$
\Big| \underset{\phi_i^* \colon \{v_i^*\} \to X_i}{\mathbb{E}} T(\phi_1^*(v_1^*), \ldots, \phi_t^*(v_t^*)) T(\phi_1^*(v_1^*), \ldots, \phi_t(v_t^{(1)})) \cdots T(\phi_1(v_1^{(t)}), \ldots, \phi_t^*(v_t^*)) \Big| > \eta^{2^t}.
$$

The expectation over t-tuples of maps $\phi_i^* \colon \{v_i^*\} \to X_i$ is the same as the expectation over t-tuples $x_i^* \in X_i$ and by defining

$$
a_i(x_i^*) := T(\phi_1(v_1^{(i)}), \ldots, x_i^*, \ldots, \phi_t(v_t^{(i)}))
$$

we see that

$$
\Big| \underset{x_1^*, \ldots, x_k^*}{\mathbb{E}} T(x_1^*, \ldots, x_k^*) \prod_{i=1}^{k} a_i(x_i^*) \Big| > \eta^{2^t},
$$

in other words, the classical bias is at least η^{2^t}. $\qquad\square$

5 Linear forms game

Before we will go in to the details of the proof of Theorem 1.5, we briefly discuss some concepts from higher order Fourier analysis. The reference for this subsection is [Tao12].

5.1 Preliminaries

Let G be a finite abelian group and $f\colon G \to \mathbb{C}$ a complex-valued function on G. First we define the multiplicative derivative Δ_h for any $h \in G$ for such functions

$$\Delta_h f(x) := f(x+h)\overline{f(x)}.$$

We define for any $s \geq 1$ the Gowers norm $\| \cdot \|_{U^s(G)}$

$$\|f\|_{U^s(G)} := \big(\mathop{\mathbb{E}}_{h_1,\ldots,h_s,x\in G} \Delta_{h_1}\cdots\Delta_{h_s}f(x)\big)^{1/2^s}. \tag{8}$$

For $s = 1$ we get the absolute value of the mean of the function

$$\|f\|_{U^1(G)} := \big(\mathop{\mathbb{E}}_{h,x\in G}\Delta_h f(x)\big)^{1/2} = |\mathop{\mathbb{E}}_{x\in G} f(x)|,$$

so technically it is not a norm, but for $s > 1$ it is indeed a norm. By the recursion

$$\|f\|_{U^{s+1}(G)}^{2^{s+1}} = \mathop{\mathbb{E}}_{h\in G}\|\Delta_h f\|_{U^s(G)}^{2^s}$$

one sees that the expectation in Equation 8 is a non-negative real.

Now let $\psi_0,\ldots,\psi_t\colon G^d \to G$ be affine linear forms, i.e. maps of the form $\psi_i(g_1,\ldots,g_m) = c_i + \sum_{j=1}^m c_{ij}g_j$ wher $c_i \in G$ and $c_{ij} \in \mathbb{Z}$.

Definition 5.1. *Let $\{\psi_0,\ldots,\psi_t\}$ be a system of affine linear forms. We say that the system has Cauchy-Schwarz complexity at most s if for any $0 \leq i \leq t$ one can partition $\{\psi_0,\ldots,\psi_t\} \setminus \{\psi_i\}$ into $s+1$ classes (empty classes are allowed) such that ψ_i does not lie in the affine linear span (over $\mathbb{Q}$) of the forms in any of these classes. The Cauchy-Schwarz complexity of the system is defined to be the least such s or ∞ if no such s exists.*

If $\psi\colon G^d \to G$ is an affine linear form, we denote by $\dot{\psi}\colon \mathbb{Q}^d \to \mathbb{Q}$ the map induced by its integer coefficients. The characteristic of G is defined to be least order of all non-identity elements. Here is an equivalent formulation of Cauchy-Schwarz complexity in terms of change of variables.

Proposition 5.2. *Let $\{\psi_0,\ldots,\psi_t\}$ be a system of affine linear forms $G^d \to G$. Suppose that the characteristic of G is sufficiently large depending on the coefficients of $\psi_0,\ldots,\psi_t$. Then the system has Cauchy-Schwarz complexity at most s if and only if for every $0 \leq i \leq t$ one can find a linear change of variables $\vec{x} = L_i(y_1,\ldots,y_{s+1},z_1,\ldots,z_d)$ on $\mathbb{Q}^d$ such that the form $\dot{\psi}_i(L_i(y_1,\ldots,y_{s+1},z_1,\ldots,z_d))$ has non-zero $y_1,\ldots,y_{s+1}$ coefficients, but all other forms $\dot{\psi}_j(L_i(y_1,\ldots,y_{s+1},z_1,\ldots,z_d))$ with $j \neq i$ have at least one vanishing $y_1,\ldots,y_{s+1}$ coefficient.*

We need also the notion of non-classical polynomials, as this is important in the inverse Gowers theorem.

Definition 5.3. *Let V be a finite dimensional vector space over a finite field of characteristic p, G an abelian group and let $P\colon V \to G$ be a map. We denote by d_h the additive derivative, i.e. $d_h P(x) := P(x+h) - P(x)$. We say that P is a non-classical polynomial of degree $\leq s$ if one has*

$$d_{h_1}\ldots d_{h_{s+1}}P(x) = 0,$$

for all $h_1,\ldots,h_{s+1} \in V$. We adopt the convention that the zero polynomial has degree $-\infty$. We denote the set of non-classical polynomial of degree $\leq s$ by $Poly_{\leq s}(V \to G)$.

The abelian group G will usually be $\mathbb{T} := \mathbb{R}/\mathbb{Z}$. In this case we have the following proposition, see [TZ12].

Proposition 5.4. *Let $P \in Poly_{\leq s}(V \to \mathbb{T})$. Then there exists $\alpha \in \mathbb{T}$ such that P takes values in the coset $\alpha + \frac{1}{p^{\lfloor \frac{s-1}{p-1} \rfloor + 1}} \mathbb{Z}/\mathbb{Z}$ of the $(p^{\lfloor \frac{s-1}{p-1} \rfloor + 1})^{th}$ roots of unity, where p is the characteristic of the ground field of V.*

A consequence of the above proposition is that low degree polynomials, in comparison with the characteristic p, are polynomials in the classical sense (up to constants), i.e. $\tilde{P}\colon V \to \mathbb{F}_p$. We will now recall Gowers inverse theorem for vector spaces over finite fields.

Theorem 5.5. *(Gowers inverse) Let $V := \mathbb{F}_p^n$ be a finite dimensional vector space over $\mathbb{F}_p$. Let $f\colon V \to \mathbb{C}$ be a function bounded in magnitude by 1 and also let $\varepsilon > 0$. If $\|f\|_{U^{s+1}} \geq \varepsilon$, then there exists a non-classical polynomial $P\colon V \to \mathbb{T}$ of degree at most s and a constant $\delta(\varepsilon, p, s) > 0$ such that*

$$|\underset{x \in V}{\mathbb{E}} f(x)e(P(x))| \geq \delta(\varepsilon, p, s).$$

Here $e\colon \mathbb{T} \to \mathbb{C} : \theta \mapsto e^{2\pi i \theta}$.

5.2 Results

We will now continue with line games. Line games, as discussed in the introduction, fall inside a larger class of games which we will describe first. For this, let Γ be a finite abelian group, let $m \geq 1$ an integer and we also have $t + 1$ affine linear forms $\psi_0, \ldots, \psi_t \colon \Gamma^m \to \Gamma$, i.e.

$$\psi_i(g_1, \ldots, g_m) = c_i + \sum_{j=1}^{m} c_{ij} g_j$$

where $(g_1, \ldots, g_m) \in \Gamma^m$, $c_i \in \Gamma$ and $c_{ij} \in \mathbb{Z}$.

Definition 5.6. *A t-player linear forms game is given by the above data together with a game map $\rho\colon \Gamma \to \{0,1\}$ as follows. The referee samples a uniform random point g from Γ^m and sends $\psi_i(g)$ to player i (players are numbered from 1 to t). The winning criterion is given by $\rho(\psi_0(g))$.*

Let G be such a game. The classical bias is given by

$$\beta(G) = \max_{a_i \colon \Gamma \to \{\pm 1\}} | \underset{g \in \Gamma^m}{\mathbb{E}} (-1)^{\rho(\psi_0(g))} \prod_{i=1}^{t} a_i(\psi_i(g))|.$$

The quantum bias is

$$\beta^*(G) = \max_{N \geq 1, A_i \colon \Gamma \to \mathrm{Obs}^{\pm 1}(\mathbb{C}^N)} \| \underset{g \in \Gamma^m}{\mathbb{E}} (-1)^{\rho(\psi_0(g))} \prod_{i=1}^{t} A_i(\psi_i(g))\|_{\mathrm{op}}.$$

Here $\mathrm{Obs}^{\pm 1}(\mathbb{C}^N)$ is the set of ± 1-valued observables on $\mathbb{C}^N$.

Remark. *It seems ad hoc that we consider such games, but these types of expressions are well studied in the context of counting linear patterns in finite abelian groups. We refer to [Tao12].*

The main technical theorem of this section is the following, from which we will deduce all other results.

Theorem 5.7. *Let G be a game as above. If the Cauchy-Schwarz complexity of $\{\psi_0, \ldots, \psi_t\}$ is at most s, we then have the inequality*

$$\beta^*(G) \leq \|(-1)^\rho\|_{U^{s+1}(\Gamma)}. \tag{9}$$

To prove this theorem, we need the following lemma.

Lemma 5.8 (Second Cauchy-Schwarz-Gowers inequality for operators). *Let $f : \Gamma \to \mathbb{C}$ be a function, $A_i : \Gamma^m \to \operatorname{End}(\mathbb{C}^N)$ for $i \in [m]$ such that $\|A_i(g)\|_{\mathrm{op}} \leq 1$ for any $g \in \Gamma^m$, A_i is independent of the i-th coordinate of g and $[A_i(g), A_j(h)] = 0$, $[A_i(g)^*, A_j(h)] = 0$ for all $i \neq j$ and $g, h \in \Gamma^m$. Then we have*

$$\|\mathop{\mathbb{E}}_{(g_1, \ldots, g_m) \in \Gamma^m} f(a_1 g_1 + \cdots + a_m g_m) \prod_{i=1}^{m} A_i(g_1, \ldots, g_m)\|_{\mathrm{op}} \leq \|f\|_{U^m(\Gamma)},$$

where a_i are non-zero integers such that the characteristic of Γ exceeds all of them.

Proof. We will prove this by induction. For $m = 1$ we have

$$\|\mathop{\mathbb{E}}_{g \in \Gamma} f(ag) A(g)\|_{\mathrm{op}} \leq |\mathop{\mathbb{E}}_{g \in \Gamma} f(ag)| = |\mathop{\mathbb{E}}_{g \in \Gamma} f(g)| = \|f\|_{U^1(\Gamma)}.$$

Here we used that A is independent of g. Assume we have proven the statement up to some integer $m \geq 1$. Then

$$\eta := \|\mathop{\mathbb{E}}_{(g_1, \ldots, g_{m+1}) \in \Gamma^{m+1}} f(a_1 g_1 + \cdots + a_{m+1} g_{m+1}) \prod_{i=1}^{m+1} A_i(g_1, \ldots, g_{m+1})\|_{\mathrm{op}}$$

$$= \|\mathop{\mathbb{E}}_{(g_2, \ldots, g_{m+1}) \in \Gamma^m} A_1(g_2, \ldots, g_{m+1}) \mathop{\mathbb{E}}_{g_1 \in \Gamma} f(a_1 g_1 + \cdots + a_{m+1} g_{m+1}) \prod_{i=2}^{m+1} A_i(g_1, \ldots, g_{m+1})\|_{\mathrm{op}},$$

we have done nothing, just rearranged and used the fact that A_1 is independent of g_1. Now write $F(g_2, \ldots, g_{m+1}) := \mathbb{E}_{g_1 \in \Gamma} f(a_1 g_1 + \cdots + a_{m+1} g_{m+1}) \prod_{i=2}^{m+1} A_i(g_1, \ldots, g_{m+1})$ so that

$$\eta = \|\mathop{\mathbb{E}}_{(g_2, \ldots, g_{m+1}) \in \Gamma^m} A_1(g_2, \ldots, g_{m+1}) F(g_2, \ldots, g_{m+1})\|_{\mathrm{op}}$$

$$\leq \|\mathop{\mathbb{E}}_{(g_2, \ldots, g_{m+1}) \in \Gamma^m} F(g_2, \ldots, g_{m+1}) F(g_2, \ldots, g_{m+1})^*\|_{\mathrm{op}}^{1/2}.$$

Here we used Proposition 4.1 and we used the fact that $\|A_1(g)\|_{\mathrm{op}} \leq 1$ for any $g \in \Gamma^{m+1}$. Then

$$
\eta \leq \| \underset{g_1, g_1', g_2, \ldots, g_{m+1}}{\mathbb{E}} f(a_1 g_1 + \cdots + a_{m+1} g_{m+1}) f(a_1 g_1' + \cdots + a_{m+1} g_{m+1})^*
$$
$$
\times \prod_{i=2}^{m+1} A_i(g_1, \ldots, g_{m+1}) A_i(g_1', \ldots, g_{m+1})^* \|_{\mathrm{op}}^{1/2}
$$
$$
\leq \big(\underset{g_1, h_1}{\mathbb{E}} \| \underset{(g_2, \ldots, g_{m+1}) \in \Gamma^m}{\mathbb{E}} \Delta_{h_1} f(a_1 g_1 + \cdots + a_{m+1} g_{m+1})
$$
$$
\times \prod_{i=2}^{m+1} A_i(g_1 + h_1, \ldots, g_{m+1}) A_i(g_1, \ldots, g_{m+1})^* \|_{\mathrm{op}} \big)^{1/2}
$$
$$
\leq \big(\underset{g_1, h_1}{\mathbb{E}} \big(\underset{h_2, \ldots, h_{m+1}, z \in \Gamma}{\mathbb{E}} \Delta_{h_{m+1}} \cdots \Delta_{h_1} f(a_1 g_1 + z))^{1/2^m} \big)^{1/2}
$$
$$
\leq \big(\underset{g_1, h_1}{\mathbb{E}} \big(\underset{h_2, \ldots, h_{m+1}, z \in \Gamma}{\mathbb{E}} \Delta_{h_{m+1}} \cdots \Delta_{h_1} f(a_1 g_1 + z))) \big)^{1/2^{m+1}}
$$
$$
= \big(\underset{h_1, h_2, \ldots, h_{m+1}, z \in \Gamma}{\mathbb{E}} \Delta_{h_{m+1}} \cdots \Delta_{h_1} f(z))^{1/2^{m+1}} = \|f\|_{U^{m+1}(\Gamma)}.
$$

In the third line we used triangle inequality to get the expectation in g_1, h_1 outside the norm. In the fifth line we used the induction hypothesis to upper bound the expression in the previous line with the Gowers norm. We then use in the sixth line Jensens inequality. $\qquad \square$

Proposition 5.9 (Generalized von Neumann inequality). *Let $f \colon \Gamma \to \mathbb{C}$ be a function, $\{\psi_0, \ldots, \psi_t\}$ a system of affine linear forms of Cauchy-Schwarz complexity s, $A_i \colon \Gamma^m \to \mathrm{End}(\mathbb{C}^N)$ for $i \in [t]$ such that $\|A_i(g)\|_{\mathrm{op}} \leq 1$ for any $g \in \Gamma^m$ and $[A_i(g), A_j(h)] = 0$, $[A_i(g)^*, A_j(h)] = 0$ for all $i \neq j$ and $g, h \in \Gamma^m$. Also assume the characteristic of Γ is sufficiently large depending on the coefficients of the affine linear forms. Then we have the inequality*

$$
\| \underset{g \in \Gamma^m}{\mathbb{E}} f(\psi_0(g)) \prod_{i=1}^{t} A_i(\psi_i(g)) \|_{\mathrm{op}} \leq \|f\|_{U^{s+1}(\Gamma)}.
$$

Proof. The system of affine linear forms has Cauchy-Schwarz complexity s so we can partition the forms $\{\psi_1, \ldots, \psi_t\}$ into $s+1$ classes $\mathcal{A}_1, \ldots, \mathcal{A}_{s+1}$ such that ψ_0 is not an affine linear combination of any forms in any class $\mathcal{A}_i$ for any i (over $\mathbb{Q}$). So one can find a linear change of variables using Proposition 5.2

$$
(g_1, \ldots, g_m) \mapsto (h_1, \ldots, h_m) + y_1 v_1 + \cdots + y_{s+1} v_{s+1}
$$

with the property that $\psi_0(y_j v_j) = a_j y_j$ where a_j is a non-zero integer and $v_j \in \mathbb{Z}^m$, but if $\psi_i \in \mathcal{A}_j$, then $\psi_i(y_j v_j) = 0$, this is where we need the large characteristic hypothesis. Now we define

$$
\tilde{A}_k(g_1, \ldots, g_m) := \prod_{j \in \mathcal{A}_k} A_j(\psi_j(g_1, \ldots, g_m)).
$$

Note that $\tilde{A}_i$ is independent of its i-th coordinate and has operator norm smaller than 1. We then

have

$$\| \underset{g \in \Gamma^m}{\mathbb{E}} f(\psi_0(g_1,\ldots,g_m)) \prod_{i=1}^{t} A_i(\psi_i(g_1,\ldots,g_m)) \|_{\mathrm{op}}$$

$$= \| \underset{g \in \Gamma^m}{\mathbb{E}} f(\psi_0(g_1,\ldots,g_m)) \prod_{i=1}^{s+1} \tilde{A}_i(g_1,\ldots,g_m) \|_{\mathrm{op}}$$

$$= \| \underset{h \in \Gamma^m}{\mathbb{E}} \underset{y_1,\ldots,y_{s+1} \in \Gamma}{\mathbb{E}} f(\psi_0(h) + a_1 y_1 + \cdots + a_{s+1} y_{s+1}) \prod_{i=1}^{s+1} \tilde{A}_i(h, y_1,\ldots,y_{s+1}) \|_{\mathrm{op}}$$

$$\leq \underset{h \in \Gamma^m}{\mathbb{E}} \|f\|_{U^{s+1}(\Gamma)} = \|f\|_{U^{s+1}(\Gamma)}.$$

In the third line we used the linear change of variables just described. Then we used the triangle inequality together with Lemma 5.8 where we need the large characteristic hypothesis. □

Remark. *If f takes values in $\{\pm 1\}$ and A_i are ± 1-valued observables, then the inequality says that the quantum bias of such games is upperbounded by the Gowers norm of the game tensor f.*

Proposition 5.9 is in full generality, i.e. for any abelian group the inequality holds. However, we will now restrict ourselves to the case where $\Gamma = \mathbb{F}_p^n$, where p is prime and $n \geq 1$ as it will make many things easier. We can use Gowers inverse theorem for vector spaces over finite fields which we recalled in the preliminaries section. We will also assume that p is sufficiently large, so that the set of non-classical polynomials coincide with the set of (classical) polynomials. Let us start giving the proof of Theorem 1.5. A t-player line game is given by a map $\tau \colon \mathbb{F}_p^n \to \{0,1\}$ which stands for the predicate together with a system of linear forms $\psi_0, \psi_i \colon (\mathbb{F}_p^n)^2 \to \mathbb{F}_p^n$ which are given by

$$\psi_0(x,y) = y \text{ and } \psi_i(x,y) = x + (i-1)y \text{ for } i = 1,\ldots,t.$$

Note that the Cauchy-Schwarz complexity of this system is at most $t-1$. For the bias of the game, it is more convenient to look at $f := (-1)^\tau$. We also need the following lemma, provided kindly to us by Shravas Rao.

Lemma 5.10. *Let $P \colon \mathbb{F}_p^n \to \mathbb{F}_p$ be a (classical) polynomial of degree $d-1$ and $p \geq d$. Then there exists d polynomials $P_i \colon \mathbb{F}_p^n \to \mathbb{F}_p$, $i = 0,\ldots,d-1$, such that*

$$P(y) = \sum_{i=0}^{d-1} P_i(x + iy).$$

Proof. The polynomial P can be represented as

$$P(x_1,\ldots,x_n) = \sum_{i=0}^{d-1} T_i(x,\ldots,x), \quad x = (x_1,\ldots,x_n),$$

where each $T_i \colon (\mathbb{F}_p^n)^i \to \mathbb{F}_p$ is an i-linear form. We will show that for each linear form T_i we can find $\alpha_0,\ldots,\alpha_{d-1}$ such that

$$T_i(y,\ldots,y) = \sum_{j=0}^{d-1} \alpha_j T_i(x + jy,\ldots,x + jy) \tag{10}$$

25

and this will be enough to construct $P_0, \ldots, P_{d-1}$. By linearity, we can rewrite the right hand side as follows,

$$\sum_{j=0}^{d-1} \sum_{s \in \{0,1\}^i} \alpha_j T_i((1-s_1)x + s_1 jy, \ldots, (1-s_i)x + s_i jy)$$

$$= \sum_{j=0}^{d-1} \sum_{s \in \{0,1\}^i} \alpha_j j^{|s|} T_i((1-s_1)x + s_1 y, \ldots, (1-s_i)x + s_i y),$$

where $|s|$ denotes the Hamming weight of s. Then 10 holds, if for $0 \leq k < i$

$$\sum_{j=0}^{d-1} \alpha_j j^k = 0 \text{ and } \sum_{j=0}^{d-1} \alpha_j j^i = 1.$$

As $d \leq p$ the $d \times d$ Vandermonde matrix associated with the sequence $1, j, \ldots, j^i$ is invertible, hence there exist unique $\alpha_0, \ldots, \alpha_{d-1}$ satisfying the above equations which concludes the proof. $\qquad \square$

The following lemma will help us later in converting complex strategies into ± 1-strategies.

Lemma 5.11. *For any $z \in \mathbb{C}$*

$$z = \frac{\pi}{2} \mathop{\mathbb{E}}_{|w|=1} [\mathrm{sgn}(z\overline{w}) \, |z| \, w],$$

where $w \in \{z \in \mathbb{C} : |z| = 1\}$ is taken uniformly at random.

Proof. Write $z = re^{i\psi}$. Then

$$\mathop{\mathbb{E}}_{|w|=1} [\mathrm{sgn}(z\overline{w})|z|w] = \frac{r}{2\pi} \int_0^{2\pi} \mathrm{sgn}(e^{-i(\phi-\psi)})e^{i\phi}d\phi$$

$$= \frac{r}{2\pi} \int_0^{2\pi} \mathrm{sgn}(e^{-i\chi})e^{i\chi+i\psi}d\chi$$

$$= \frac{z}{2\pi} \int_0^{2\pi} \mathrm{sgn}(e^{-i\chi})e^{i\chi}d\chi$$

$$= \frac{2z}{\pi}.$$

$\qquad \square$

Proof of Theorem 1.5. By Proposition 5.9 and the hypothesis that the game has entangled value $\varepsilon > 0$ implies that $\|f\|_{U^t} > \varepsilon$. Then by Gowers inverse Theorem 5.5 and assumption that $p > t$ there exists a constant $\delta = \delta(\varepsilon, p, t) > 0$ and a (classical) polynomial of degree at most $t - 1$ such that

$$\left| \mathop{\mathbb{E}}_{x \in \mathbb{F}_p^n} f(x)e(P(x)) \right| > \delta.$$

We now want to convert this presence of structure into a classical strategy. First by Lemma 5.10 we can find t polynomials P_i for $i = 1, \ldots, t$ such that

$$P(y) = \sum_{i=1}^t P_i(x + (i-1)y).$$

26

This implies

$$|\mathop{\mathbb{E}}_{x,y\in\mathbb{F}_p^n} f(\psi_0(x,y))e(P(\psi_0(x,y)))| = |\mathop{\mathbb{E}}_{x,y\in\mathbb{F}_p^n} f(\psi_0(x,y)) \prod_{i=1}^{t} e(P_i(\psi_i(x,y)))| > \delta.$$

The polynomials are not classical strategies yet, we can turn it into ± 1-strategy using Lemma 5.11 at a loss of a factor $2^t/\pi^t$. $\qquad\square$

5.3 Parallel repetition

Let $f\colon \Gamma \to \{1,-1\}$ be a function, representing the predicate. We want to consider k-fold XOR parallel repetition. The predicate for this is $f^k\colon \Gamma^k \to \{1,-1\}$ defined by

$$f^k(g_1,\ldots,g_k) := \prod_{i=1}^{k} f(g_i).$$

Lemma 5.12. *We have that*

$$\|f^k\|_{U^{s+1}(\Gamma^k)} = \|f\|_{U^{s+1}(\Gamma)}^k.$$

Proof. This follows immediately from the definition of Gowers norm. $\qquad\square$

Let $\{\psi_0,\ldots,\psi_t\}$ be linear forms $\Gamma^m \to \Gamma$ which together with f define the game G. The linear forms corresponding with k-fold XOR parallel repetition are denoted by $\{\psi_0^k,\ldots,\psi_t^k\}$ which are maps $(\Gamma^m)^k \to \Gamma^k$ and are given by

$$\psi_i^k(g^1,\ldots,g^k) := (\psi_i(g^1),\ldots,\psi_i(g^k)), \text{ where } g^i \in \Gamma^m.$$

Note that if $\{\psi_0,\ldots,\psi_t\}$ has Cauchy-Schwarz complexity at most s, then the Cauchy-Schwarz complexity of $\{\psi_0^k,\ldots,\psi_t^k\}$ is also at most s. Denote by $G^{\oplus k}$ the k-fold XOR parallel repetition, then we have as an immediate consequence of Proposition 5.9 together with Lemma 5.12 the following upper bound

$$\beta^*(G^{\oplus k}) \le \|f\|_{U^{s+1}(\Gamma)}^k.$$

If G is an XOR game, denote by G^k the k-fold parallel repetition. If S is a strategy (classical or quantum) for a game G, denote by $\omega(G,S)$ the winning probability using strategy S. Also denote by $\varepsilon(G,S) := 2\omega(G,S) - 1$ the bias of this strategy. To prove Lemma 1.6, we use the following lemma, which is a straightforward generalization of the 2-player version in [CSUU08] (lemma 8 in that paper) to any number of players.

Lemma 5.13. *Let G be an XOR game assume. Let S be any strategy for G^k. For each $M \subset [k]$, we denote by S_M the following strategy for the XOR parallel repetition $\oplus_{i\in M}G$: (1) Run strategy S, yielding answers $a_i^1,\ldots,a_i^k$ for player $i = 1,\ldots,t$. (2) Player i outputs $\sum_{j\in M} a_i^j \mod 2$. We then have*

$$\omega(G^k,S) = \frac{1}{2^k} \sum_{M\subset[k]} \varepsilon(\oplus_{i\in M}G,S_M).$$

Proof of Lemma 1.6. Let S be the quantum strategy that achieves the maximum winning probability of the game G^k. We then use Lemma 5.13,

$$\omega^*(G^k) = \omega(G^k, S) = \frac{1}{2^k} \sum_{M \subset [k]} \varepsilon(\oplus_{i \in M} G, S_M)$$

$$\leq \frac{1}{2^k} \sum_{M \subset [k]} \beta^*(G^{\oplus|M|}) = \frac{1}{2^k} \sum_{l=0}^{k} \beta^*(G^{\oplus l}) \binom{k}{l}$$

$$\leq \frac{1}{2^k} \sum_{l=0}^{k} \binom{k}{l} \|f\|_{U^{s+1}(\Gamma)} = \left(\frac{1 + \|f\|_{U^{s+1}(\Gamma)}}{2} \right)^k.$$

$\square$

6   Near-perfect strategies for 2-player unique games

In this section we prove Theorem 1.7. Consider a unique game where π_{xy} is the matching between the players' answers on inputs x, y, so that when the first player answers i they win if the second player answers $j = \pi_{xy}(i)$. Let us start by writing down an expression for the entangled winning probability when the players use a shared state $|\psi\rangle$ and projectors $\Pi_i^{(x)}, \Pi_j^{(y)}$ for inputs x, y and outputs i, j. For finite-dimensional systems we can always assume that a strategy is of such a form. The winning probability is given by

$$\mathbb{E}_{x,y} \sum_{i=1}^{k} \Pr(\text{answer } i, \pi_{xy}(i) \mid \text{input } x, y) = \mathbb{E}_{x,y} \sum_{i=1}^{k} \langle \psi | \Pi_i^{(x)} \otimes \Pi_{\pi_{xy}(i)}^{(y)} | \psi \rangle.$$

Now define vectors $|u_i^{(x)}\rangle = (\Pi_i^{(x)} \otimes \mathrm{Id})|\psi\rangle$ and $|v_j^{(y)}\rangle = (\mathrm{Id} \otimes \Pi_j^{(y)})|\psi\rangle$, then we can write the winning probability as

$$\mathbb{E}_{x,y} \sum_{i=1}^{k} \langle u_i^{(x)} | v_{\pi_{xy}(i)}^{(y)} \rangle \geq 1 - \varepsilon \tag{11}$$

where we use the assumption that there is a strategy with entangled value at least $1 - \varepsilon$. The vectors have the following properties:

$$\forall x, y, \forall i \neq j \qquad \langle u_i^{(x)} | u_j^{(x)} \rangle = \langle v_i^{(y)} | v_j^{(y)} \rangle = 0 \qquad \text{(orthogonal projectors)}$$

$$\forall x, y \qquad \sum_{i=1}^{k} \|u_i^{(x)}\|^2 = \sum_{i=1}^{k} \|v_i^{(y)}\|^2 = 1 \qquad \text{(projectors sum to identity)}$$

$$\forall x, y, \forall i, j \qquad \langle u_i^{(x)} | v_j^{(y)} \rangle \geq 0 \qquad \text{(projectors are Hermitian)}$$

By using $\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\langle u | v \rangle$ (for real-valued inner products) we can write (11) as

$$\frac{1}{2} \mathbb{E}_{x,y} \sum_{i=1}^{k} \|u_i^{(x)} - v_{\pi_{xy}(i)}^{(y)}\|^2 \leq \varepsilon. \tag{12}$$

It is possible to maximize expression (11) (or equivalently minimize (12)) over vectors with the given properties. This optimization problem is an SDP and can be solved in polynomial time but will generally not yield a quantum strategy as not all such vectors can be attained by quantum strategies. Our goal will be to extract from the vectors a classical strategy, something known as rounding, such that its winning probability is high.

As stated in the introduction, one can get some intuition by considering the $\varepsilon = 0$ case. There (12) yields $|u_i^{(x)}\rangle = |v_{\pi_{xy}(i)}^{(y)}\rangle$ for each x, y and i. Using shared randomness the players sample a random vector $|g\rangle$ and compute the overlaps $\xi_i^{(x)} = \langle g|u_i^{(x)}\rangle$ and $\xi_i^{(y)} = \langle g|v_i^{(y)}\rangle$ respectively. The players will have the same values $\xi_i^{(x)} = \xi_{\pi_{xy}(i)}^{(y)}$ so both players can output the answer i for which their overlap has the largest value.

For $\varepsilon > 0$ the sets of vectors will not be exactly equal and therefore the values $\xi_i^{(x)}, \xi_{\pi_{xy}(i)}^{(y)}$ will be close but not exactly equal. The discrepancy in these values will be bigger for vectors $|u_i^{(x)}\rangle$ with a small norm. In Section 2 of [CMM06] a rounding algorithm is provided that solves these issues. Note that we write $u_i^{(x)}, v_j^{(y)}$ for the vectors belonging to questions x, y and answers i, j whereas in [CMM06] these vectors are instead denoted by u_i, v_j where u, v are the questions and i, j the answers. The only difference between their SDP and the above one is that they have an additional constraint $0 \leq \langle u_i^{(x)}|v_{\pi_{xy}(i)}^{(y)}\rangle \leq |u_i^{(x)}|^2$ (constraint (5) in their paper). This constraint does not necessarily hold in the quantum setting so we will drop it and adapt their proofs to work without this constraint.

The following is **Rounding Algorithm 2** from Section 4 of [CMM06], adapted to our notation.
Rounding algorithm
Input: A solution of the SDP with objective value $1 - \varepsilon$.
Output: A classical strategy: $a(x)$ and $b(y)$
Define $[x]_r$ as the function that rounds x up or down depending on whether the fractional part of x is greater or less than r. If r is uniform random on $[0, 1]$ then the expected value of $[x]_r$ is x.

1. Define $|\tilde{u}_i^{(x)}\rangle = |u_i^{(x)}\rangle / \|u_i^{(x)}\|$ if $\|u_i^{(x)}\| \neq 0$, otherwise $|\tilde{u}_i^{(x)}\rangle = 0$.

2. Pick $r \in [0, 1]$ uniformly at random.

3. Pick random independent Gaussian vectors $|g_1\rangle, ..., |g_{2k}\rangle$ with independent components distributed as $\mathcal{N}(0, 1)$.

4. For each question x:

   (a) Set $s_i^{(x)} = \left[2k \cdot \|u_i^{(x)}\|^2\right]_r$.

   (b) For each i project $s_i^{(x)}$ vectors $|g_1\rangle, ..., |g_{s_i^{(x)}}\rangle$ to $|\tilde{u}_i^{(x)}\rangle$:

   $$\xi_{i,s}^{(x)} = \langle g_s|\tilde{u}_i^{(x)}\rangle, \quad s = 1, 2, ..., s_i^{(x)}$$

   (c) Select the $\xi_{i,s}^{(x)}$ with the largest absolute value. Assign $a(x) = i$.

5. Repeat the previous step for each question y but with the vectors $|v_j^{(y)}\rangle$ to obtain $b(y)$.

29

The intuition behind the algorithm is as follows. Similar to the $\varepsilon = 0$ case, the values $\xi_{i,s}^{(x)}$ and $\xi_{\pi_{xy}(i),s}^{(y)}$ will be close. Vectors $|u_i^{(x)}\rangle$ and $|u_j^{(x)}\rangle$ for different answers $i \neq j$ are orthogonal and their corresponding values $\xi_{i,s}^{(x)}$ and $\xi_{j,s}^{(x)}$ are therefore independent. For vectors with small norm, the values $\xi_{i,s}^{(x)}$ and the matching $\xi_{\pi_{xy}(i),s}^{(x)}$ will be less correlated. Therefore we sample more Gaussian vectors for answers corresponding to a high norm (step 4a).

To prove Theorem 1.7 we have to show that the result of the above rounding algorithm is a strategy with winning probability $1 - \mathcal{O}(\sqrt{\varepsilon \log k})$. This is exactly the result of Theorem 4.5 of [CMM06] with the exception of the additional constraint mentioned before. This modification requires a different proof of Lemma 4.2 and 4.3 in [CMM06] but leaves the remaining part of their proof unchanged. We therefore only prove these Lemma's and refer the reader to Section 4 of [CMM06] for the remainder of the proof.

We adopt their definitions

$$\varepsilon_{xy} = \frac{1}{2} \sum_{i=1}^{k} \|u_i^{(x)} - v_{\pi_{xy}(i)}^{(y)}\|^2,$$

$$\varepsilon_{xy}^i = \frac{1}{2} \|\tilde{u}_i^{(x)} - \tilde{v}_{\pi_{xy}(i)}^{(y)}\|^2,$$

where $\tilde{u}_i^{(x)}$ and $\tilde{v}_i^{(y)}$ were defined in step 1 of the rounding algorithm. Note that $\mathbb{E}_{x,y}\,\varepsilon_{xy} \leq \varepsilon$.

Lemma 6.1 (Originally Lemma 4.2). *The probability that the rounding algorithm gives a correct assignment to the questions x, y is $1 - \mathcal{O}(\sqrt{\varepsilon_{xy} \log k})$.*

Proof. If $\varepsilon_{xy} \geq 1/128$ then the statement follows trivially since this can be hidden in the big-O. Therefore assume $\varepsilon_{xy} \leq 1/128$. Define

$$M = \left\{ (i,s) \; : \; i \in [k] \, , \, s \leq \min(s_i^{(x)}, s_{\pi_{xy}(i)}^{(y)}) \right\},$$

$$M_c = \left\{ (i,s) \; : \; i \in [k] \, , \, \min(s_i^{(x)}, s_{\pi_{xy}(i)}^{(y)}) < s \leq \max(s_i^{(x)}, s_{\pi_{xy}(i)}^{(y)}) \right\}.$$

The set M contains the pairs (i,s) for which both $\xi_{i,s}^{(x)}$ and $\xi_{\pi_{xy}(i),s}^{(y)}$ are defined and the set M_c contains the pairs for which only one of these is defined.

We need the following two lemmas to continue.

Lemma 6.2. *When $\varepsilon_{xy} \leq 1/128$ then $\mathbb{E}_r[|M_c|] \leq 4k\sqrt{2\varepsilon_{xy}}$ and $|M| \geq k/2$.*

This was originally Lemma 4.3 and it stated $E_r[|M_c|] \leq 4k\varepsilon_{xy}$.

Proof. The expected value of $|s_i^{(x)} - s_{\pi_{xy}(i)}^{(y)}|$ is given by

$$\mathbb{E}_r \left| \left[2k \cdot \|u_i^{(x)}\|^2 \right]_r - \left[2k \cdot \|v_{\pi_{xy}(i)}^{(y)}\|^2 \right]_r \right| = 2k \left| \|u_i^{(x)}\|^2 - \|v_{\pi_{xy}(i)}^{(y)}\|^2 \right|.$$

By the triangle inequality

$$\left| \|u_i^{(x)}\|^2 - \|v_{\pi_{xy}(i)}^{(y)}\|^2 \right| = \left| \|u_i^{(x)}\| - \|v_{\pi_{xy}(i)}^{(y)}\| \right| \left(\|u_i^{(x)}\| + \|v_{\pi_{xy}(i)}^{(y)}\| \right)$$

$$\leq \|u_i^{(x)} - v_{\pi_{xy}(i)}^{(y)}\| \left(\|u_i^{(x)}\| + \|v_{\pi_{xy}(i)}^{(y)}\| \right),$$

and by using Cauchy-Schwarz twice we have

$$\sum_{i=1}^{k}\|u_i^{(x)}-v_{\pi_{xy}(i)}^{(y)}\|\left(\|u_i^{(x)}\|+\|v_{\pi_{xy}(i)}^{(y)}\|\right)\leq\sqrt{\sum_{i=1}^{k}\|u_i^{(x)}-v_{\pi_{xy}(i)}^{(y)}\|^2}\sqrt{\sum_{i=1}^{k}\left(\|u_i^{(x)}\|+\|v_{\pi_{xy}(i)}^{(y)}\|\right)^2}$$

$$=\sqrt{2\varepsilon_{xy}}\sqrt{\sum_{i=1}^{k}\left(\|u_i^{(x)}\|^2+\|v_{\pi_{xy}(i)}^{(y)}\|^2+2\|u_i^{(x)}\|\|v_{\pi_{xy}(i)}^{(y)}\|\right)}$$

$$\leq\sqrt{2\varepsilon_{xy}}\sqrt{1+1+2}=2\sqrt{2\varepsilon_{xy}}.$$

The proof follows from $|M_c|=\sum_{i=1}^{k}|s_i^{(x)}-s_{\pi_{xy}(i)}^{(y)}|$. For the second part of the lemma, observe that

$$\min(s_i^{(x)},s_{\pi_{xy}(i)}^{(y)})\geq 2k\,\min(\|u_i^{(x)}\|^2,\|v_{\pi_{xy}(i)}^{(y)}\|^2)-1$$

$$\geq 2k\left(\|u_i^{(x)}\|^2-\left|\|u_i^{(x)}\|^2-\|v_{\pi_{xy}(i)}^{(y)}\|^2\right|\right)-1.$$

Therefore we have

$$|M|=\sum_{i=1}^{k}\min(s_i^{(x)},s_{\pi_{xy}(i)}^{(y)})\geq\sum_{i=1}^{k}\left(2k\,\|u_i^{(x)}\|^2-2k\left|\|u_i^{(x)}\|^2-\|v_{\pi_{xy}(i)}^{(y)}\|^2\right|-1\right)$$

$$\geq 2k-4k\sqrt{2\varepsilon_{xy}}-k\geq k/2,$$

where we used $\varepsilon_{xy}\leq 1/128$. $\qquad\square$

Lemma 6.3. *The following inequality holds*

$$\mathop{\mathbb{E}}_{r}\left[\frac{1}{|M|}\sum_{(i,s)\in M}\varepsilon_{xy}^{i}\right]\leq 4\varepsilon_{xy}$$

Proof. This is Lemma 4.4 in [CMM06]. $\qquad\square$

We now continue the proof of Lemma 6.1. First consider a fixed value of r (picked in step 2 of the rounding algorithm. Consider the sequences $\xi_{i,s}^{(x)}$ and $\xi_{\pi_{xy}(i),s}^{(y)}$ where the indices (i,s) run over all $(i,s)\in M$. We apply Theorem 4.1 of [CMM06] to these sequences and get that the probability that the largest absolute value in the first sequence has the same index as the largest absolute value in the second sequence is

$$1-\mathcal{O}\left(\sqrt{\log|M|\cdot\frac{1}{|M|}\sum_{(i,s)\in M}\varepsilon_{xy}^{i}}\right).$$

By Jensen's inequality we have

$$\mathop{\mathbb{E}}_{r}\left[1-\mathcal{O}\left(\sqrt{\log|M|\cdot\frac{1}{|M|}\sum_{(i,s)\in M}\varepsilon_{xy}^{i}}\right)\right]\geq 1-\mathcal{O}\left(\sqrt{\mathop{\mathbb{E}}_{r}\left[\log|M|\cdot\frac{1}{|M|}\sum_{(i,s)\in M}\varepsilon_{xy}^{i}\right]}\right)$$

$$\geq 1-\mathcal{O}\left(\sqrt{\varepsilon_{xy}\log k}\right)$$

31

where the second inequality follows from $|M| \leq 3k$ and Lemma 6.3. In the rounding algorithm, the largest $\xi_{i,s}^{(x)}$ is picked not only among the $(i,s) \in M$ but also $(i,s) \in M_c$. However, the probability that the index for the largest value is in M_c is at most

$$\mathbb{E}_r \left[\frac{|M_c|}{|M|} \right] \leq \frac{4k\sqrt{2\varepsilon_{xy}}}{k/2} = 8\sqrt{2\varepsilon_{xy}},$$

by Lemma 6.2. Therefore by the union bound, the probability that the answers match is at least

$$1 - \mathcal{O}(\sqrt{\varepsilon_{xy}\log k}) - 8\sqrt{2\varepsilon_{xy}} = 1 - \mathcal{O}(\sqrt{\varepsilon_{xy}\log k}),$$

which finishes the proof. $\qquad\square$

With these modified lemmas, Theorem 4.5 of [CMM06] shows that the rounding algorithm gives a classical strategy that wins the game with probability $1 - \mathcal{O}(\sqrt{\varepsilon \log k})$.

Acknowledgements

References

[BBC+01]   Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. Quantum lower bounds by polynomials. *Journal of the ACM*, 48(4):778–797, 2001. Earlier version in FOCS'98. [5]

[BBLV13]   Jop Briët, Harry Buhrman, Troy Lee, and Thomas Vidick. Multipartite entanglement in XOR games. *Quantum Information & Computation*, 13(3-4):334–360, 2013. [3]

[BCL+06]   Christian Borgs, Jennifer Chayes, László Lovász, Vera T Sós, and Katalin Veszter-gombi. Counting graph homomorphisms. In *Topics in discrete mathematics*, pages 315–371. Springer, 2006. [8]

[Bel64]   J.S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1(3):195–200, 1964. [2]

[BMMN11]   Mark Braverman, Konstantin Makarychev, Yury Makarychev, and Assaf Naor. The grothendieck constant is strictly smaller than Krivines bound. In *IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS '11)*, pages 453–462, 2011. [2]

[BOGKW88]   Michael Ben-Or, Shafi Goldwasser, Joe Kilian, and Avi Wigderson. Multi-prover interactive proofs: how to remove intractability assumptions. In *20th Annual ACM Symposium on Theory of Computing (STOC '88)*, pages 113–131, 1988. [2]

[Boy04]   Michel Boyer. Extended GHZ n-player games with classical probability of winning tending to 0. *arXiv preprint arXiv:0408090*, 09 2004. [4,10,12]

[BV13]   Jop Briët and Thomas Vidick. Explicit lower and upper bounds on the entangled value of multiplayer XOR games. *Communications in Mathematical Physics*, 321(1):181–207, 2013. [3]

[CHPS12]   David Conlon, Hiêp Hàn, Yury Person, and Mathias Schacht. Weak quasi-randomness for uniform hypergraphs. *Random Structures & Algorithms*, 40(1):1–38, 2012. [8]

[CHTW04]   Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *19th Annual IEEE Conference on Computational Complexity (CCC 2004), 21-24 June 2004, Amherst, MA, USA*, pages 236–249, 2004. [4]

[CL17]   David Conlon and Joonkyung Lee. Finite reflection groups and graph norms. *Advances in Mathematics*, 315:130–165, 2017. [8]

[CMM06]   Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for unique games. In *Proceedings of the Thirty-eighth Annual ACM Symposium on Theory of Computing*, STOC '06, pages 205–214, New York, NY, USA, 2006. ACM. [7,9,29,30,31,32]

[CSUU08]   Richard Cleve, William Slofstra, Falk Unger, and Sarvagya Upadhyay. Perfect parallel repetition theorem for quantum xor proof systems. *Computational Complexity*, 17(2):282–299, 2008. [27]

[DHVY16]   Irit Dinur, Prahladh Harsha, Rakesh Venkat, and Henry Yuen. Multiplayer parallel repetition for expander games. *arXiv preprint arXiv:1610.08349*, 2016. [6,11]

[DMR09]   Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional hardness for approximate coloring. *SIAM J. Comput.*, 39(3):843–873, 2009. [7]

[Gro53]   Alexander Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques (French). *Bol. Soc. Mat. São Paulo*, 8:1–79, 1953. [2]

[Hal27]   Philip Hall. The distribution of means for samples of size N drawn from a population in which the variate takes values between 0 and 1, all such values being equally probable. *Biometrika*, 19(3/4):240–245, 1927. [13]

[Hat09]   Hamed Hatami. *On generalizations of Gowers norms*. PhD thesis, University of Toronto, 2009. [8,14]

[Hat10]   Hamed Hatami. Graph norms and Sidorenko's conjecture. *Israel Journal of Mathematics*, 175(1):125–150, 2010. [8]

[IKP+08]   Tsuyoshi Ito, Hirotada Kobayashi, Daniel Preda, Xiaoming Sun, and Andrew C. C. Yao. Generalized Tsirelson inequalities, commuting-operator provers, and multi-prover interactive proof systems. In *Proceedings of the 2008 IEEE 23rd Annual Conference on Computational Complexity*, CCC '08, pages 187–198. IEEE Computer Society, 2008. [4,5]

[Irw27]   J. O. Irwin. On the frequency distribution of the means of samples from a population having any law of frequency with finite moments, with special reference to pearson's type ii. *Biometrika*, 19(3/4):225–239, 1927. [13]

[Kho02]   Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thiry-fourth Annual ACM Symposium on Theory of Computing (STOC '02)*, STOC '02, pages 767–775, New York, NY, USA, 2002. ACM. [2,4,6]

[KKMO07]   Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable csps? *SIAM J. Comput.*, 37(1):319–357, 2007. [7]

[KLL+15]   Iordanis Kerenidis, Sophie Laplante, Virginie Lerays, Jérémie Roland, and David Xiao. Lower bounds on information complexity via zero-communication protocols and applications. *SIAM J. Comp.*, 44(5), 2015. [2]

[KR08]   Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within 2-epsilon. *J. Comput. Syst. Sci.*, 74(3):335–349, 2008. [7]

[Kri77]   Jean-Louis Krivine. Sur la constante de grothendieck. *C. R. Acad. Sci. Paris Sér. A-B*, 284(8):A445A446, 1977. [2]

[KRT08]   Julia Kempe, Oded Regev, and Ben Toner. Unique games with entangled provers are easy. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 457–466, 2008. [7,9]

[LS09]   Nati Linial and Adi Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Structures and Algorithms*, 34:368–394, 2009. [2]

[PGWP+08] David Pérez-García, Michael Wolf, Carlos Palazuelos, Ignacio Villanueva, and Marius Junge. Unbounded violation of tripartite Bell inequalities. *Communications in Mathematical Physics*, 279:455, 2008. [2,3]

[Raz99]   Ran Raz. Exponential separation of quantum and classical communication complexity. In *31st annual ACM symposium on theory of computing*, pages 358–367, 1999. [2]

[SZ08]   Yaoyun Shi and Yufan Zhu. Tensor norms and the classical communication complexity of nonlocal quantum measurement. *SIAM J. Comput.*, 38(3):753–766, 2008. [2]

[Tao12]   Terence Tao. *Higher order Fourier analysis*, volume 142. American Mathematical Soc., 2012. [20,22]

[Tsi87]   Boris S. Tsirelson. Quantum analogues of the Bell inequalities. The case of two spatially separated domains. *J. Soviet Math.*, 36:557–570, 1987. [2]

[TZ12]   Terence Tao and Tamar Ziegler. The inverse conjecture for the Gowers norm over finite fields in low characteristic. *Annals of Combinatorics*, 16(1):121–188, 2012. [9,22]

[WHKN18]   Adam Bene Watts, Aram W Harrow, Gurtej Kanwar, and Anand Natarajan. Algorithms, bounds, and strategies for entangled XOR games. *arXiv preprint arXiv:1801.00821*, 2018. [3,4]

[Zuk93]   M. Zukowski. Bell theorem involving all settings of measuring apparatus. *Phys. Lett. A*, 177(290), 1993. [3]