

# Explicit lower and upper bounds on the entangled value of multiplayer XOR games

Jop Briët\*

Thomas Vidick†

## Abstract

XOR games are the simplest model in which the nonlocal properties of entanglement manifest themselves. When there are two players, it is well known that the bias — the maximum advantage over random play — of entangled players can be at most a constant times greater than that of classical players. Recently, Pérez-García et al. [Comm. Math. Phys. 279 (2), 2008] showed that no such bound holds when there are three or more players: the advantage of entangled players over classical players can become unbounded, and scale with the number of questions in the game. Their proof relies on non-trivial results from operator space theory, and gives a non-explicit existence proof, leading to a game with a very large number of questions and only a loose control over the local dimension of the players' shared entanglement.

We give a new, simple and explicit (though still probabilistic) construction of a family of three-player XOR games which achieve a large quantum-classical gap (QC-gap). This QC-gap is exponentially larger than the one given by Pérez-García et al. in terms of the size of the game, achieving a QC-gap of order  $\sqrt{N}$  with  $N^2$  questions per player. In terms of the dimension of the entangled state required, we achieve the same (optimal) QC-gap of  $\sqrt{N}$  for a state of local dimension  $N$  per player. Moreover, the optimal entangled strategy is very simple, involving observables defined by tensor products of the Pauli matrices.

Additionally, we give the first upper bound on the maximal QC-gap in terms of the number of questions per player, showing that our construction is only quadratically off in that respect. Our results rely on probabilistic estimates on the norm of random matrices and higher-order tensors which may be of independent interest.

## 1 Introduction

Multiplayer games, already a very successful abstraction in theoretical computer science, were first proposed as an ideal framework in which to study the nonlocal properties of entanglement by Cleve et al. [CHTW04]. Known as *nonlocal*, or *entangled*, games, they can be thought of as an interactive re-framing of the familiar setting of Bell inequalities: a referee (the experimentalist) interacts with a number of players (the devices). The referee first sends a classical question (a

---

\*CWI. Supported by an NWO Vici grant and the EU project QCS. Email: j.briet@cwi.nl.

†Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology. Supported by the National Science Foundation under Grant No. 0844626. Email: vidick@csail.mit.edu

setting) to each player. The players are all-powerful (there is no restriction on the shared state or the measurements applied) but not allowed to communicate: each of them must make a local measurement on his or her part of a shared entangled state, and provide a classical answer (the outcome) to the referee’s question. The referee then decides whether to accept or reject the players’ answers (he evaluates the Bell functional).

In their paper, Cleve et al. gave an in-depth study of the simplest class of multiplayer games, *two-player XOR games*. The XOR property refers to the fact that in such games each player answers with a single bit, and the referee’s acceptance criterion only depends on the parity of the bits he receives as answers. One of the most fundamental Bell inequalities, the CHSH inequality [CHSH69], fits in this framework. In the corresponding XOR game the acceptance criterion dictates that the parity of the players’ answers must equal the product of their questions, a uniform i.i.d. bit each. The laws of quantum mechanics predict that the CHSH game has the following striking property: there is a quantum strategy in which the players share a simple entangled state — a single EPR pair — and use it to achieve a strictly higher success probability than the best classical, unentangled strategy: roughly 85%, as compared to 75%. This example demonstrates that quantum mechanics is *nonlocal*: predictions made by the theory cannot be reproduced classically, or more generally by any local hidden variable model, a “paradox” most famously put forward by Einstein, Podolsky and Rosen [EPR35].

Any XOR game  $G$  can be won with probability  $1/2$  by players who independently answer each question with the outcome of a random coin flip. It is therefore natural to measure the success of quantum (resp. classical) players through their maximum achievable *bias*  $\beta^*(G)$  (resp.  $\beta(G)$ ), defined as their maximum winning probability in the game, *minus* the success probability that would be achieved by random play. As has become standard practice, we will measure the advantage of quantum over classical players through the ratio  $\beta^*(G)/\beta(G)$ , referred to as the quantum-classical gap, or *QC-gap* for short.<sup>1</sup> The CHSH example demonstrates the existence of a game for which  $\beta^*(G) \geq \sqrt{2}\beta(G)$ , and Tsirelson [Tsi87] proved that this gap was close to best possible. By making a connection to the celebrated *Grothendieck inequality* he showed that for any two-player XOR game  $G$ , we have  $\beta^*(G)/\beta(G) \leq K_G^{\mathbb{R}}$ , where  $K_G^{\mathbb{R}}$  is the real Grothendieck constant.<sup>2</sup> The exact value of  $K_G^{\mathbb{R}}$  is unknown, and the best upper bound currently known,  $K_G^{\mathbb{R}} \lesssim 1.78$ , appeared in recent work of Braverman et. al. [BMMN11]. Although experiments based on the CHSH game have been performed [AGR81, ADR82], the relatively small gap forces the use of state-of-the-art devices in terms of precision and timing in order to differentiate a truly nonlocal strategy from one that can be explained by local hidden-variable models. In order to observe larger quantum-classical gaps, more general classes of games need to be considered, prompting a question that has driven much recent research in this area: *For a given QC-gap, what is the simplest game (in terms of the number of players, questions and answers) which demonstrates such a QC-gap (if one at all exists)?*

There are two main directions in which one can look for generalizations of two-player XOR games. The first is to increase the number of possible answers from each player. This option has

---

<sup>1</sup>See Section 5 for a brief discussion of other ways of measuring the quantum advantage, such as through the difference  $\beta^*(G) - \beta(G)$ .

<sup>2</sup>The subscript  $G$  in  $K_G^{\mathbb{R}}$  stands for “Grothendieck”, and is not related to the game  $G$ !

so far been the preferred one, and has by now been relatively well explored [CHTW04, KRT10, JPPG<sup>+</sup>10, JP11, Reg11, BRSdW11]. In particular it is known that the largest possible quantum-classical gap is bounded by a constant times the minimum of the number of questions, the number of answers, and the local dimension of the players [JPPG<sup>+</sup>10], and there are explicit constructions of games (i.e., games whose existence is proved through a constructive proof) which come close to achieving these bounds [BRSdW11]. Unfortunately, these games require the players to perform complex measurements, involving large numbers of outcomes, making them ill-suited to experiment.

The second possible avenue for generalization consists in increasing the number of players, while remaining in the simple setting of binary answers and an XOR-based acceptance criterion. Our limited understanding of multipartite entanglement makes this setting more challenging, and for a long time little more than small, constant-size examples were known [Mer90, Zuk93]. However, recently, Pérez-García et al. [PGWP<sup>+</sup>08] discovered that adding even just one player allowed for a very different scaling of the QC-gap. They demonstrated the existence of an infinite family of three-player XOR games  $(G_N)_{N \in \mathbb{N}}$  for which  $\lim_{N \rightarrow \infty} \beta^*(G_N) / \beta(G_N) = +\infty$  — an unbounded gap! This exciting result demonstrated for the first time that very large violations could be observed even in the relatively simple context of three-player XOR games.

The results in [PGWP<sup>+</sup>08] were proved by establishing a surprising connection between XOR games and certain natural norms on the tensor product of operator spaces, enabling the authors to leverage powerful techniques from the latter area in order to establish their results on XOR games. Since their seminal paper, similar techniques have been successfully applied to other settings, such as general two-player games [JPPG<sup>+</sup>10] and games with quantum communication [CJPPG11].

For the games  $G_N$  from [PGWP<sup>+</sup>08], however, the above-mentioned techniques have a few somewhat unfortunate consequences. First of all, these techniques resulted in a highly non-explicit existence proof. While Pérez-García et al. show the *existence* of the games, it seems quite hard to even get the slightest idea of what the games would look like. Moreover, their use of the theory of operator spaces gives a very large game, with an exponential (in the QC-gap) number of questions per player. Finally, the strategies required of the players to achieve the promised QC-gap are not explicitly known, and may for instance require an entangled state with unbounded dimension on two of the players; only the first player’s dimension is controlled. We note that after the completion of our work, but independently from it, Pisier [Pis12a, Pis12b] showed that the construction in [PGWP<sup>+</sup>08] could be improved to require only a polynomial number of questions to each player, and that one could keep a control of the entanglement dimension on all three players. The resulting parameters, however, are still worse than the ones that we achieve here.

## 1.1 Our results

In this paper we give a new and improved proof of the existence of a family of three-player XOR games for which the QC-gap is unbounded. Our proof technique uses the probabilistic method: we describe a simple *probabilistic* procedure that outputs a game with the desired properties *with high probability*. As such it is much more explicit than previous results [PGWP<sup>+</sup>08], albeit not

fully constructive. Our construction is outlined in Section 1.2 below. For a desired ratio  $\sqrt{N}$ , our game has order  $N^2$  questions per player, which, as we show, is within a factor  $\tilde{O}(N)$  of the smallest number possible. Moreover, to achieve such a gap entangled players only need to use Pauli observables and an entangled state of local dimension  $N$  per player. The simplicity of our construction enables us to give concrete values for most of the parameters, leading to a rigorous control of the constants involved. We prove the following:

**Theorem 1.** *For any integer  $n$  and  $N = 2^n$  there exists a three-player XOR game  $G_N$ , with  $N^2$  questions per player, such that  $\beta^*(G_N) \geq \Omega(\sqrt{N} \log^{-5/2} N) \beta(G_N)$ . Moreover, there is an entangled strategy which achieves a bias of  $\Omega(\sqrt{N} \log^{-5/2} N) \beta(G_N)$ , uses an entangled state of local dimension  $N$  per player, and in which the players' observables are tensor products of  $n$  Pauli matrices.*

Additionally, we prove that the dependence of the QC-gap on the number of questions obtained in Theorem 1 is close to optimal.<sup>3</sup> This improves upon an independent previous result by Loubenets [Lou12], who showed that  $\beta^*(G) \leq (2Q - 1)^2 \beta(G)$ .

**Theorem 2.** *For any 3-player XOR game  $G$  in which there are at most  $Q$  possible questions to the third player,*

$$\beta^*(G) \leq \sqrt{Q} K_G^{\mathbb{R}} \beta(G),$$

where  $K_G^{\mathbb{R}} < 1.783$  is the real Grothendieck constant.

Finally, we also show that the dependence on the local dimension of the entangled state is optimal, re-proving in a simpler language a result first proved in [PGWP<sup>+</sup>08].

**Theorem 3.** *Let  $G$  be a 3-player XOR game in which the maximal entangled bias  $\beta^*(G)$  is achieved by a strategy in which the third player's local dimension is  $d$ . Then*

$$\beta^*(G) \leq \sqrt{3d} (K_G^{\mathbb{C}})^{3/2} \beta(G),$$

where  $K_G^{\mathbb{C}} < 1.405$  is the complex Grothendieck constant.

**Generalizations.** While we present our results in the case of three-player XOR games, they have straightforward extensions to an arbitrary number of players. In particular, one can show that the following holds, for any  $r \geq 3$ :

1. For any integer  $N$  that is a power of 2, there exists a  $r$ -player XOR game  $G$ , with  $N^2$  questions per player, such that  $\beta^*(G) \geq \Omega((N \log^{-5} N)^{(r-2)/2}) \beta(G)$ , and there is an entangled strategy achieving this gap that involves only  $N$ -dimensional Pauli observables.
2. If  $G$  is a  $r$ -player XOR game in which at least  $r - 2$  of the players have at most  $Q$  possible questions each, then  $\beta^*(G) \leq O(Q^{(r-2)/2}) \beta(G)$ .
3. If  $G$  is a  $r$ -player XOR game in which the shared state of the players is restricted to have local dimension  $d$  on at least  $r - 2$  of the players, then  $\beta^*(G) \leq O(d^{(r-2)/2}) \beta(G)$ .

---

<sup>3</sup>A similar result was recently communicated to us by Carlos Palazuelos [Pal11].

**Applications to operator space theory.** The original motivation for the construction in [PGWP<sup>+</sup>08] was to show that a certain trilinear extension of Grothendieck’s inequality does not hold. Our construction leads to an improved obstruction: the three-player XOR game constructed in Theorem 1 can be used to show that two different norms are not equivalent for the space of trilinear functionals on  $\ell_\infty \times \ell_\infty \times \ell_\infty$ . More precisely, Theorem 1 implies that for any  $N = 2^n$  there is a trilinear functional  $T : \ell_\infty^{N^2} \times \ell_\infty^{N^2} \times \ell_\infty^{N^2} \rightarrow \mathbb{C}$  such that

$$\|T\|_{\text{cb}} \geq \Omega(\sqrt{N} \log^{-5/2} N) \|T\|, \quad (1)$$

and the  $N$ th amplifications in the completely bounded norm suffice. Put differently, the injective and minimal tensor norms are inequivalent on  $\ell_1 \otimes \ell_1 \otimes \ell_1$ . This improves on the estimate from [PGWP<sup>+</sup>08], in which the bound was logarithmic in  $N$ . For more details and background on relevant aspects of Grothendieck’s inequality we refer to the excellent survey [Pis12a], and to Section 20 in particular for the connection with XOR games.

In addition, Pisier [Pis12b] recently applied our result to prove an almost-tight estimate on the norm of the re-ordering map

$$J : (H_1 \otimes_2 K_1) \otimes_\epsilon \cdots \otimes_\epsilon (H_\ell \otimes_2 K_r) \rightarrow (H_1 \otimes_2 \cdots \otimes_2 H_r) \otimes_\epsilon (K_1 \otimes_2 \cdots \otimes_2 K_r), \quad (2)$$

where  $H_i, K_i$  are  $N$ -dimensional Hilbert spaces, proving that  $\|J\| = \tilde{\Omega}(N^{r-1})$ , where the  $\tilde{\Omega}$  notation ignores possible poly-logarithmic factors.

## 1.2 Proof overview and techniques

**Lower bound.** Our construction of a three-player XOR game  $G_N$  proceeds through two independent steps. In the first step we assume given a 3-tensor  $T = T_{(i,i'),(j,j'),(k,k')}$  of dimension  $N^2 \times N^2 \times N^2$ , where  $N$  is a power of 2. Based on  $T$ , we define a three-player XOR game  $G_N = G(T)$ . Questions in this game are  $N$ -dimensional Pauli matrices  $P, Q, R$ , and the corresponding game coefficient<sup>4</sup> is defined as

$$G(P, Q, R) = \langle T, P \otimes Q \otimes R \rangle := \sum_{(i,i'),(j,j'),(k,k')} T_{(i,i'),(j,j'),(k,k')} P_{i,i'} Q_{j,j'} R_{k,k'}.$$

This definition results in a game whose entangled and classical biases can be directly related to *spectral properties* of the tensor  $T$ . On the one hand we show that the classical bias  $\beta(G_N)$  reflects the tripartite structure of  $T$ , and is upper-bounded by the norm of  $T$  as a trilinear operator. On the other hand we show that the entangled bias  $\beta^*(G_N)$  is *lower-bounded* by the norm of  $T$  as a *matrix* — a bilinear operator on  $N^3$ -dimensional vectors, obtained by pairing up the indices  $(i, j, k)$  and  $(i', j', k')$ . This new connection reduces the problem of constructing a game with large QC-gap to constructing a tensor  $T$  with appropriate spectral properties.

The second step of the proof is our main technical contribution. We give a probabilistic construction of a 3-tensor  $T$  having large norm when seen as a bilinear operator (giving a large entangled bias), but low norm when seen as a trilinear operator (giving a low classical bias). To this

---

<sup>4</sup>The equation below defines a complex number. Taking its real or imaginary part would result in a Bell functional, which can in turn easily be transformed into an XOR game through a proper normalization.

end, we simply take  $T$  to correspond to an (almost) rank-1 matrix: letting  $(g_{ijk})$  be a random  $N^3$ -dimensional vector with i.i.d. entries distributed as standard Gaussians,<sup>5</sup> the  $(i, i'), (j, j'), (k, k')$ -th entry of  $T$  is  $g_{ijk} g_{i'j'k'}$  if  $i \neq i', j \neq j'$  and  $k \neq k'$ , and 0 otherwise. The fact that  $T$ , when seen as a matrix, is close to having rank 1 makes it easy to lower bound its spectral norm. An upper bound on the norm of  $T$  as a trilinear operator is proved in two steps. In the first step we apply a concentration bound due to Latała to show that for any fixed Hermitian  $X, Y, Z$  with Frobenius norm at most 1, the product  $|\langle T, X \otimes Y \otimes Z \rangle|$  is highly concentrated around its expected value, where the concentration is over the random choice of  $T$ . We then conclude by a union bound, using a delicate  $\varepsilon$ -net construction based on a decomposition of Hermitian matrices with Frobenius norm at most 1 as linear combinations of (normalized, signed) projectors.

**Upper bounds.** We prove upper bounds on the largest possible QC-gap achievable by any three-player XOR game, both as a function of the local dimension of an optimal strategy, and of the number of questions per player in the game. Both bounds follow the same overall proof strategy: using a decoupling argument, we show that the third player can be restricted to applying a classical strategy while incurring only a bounded factor loss in the bias. We conclude by applying (the easy direction of) Tsirelson’s Theorem and Grothendieck’s inequality (see Section 2.6) to show that the first two players can be made classical at a further loss of a constant factor only.

**Organization of the paper.** We start with some preliminaries in Section 2. We describe our construction of a game with unbounded QC-gap in Section 3. Our upper bounds on the QC-gap as a function of the number of questions and the local dimension are proved in Section 4. We conclude with some open questions in Section 5.

## 2 Preliminaries

### 2.1 Notation

For a positive integer  $N$  we define  $[N] := \{1, \dots, N\}$ . For a positive integer  $K$  we denote by  $[N]^K$  the Cartesian product of the set  $[N]$  with itself  $K$  times (i.e.,  $[N] \times \dots \times [N]$ ).

For a subset  $\mathcal{W} \subseteq \mathcal{V}$  of a normed vector space  $(\mathcal{V}, \|\cdot\|)$  we let  $S(\mathcal{W}) := \{X \in \mathcal{W} : \|X\| = 1\}$  be the unit sphere,  $B(\mathcal{W}, \tau) := \{X \in \mathcal{W} : \|X\| \leq \tau\}$  the ball of radius  $\tau$  and  $B(\mathcal{W}) := B(\mathcal{W}, 1)$  the unit ball. We let  $\|\cdot\|_2$  denote the usual Euclidean norm. Throughout we endow  $\mathbb{C}^N$  with this norm.

We will usually use  $g \sim \mathcal{N}(0, 1)$  to denote a real-valued random variable distributed according to a standard normal (Gaussian) distribution (i.e., a variable with mean 0 and variance 1), and  $|g\rangle \sim \mathcal{N}(0, 1)^N$  for an  $N$ -dimensional vector whose entries are i.i.d. standard normal random variables.

---

<sup>5</sup>Our results also hold with the Gaussians replaced by i.i.d. standard Bernoulli random variables.

**Matrices.** Throughout  $\mathcal{H}$  will denote a  $N$ -dimensional complex Hilbert space. We identify the set of linear operators  $L(\mathcal{H})$  on  $\mathcal{H}$  with the set  $\text{Mat}(N)$  of complex  $N$ -by- $N$  matrices. Let  $\text{Herm}(\mathcal{H}) = \{X \in L(\mathcal{H}) : X^\dagger = X\}$  be the subset of Hermitian operators and  $\text{Obs}(\mathcal{H}) \subseteq \text{Herm}(\mathcal{H})$  be the Hermitian operators with all eigenvalues in  $\{-1, 1\}$ . In other words,  $\text{Obs}(\mathcal{H})$  is the set of  $\{-1, 1\}$ -valued observables on  $\mathcal{H}$ . Note that operators in  $\text{Obs}(\mathcal{H})$  are unitary and square to the identity. We will use the notation  $\text{Herm}(N)$  and  $\text{Obs}(N)$  when we think of the operators' matrix representation. The space of matrices  $\text{Mat}(N)$  is a Hilbert space for the inner product  $(A, B) \mapsto \langle A, B \rangle := \text{Tr}(AB^\dagger)$ . The resulting norm is the *Frobenius norm*  $A \mapsto \|A\|_F := \sqrt{\text{Tr}(AA^\dagger)}$ . Throughout we tacitly endow  $\text{Mat}(N)$  with the Frobenius norm, so the balls and sphere are always defined with respect to this norm. Note that if we let the singular values of a matrix  $X \in \text{Mat}(N)$  be  $\sigma_1(X) \geq \dots \geq \sigma_N(X)$ , then  $\|X\|_F^2 = \sigma_1(X)^2 + \dots + \sigma_N(X)^2$ . We recall that for each eigenvalue  $\lambda$  of a Hermitian matrix  $X$  there is a corresponding singular value  $\sigma = |\lambda|$ . We denote by  $\|\cdot\|_\infty = \sigma_1(X)$  the operator norm on  $\text{Mat}(N)$ . Let  $\text{Proj}(N)_k \subseteq \text{Herm}(N)$  be the set of rank- $k$  (orthogonal) projectors on  $\mathbb{C}^N$  and let  $\overline{\text{Proj}}(N)_k = \text{Proj}(N)_k / \sqrt{k}$  be the set of rank- $k$  projectors that are normalized with respect to the Frobenius norm. Define the set of all  $N$ -dimensional normalized projectors by  $\overline{\text{Proj}}(N) = \bigcup_{k=1}^N \overline{\text{Proj}}(N)_k$ .

If  $N = 2^n$  for some positive integer  $n$ , we let  $\mathcal{P}_n := \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}^{\otimes n}$  be the set of  $n$ -fold tensor products of Pauli matrices. The letters  $P, Q, R$  will usually denote elements of  $\mathcal{P}_n$ . We have  $|\mathcal{P}_n| = N^2$ , and for  $P, Q \in \mathcal{P}_n$  we have  $\langle P, Q \rangle = N \delta_{P, Q}$ : the set  $\mathcal{P}_n$  forms an orthogonal basis of observables for  $\text{Mat}(N)$ .

## 2.2 Tensors

Given positive integers  $r, N_1, \dots, N_r$ , an  $r$ -tensor of dimensions  $N_1 \times \dots \times N_r$  is a map of the form  $T : [N_1] \times \dots \times [N_r] \rightarrow \mathbb{C}$ . Every element  $T(i_1, \dots, i_r)$  of such a tensor is specified by an  $r$ -tuple of indices  $(i_1, \dots, i_r) \in [N_1] \times \dots \times [N_r]$ . We will mostly deal with 3-tensors of dimensions  $N^2 \times N^2 \times N^2$  for some  $N \in \mathbb{N}$ . In this case we index the elements by three pairs of indices  $(i, i')$ ,  $(j, j')$  and  $(k, k') \in [N]^2$ . We will think of such a tensor in two different ways: as a *bilinear* functional acting on  $N^3$ -dimensional complex vectors, and as a *trilinear* functional acting on Hermitian  $N \times N$  matrices. For the sake of concreteness we now describe in detail how these two perspectives relate to each other.

**The bilinear view.** Let  $T$  be a 3-tensor of dimension  $N^2 \times N^2 \times N^2$ . The dimensions of the tensor  $T$  allow us to view it as an  $N^3$ -by- $N^3$  complex matrix. Correspondingly, we define the spectral norm of  $T$  by

$$\|T\|_{3,3} := \max_{x, y \in \mathcal{S}(\mathbb{C}^{N^3})} \left| \sum_{(i,j,k),(i',j',k') \in [N]^3} T_{(i,i'),(j,j'),(k,k')} x_{i,j,k} y_{i',j',k'} \right|.$$

Suppose that for some  $n \in \mathbb{N}$ , we have  $N = 2^n$ . Since the set  $\mathcal{P}_{N^3} = \{X \otimes Y \otimes Z : X, Y, Z \in \mathcal{P}_N\}$  is an orthogonal basis for  $\text{Mat}(N^3)$ , we can define the "Fourier coefficient" of  $T$  at  $(P, Q, R)$  as

$$\widehat{T}(P, Q, R) := \langle T, P \otimes Q \otimes R \rangle = \sum_{(i,i'),(j,j'),(k,k') \in [N]^2} T_{(i,i'),(j,j'),(k,k')} P_{i,i'} Q_{j,j'} R_{k,k'}.$$

With this definition,  $T$  can be written as

$$T = N^{-3} \sum_{P, Q, R \in \mathcal{P}_N} \widehat{T}(P, Q, R) P \otimes Q \otimes R.$$

**The trilinear view.** Let  $T$  be a 3-tensor of dimensions  $N^2 \times N^2 \times N^2$ . We can associate with  $T$  a trilinear functional  $L_T : \text{Herm}(N) \times \text{Herm}(N) \times \text{Herm}(N) \rightarrow \mathbb{C}$  defined by

$$L_T(X, Y, Z) = \langle T, X \otimes Y \otimes Z \rangle = \sum_{(i,j,k),(i',j',k') \in [N]^3} T_{(i,i'),(j,j'),(k,k')} X_{i,i'} Y_{j,j'} Z_{k,k'},$$

where  $X, Y, Z \in \text{Herm}(N)$ . The operator norm of  $L_T$  induces the following norm on  $T$

$$\|T\|_{2,2,2} := \max_{X, Y, Z \in \mathcal{B}(\text{Herm}(N))} |L_T(X, Y, Z)| = \max_{X, Y, Z \in \mathcal{B}(\text{Herm}(N))} |\langle T, X \otimes Y \otimes Z \rangle|.{}^6$$

### 2.3 XOR games

An  $r$ -player XOR game with  $N$  questions per player is fully specified by a joint probability distribution  $\pi$  on  $[N]^r$  and an  $r$ -tensor  $M : [N]^r \rightarrow \{-1, 1\}$ . The *classical bias* of an XOR game  $G = (\pi, M)$  is defined by

$$\beta(G) := \max_{\chi_1, \dots, \chi_r : [N] \rightarrow \{-1, 1\}} \mathbb{E}_{(q_1, \dots, q_r) \sim \pi} \left[ M(q_1, \dots, q_r) \chi_1(q_1) \cdots \chi_r(q_r) \right].$$

The maps  $\chi_1, \dots, \chi_r$  in the above maximum are referred to as *strategies*: they should be interpreted as giving the players' answers to the questions  $q_1, \dots, q_r$ , respectively. The *entangled bias* of  $G$  is defined by

$$\beta^*(G) := \sup_{\substack{d \in \mathbb{N}, |\Psi\rangle \in \mathcal{S}(\mathbb{C}^{d^r}) \\ A_1, \dots, A_r : [N] \rightarrow \text{Obs}(d)}} \mathbb{E}_{(q_1, \dots, q_r) \sim \pi} \left[ M(q_1, \dots, q_r) \langle \psi | A_1(q_1) \otimes \cdots \otimes A_r(q_r) | \psi \rangle \right].$$

In the sequel it will be convenient to merge  $\pi$  and  $M$  into a single tensor  $T : [N]^r \rightarrow \mathbb{R}$  defined by  $T(q_1, \dots, q_r) = \pi(q_1, \dots, q_r) M(q_1, \dots, q_r)$ . Conversely, any tensor  $T : [N]^r \rightarrow \mathbb{R}$  defines (up to normalization) an XOR game by setting the distribution to  $\pi(q_1, \dots, q_r) = |T(q_1, \dots, q_r)|$  and the game tensor to  $M(q_1, \dots, q_r) = \text{sign}(T(q_1, \dots, q_r))$ .

### 2.4 $\varepsilon$ -nets

Our probabilistic proof of the existence of a game for which there is a large QC-gap relies on the construction of specific  $\varepsilon$ -nets over Hermitian matrices, which we describe in this section.

**Definition 4.** An  $\varepsilon$ -net for a subset  $\mathcal{W}$  of a metric space  $(\mathcal{V}, d)$  is a finite set  $W \subseteq \mathcal{V}$  such that for every  $x \in \mathcal{W}$ , there exists an  $s \in W$  such that  $d(x, s) \leq \varepsilon$ .

<sup>6</sup>The restriction to Hermitian matrices in this definition is not essential, but it will be convenient later on.



**Fact 5.** For every  $N \in \mathbb{N}$  and any  $\varepsilon > 0$  there exists an  $\varepsilon$ -net  $S_\varepsilon$  for  $S(\mathbb{C}^N)$  of cardinality  $|S_\varepsilon| \leq (1 + 2/\varepsilon)^N$ .

*Proof.* This well-known fact follows from a volume argument: we can choose  $S_\varepsilon$  so that the balls with radius  $\varepsilon/2$  centered at the points in  $S_\varepsilon$  are disjoint. (See e.g. [Pis99, Lemma 4.10].)  $\square$

The following lemma shows that for any  $\varepsilon > 0$  and  $N > 1$ , an  $\varepsilon/(4\sqrt{\ln N})$ -net over (normalized, signed)  $N$ -dimensional projections automatically induces an  $\varepsilon$ -net over  $N$ -dimensional Hermitian matrices with Frobenius norm at most 1. The lemma follows from a well known equivalence between the unit ball of normalized projections and the unit ball corresponding to the matrix norm derived from the Lorentz-sequence semi-norm  $\ell_{2,1}$ . We give a self-contained proof below.

**Lemma 6.** Let  $N > 1$  and  $X \in B(\text{Herm}(N))$ . Then  $X$  can be decomposed as a linear combination

$$X = \sum_x \lambda_x X_x,$$

where each  $X_x \in \overline{\text{Proj}}(N)$  is a normalized projector and  $\sum_x |\lambda_x| \leq 4\sqrt{\ln N}$ .

*Proof.* Let  $X = \sum_i \lambda_i |u_i\rangle\langle u_i|$  be the spectral decomposition of a Hermitian matrix  $X$  with norm

$$\|X\|_F^2 = \sum_i \lambda_i^2 \leq 1. \quad (3)$$

For every  $t \in [-1, 1]$ , let  $P_t$  be the projector on  $\text{Span}\{|u_i\rangle : \lambda_i \in [-1, -t]\}$  if  $t < 0$  and the projector on  $\text{Span}\{|u_i\rangle : \lambda_i \in (t, 1]\}$  if  $t \geq 0$ . Then the following holds:

$$X = \int_{t=-1}^1 \text{sign}(t) P_t dt = \int_{t=-1}^1 \sqrt{\text{rank}P_t} \frac{\text{sign}(t) P_t}{\sqrt{\text{rank}P_t}} dt, \quad (4)$$

where the integral is taken coefficient-wise.<sup>7</sup> By a direct calculation,

$$\int_{t=-1}^1 |t| \text{Tr}(P_t) dt = \frac{1}{2} \sum_i \lambda_i^2 \leq \frac{1}{2}, \quad (5)$$

where the last equality follows from (3). Eq. (4) shows that  $X$  may be written as a non-negative linear combination of the  $\text{sign}(t) P_t / \sqrt{\text{rank}P_t}$  with coefficients summing up to

$$\begin{aligned} \int_{t=-1}^1 \sqrt{\text{rank}P_t} dt &\leq \int_{t=-1/\sqrt{N}}^{1/\sqrt{N}} \sqrt{N} dt + \int_{t=-1}^{-1/\sqrt{N}} \sqrt{\text{rank}P_t} dt + \int_{t=1/\sqrt{N}}^1 \sqrt{\text{rank}P_t} dt \\ &\leq 2 + \left(2 \int_{t=1/\sqrt{N}}^1 \frac{1}{t} dt\right)^{1/2} \left(\int_{t=-1}^{-1/\sqrt{N}} (-t) \text{rank}P_t dt + \int_{t=1/\sqrt{N}}^1 t \text{rank}P_t dt\right)^{1/2} \\ &\leq 2 + \sqrt{\ln N/2}, \end{aligned}$$

where the first inequality uses  $\text{rank}P_t \leq \min(N, t^{-2})$  for every  $t$ , the second inequality follows from Cauchy-Schwarz and the last uses (5), together with  $\text{rank}P_t = \text{Tr}P_t$ .  $\square$

<sup>7</sup>The coefficients of  $P_t$  are step functions, so the integral is well-defined.

The following lemma gives a straightforward construction of an  $\varepsilon$ -net for the set of normalized rank- $k$  projectors on  $\mathbb{C}^d$  (see e.g. [Sza82] for more general constructions of nets on Grassmannian spaces).

**Lemma 7.** *For every  $k \in [N]$  and any  $0 < \varepsilon \leq 1$  there exists a set*

$$\mathcal{Z}_\varepsilon^k \subseteq \bigcup_{\ell=1}^k \overline{\text{Proj}}(N)_\ell$$

of size  $|\mathcal{Z}_\varepsilon^k| \leq 2(5/\varepsilon)^{kN}$ , such that for any  $X \in \overline{\text{Proj}}(N)_k$  there is an  $\tilde{X} \in \mathcal{Z}_\varepsilon^k$  satisfying  $\|X - \tilde{X}\|_F \leq \varepsilon$ .

*Proof.* Let  $\eta = \varepsilon/\sqrt{2}$  and  $S_\eta$  be an  $\eta$ -net for the unit sphere  $S(\mathbb{C}^N)$ . For every  $k$ -subset  $T \subseteq S_\eta$  let  $Y_T$  be the projector on the space spanned by the vectors in  $k$  and let  $\bar{Y}_T = Y_T/\sqrt{\text{rank}Y}$ . Define the set  $\mathcal{Z}_\varepsilon^k$  by

$$\mathcal{Z}_\varepsilon^k = \{\bar{Y}_T : T \subseteq S_\eta, |T| = k\}.$$

Note that for any  $\bar{Y} \in \mathcal{Z}_\varepsilon^k$ , we have  $\text{rank}(\bar{Y}) \leq k$ . Moreover, by the upper bound on the minimal size of  $S_\eta$  from Fact 5, we have

$$|\mathcal{Z}_\varepsilon^k| \leq \binom{|S_\eta|}{k} \leq \binom{(3/\eta)^N}{k} \leq 2 \left(\frac{5}{\varepsilon}\right)^{kN}.$$

Fix  $X \in \overline{\text{Proj}}(N)_k$  and let  $|\phi_1\rangle, \dots, |\phi_k\rangle \in S(\mathbb{C}^N)$  be orthonormal eigenvectors of  $X$  with eigenvalue  $1/\sqrt{k}$ . Let  $|\psi_1\rangle, \dots, |\psi_k\rangle \in S_\eta$  be the vectors closest to  $|\phi_1\rangle, \dots, |\phi_k\rangle$  (resp.) with respect to the Euclidean distance. Let  $Y$  be the projector on the space spanned by the  $|\psi_1\rangle, \dots, |\psi_k\rangle$  and let  $\bar{Y} = Y/\sqrt{\text{rank}Y}$ . Clearly,  $\bar{Y} \in \mathcal{Z}_\varepsilon^k$ . Since  $Y$  is positive semidefinite and for every  $i = 1, \dots, k$ , the vector  $|\psi_i\rangle$  is an eigenvector of  $Y$  with eigenvalue 1,

$$\langle \phi_i | Y | \phi_i \rangle \geq |\langle \phi_i | \psi_i \rangle|^2 \geq 1 - \eta^2,$$

where the second inequality follows since  $|\psi_i\rangle$  is closest to  $|\phi_i\rangle$  in the  $\varepsilon$ -net.<sup>8</sup> By definition of the Frobenius norm and the fact that  $X$  and  $\bar{Y}$  are Hermitian, we get

$$\begin{aligned} \|X - \bar{Y}\|_F^2 &= \|X\|_F^2 + \|\bar{Y}\|_F^2 - 2\text{Tr}(X\bar{Y}) \\ &\leq 2 - 2\text{Tr}(X\bar{Y}) \\ &\leq 2 \left(1 - \frac{1}{k} \sum_{i=1}^k \langle \phi_i | Y | \phi_i \rangle\right) \\ &\leq 2 \left(1 - (1 - \eta^2)\right) \\ &= \varepsilon^2, \end{aligned}$$

and the lemma is proved. □

---

<sup>8</sup>Notice that for any complex unit vectors  $x, y$ , we have  $\|x - y\|^2 = 2 - 2\Re(\langle x, y \rangle)$  and  $|\langle x, y \rangle|^2 = \Re(\langle x, y \rangle)^2 + \Im(\langle x, y \rangle)^2$ .

**Definition 8.** For every triple of integers  $(k, \ell, m) \in [N]^3$  and any real number  $0 < \varepsilon \leq 1$ , define

$$\mathcal{Z}_\varepsilon^{(k, \ell, m)} = \{X \otimes Y \otimes Z : (X, Y, Z) \in \mathcal{Z}_\varepsilon^k \times \mathcal{Z}_\varepsilon^\ell \times \mathcal{Z}_\varepsilon^m\},$$

and  $\mathcal{Z}_\varepsilon = \bigcup_{(k, \ell, m) \in [N]^3} \mathcal{Z}_\varepsilon^{(k, \ell, m)}$ .

**Proposition 9.** For any  $\varepsilon > 0$  and  $d > 1$ , the set  $\mathcal{Z}_\varepsilon$  is a  $3\varepsilon$ -net for the set of matrices  $X \otimes Y \otimes Z$  where  $(X, Y, Z) \in \overline{\text{Proj}}(N) \times \overline{\text{Proj}}(N) \times \overline{\text{Proj}}(N)$ , with respect to the distance function defined by the Frobenius norm.

*Proof.* Let  $1 \leq k, \ell, m \leq N$  and  $X \in \overline{\text{Proj}}(N)_k$ ,  $Y \in \overline{\text{Proj}}(N)_\ell$  and  $Z \in \overline{\text{Proj}}(N)_m$ . Let  $\tilde{X} \in \mathcal{Z}_\varepsilon^k$ ,  $\tilde{Y} \in \mathcal{Z}_\varepsilon^\ell$  and  $\tilde{Z} \in \mathcal{Z}_\varepsilon^m$  be the closest elements in the nets to (resp.)  $X, Y$  and  $Z$  in Frobenius distance. Using the trivial identity  $A \otimes B - \tilde{A} \otimes \tilde{B} = A \otimes (B - \tilde{B}) - (A - \tilde{A}) \otimes \tilde{B}$  twice in a row, and the triangle inequality, we can upper bound the distance  $\|X \otimes Y \otimes Z - \tilde{X} \otimes \tilde{Y} \otimes \tilde{Z}\|_F$  by

$$\|X \otimes Y \otimes (Z - \tilde{Z})\|_F + \|X \otimes (Y - \tilde{Y}) \otimes \tilde{Z}\|_F + \|(X - \tilde{X}) \otimes \tilde{Y} \otimes \tilde{Z}\|_F.$$

Since for any  $A, B$ ,  $\|A \otimes B\|_F = \|A\|_F \|B\|_F$ , the quantity above is less than  $3\varepsilon$ .  $\square$

## 2.5 Deviation bounds

In this section we collect some useful large deviation bounds.

**Fact 10** (Gaussian tail bound). Let  $g \sim \mathcal{N}(0, 1)$  be a standard normal random variable. Then for any  $t \geq 0$ ,

$$\Pr[|g| \geq t] \leq 2e^{-t^2/2}.$$

**Fact 11** (Hoeffding's inequality). Let  $h_1, \dots, h_N$  be independent centered random variables such that for every  $i \in [N]$ , we have  $\Pr[h_i \in [a_i, b_i]] = 1$ . Then for any  $t \geq 0$ ,

$$\Pr\left[\left|\sum_{i=1}^N h_i\right| \geq t\right] \leq 2e^{-2t^2 / \sum_i (b_i - a_i)^2}.$$

**Fact 12** (Bernstein's inequality, see eg. Prop. 16 in [Ver10]). Let  $h_1, \dots, h_N$  be independent centered random variables and  $K > 0$  be such that  $\Pr[|h_i| \geq t] \leq e^{1-t/K}$  for all  $i$  and  $t \geq 0$ . Then for any  $a \in \mathbb{R}^N$  and  $t \geq 0$ ,

$$\Pr\left[\left|\sum_{i=1}^N a_i h_i\right| \geq t\right] \leq 2e^{-\frac{1}{4e} \min\left(\frac{t^2}{2eK^2 \|a\|_2^2}, \frac{t}{K \|a\|_\infty}\right)}.$$

**Corollary 13** ( $\chi^2$  tail bound). Let  $|g\rangle$  be a random vector distributed according to  $\mathcal{N}(0, 1)^N$ . Then for every  $t \geq 0$ ,

$$\Pr\left[\left|\| |g\rangle \|_2^2 - N\right| \geq t\right] \leq 2e^{-\frac{1}{8e} \min\left(\frac{t^2}{4eN}, t\right)}.$$

*Proof.* Write  $|g\rangle = g_1|1\rangle + \dots + g_N|N\rangle$  where  $g_1, \dots, g_N$  are i.i.d. standard normal random variables. By Fact 10, for every  $i$  the  $g_i$  satisfy that for every  $t \geq 0$ ,

$$\begin{aligned} \Pr[|g_i^2 - 1| \geq t] &= \Pr[g_i^2 \geq t + 1] + \Pr[g_i^2 \leq 1 - t] \\ &\leq ee^{-(t+1)/2}, \end{aligned}$$

where the factor  $e$  in front ensures that the bound is trivial whenever the second term  $\Pr(g_i^2 \leq 1 - t)$  is nonzero. Hence the random variables  $h_i := g_i^2 - 1$  satisfy the hypothesis of Fact 12 with  $K = 2$ , which immediately gives the claimed bound.  $\square$

**Corollary 14** (Projections of Bernoulli vectors). *Let  $\varepsilon_{ij}, i, j \in [N]$  be i.i.d. Bernoulli random variables, and  $a \in \mathbb{R}^N$ . Then*

$$\Pr \left[ \left| \sum_{j=1}^N \left( \sum_{i=1}^N a_i \varepsilon_{ij} \right)^2 - N \|a\|_2^2 \right| > t \right] \leq 2e^{-\frac{1}{4e} \min \left( \frac{t^2}{8e \|a\|_2^4 N}, \frac{t}{2 \|a\|_2^2} \right)}.$$

*Proof.* For any  $j \in [N]$  let  $\eta_j = (\sum_{i=1}^N a_i \varepsilon_{ij})^2 - \|a\|_2^2$ . The  $\eta_j$  are independent centered random variables, and by Fact 11 they satisfy a tail bound as required by Fact 12, with  $K = 2\|a\|_2^2$ . The corollary follows.  $\square$

The following is a special case of a result due to Latała (see Corollary 1 in [Lat06]).

**Corollary 15.** *Let  $A \in \text{Herm}(N)$  be a Hermitian matrix and  $|g\rangle \sim \mathcal{N}(0, 1)^N$ . Then, for any  $t \geq 0$ ,*

$$\Pr \left[ |\langle g|A|g\rangle - \text{Tr}(A)| \geq t \right] \leq 2e^{-\frac{1}{24e} \min \left( \frac{t^2}{12e \|A\|_F^2}, \frac{t}{\|A\|_\infty} \right)}.$$

*Proof.* Since  $A$  is Hermitian, it is unitarily diagonalizable:  $A = UDU^\dagger$  where  $D = \text{diag}(\lambda_i)$ , and the  $\lambda_i$  are its real eigenvalues. Then

$$\langle g|A|g\rangle = \sum_{i=1}^N \lambda_i |\langle i|U|g\rangle|^2,$$

where the  $g_i$  are the standard normal distributed coefficients of the random vector  $|g\rangle$ . Since the rows of  $U$  are orthogonal, the  $\langle i|U|g\rangle$  are independent random variables. Moreover, since  $|g\rangle$  is real, we have  $|\langle i|U|g\rangle|^2 = (\Re(\langle i|U|g\rangle))^2 + (\Im(\langle i|U|g\rangle))^2$ , where  $\Re(\langle i|U|g\rangle)$  and  $\Im(\langle i|U|g\rangle)$  are the real and imaginary parts of the unit vector  $\langle i|U|g\rangle$  forming the  $i$ -th row of  $U$ . By rotation invariance, we have that for arbitrary  $|x\rangle \in \mathbb{R}^N$ , the random variable  $\langle x|g\rangle$  is distributed as  $\mathcal{N}(0, \|x\|^2)$ . It follows from Fact 10 that for every  $i \in [N]$ , we have

$$\begin{aligned} \Pr[|\langle i|U|g\rangle|^2 \geq t] &\leq \Pr[(\Re(\langle i|U|g\rangle))^2 \geq t/2] + \Pr[(\Im(\langle i|U|g\rangle))^2 \geq t/2] \\ &\leq 2e^{-t/(4\|\Re(\langle i|U|g\rangle)\|^2)} + 2e^{-t/(4\|\Im(\langle i|U|g\rangle)\|^2)} \\ &\leq 4e^{-t/4} \end{aligned}$$

Hence we can apply Fact 12 with  $K = 4(\ln(4/e) + 1) \leq 6$  to obtain for any  $t \geq 0$ :

$$\Pr \left[ \left| \sum_{i=1}^N \lambda_i \left( |\langle i|U|g\rangle|^2 - \mathbb{E}[|\langle i|U|g\rangle|^2] \right) \right| \geq t \right] \leq 2e^{-\frac{1}{24e} \min\left(\frac{t^2}{12e\|A\|_F^2}, \frac{t}{\|A\|_\infty}\right)}$$

which proves the claim since  $\mathbb{E}[|\langle i|U|g\rangle|^2] = 1$  for every  $i \in [N]$  and  $\sum_{i=1}^N \lambda_i = \text{Tr}(A)$ .  $\square$

Finally, we state without proof an analogue of the preceding Corollary which applies to Bernoulli random variables, and is a special case of a result of Hanson and Wright [HW71].

**Theorem 16.** *There exists a constant  $D > 0$  such that the following holds. Let  $A \in \text{Herm}(N)$  be a Hermitian matrix and  $\varepsilon_{ij}$  i.i.d. Bernoulli random variables. Then, for any  $t \geq 0$ ,*

$$\Pr \left[ \left| \sum_{i,j} A_{ij} \varepsilon_{ij} - \text{Tr}(A) \right| \geq t \right] \leq 2e^{-C \min\left(\frac{t^2}{\|A\|_F^2}, \frac{t}{\|A\|_\infty}\right)}.$$

## 2.6 Grothendieck's inequality

We use the following version of Grothendieck's inequality [Gro53]. The bounds on the constants involved come from [Haa87] and [BMMN11].

**Theorem 17** (Grothendieck's inequality). *There exists a universal constant  $K_G^{\mathbb{R}} < 1.783$  such that the following holds. Let  $N$  and  $d$  be positive integers. Then, for any matrix  $M \in \text{Mat}(N)$  with real coefficients and any complex unit vectors  $x_1, \dots, x_N, y_1, \dots, y_N \in \mathbb{S}(\mathbb{C}^d)$ , we have*

$$\left| \sum_{i,j=1}^N M_{ij} \langle x_i, y_j \rangle \right| \leq K_G^{\mathbb{R}} \max_{\chi, v: [N] \rightarrow \{-1,1\}} \sum_{i,j=1}^N M_{ij} \chi(i) v(j), \quad (6)$$

If we allow  $\chi, v$  on the right-hand side of (6) to take values in the set of all complex numbers with modulus (at most) 1, then the constant  $K_G^{\mathbb{R}}$  may be replaced by the complex Grothendieck constant  $K_G^{\mathbb{C}} < 1.405$ .

## 3 Unbounded gaps

This section is devoted to the proof of Theorem 1. The theorem is proved in two steps. In the first step we associate a three player XOR game  $G$  to any 3-tensor  $T$ , and relate the quantum-classical gap for that game to spectral properties of  $T$ . We emphasize that the game  $G = G(T)$  is *not* defined from  $T$  in the most straightforward way (using  $T$  as the game tensor), but through a more delicate transformation, based on the use of the Fourier transform, which is exposed in Section 3.1.

**Proposition 18.** *Let  $n$  be an integer and let  $N = 2^n$ . Let  $T$  be any 3-tensor of dimensions  $N^2 \times N^2 \times N^2$ . Then there exists a 3-player XOR game  $G = G(T)$  such that*

$$\frac{\beta^*(G)}{\beta(G)} \geq \frac{1}{4N^{3/2}} \frac{\|T\|_{3,3}}{\|T\|_{2,2,2}}.$$

Moreover, in the game  $G$  there are  $N^2$  questions to each player, and there is an entangled strategy which achieves the claimed violation and uses only  $N$ -dimensional Pauli observables.

In the second step we show the existence of a tensor  $T$  such that  $\|T\|_{3,3}/\|T\|_{2,2,2}$  is large.

**Proposition 19.** *There is a constant  $C > 0$  such that for any integer  $N$  there exists a 3-tensor  $T$  of dimensions  $N^2 \times N^2 \times N^2$  such that*

$$\frac{\|T\|_{3,3}}{\|T\|_{2,2,2}} \geq CN^2 \log^{-5/2} N.$$

Theorem 1 trivially follows from the two propositions above. While we have not made the constants in the preceding propositions completely explicit, it is not hard to extract numerical values from our proofs; in particular we give precise estimates for all our probabilistic arguments. Proposition 18 is proved in Section 3.1, and Proposition 19 is proved in Section 3.2.

### 3.1 Pauli XOR games

Let  $T$  be a complex 3-tensor of dimensions  $N^2 \times N^2 \times N^2$ , where  $N = 2^n$  and  $n$  is an arbitrary integer. Based on  $T$  we define a three-player XOR game  $G = G(T)$  with the following properties:

1. There are  $N^2$  questions per player,
2. The best classical strategy for game  $G(T)$  achieves a bias of at most  $N^{9/2}\|T\|_{2,2,2}$ ,
3. There is a entangled strategy which uses only Pauli matrices as observables and entanglement of local dimension  $N$  per player and achieves a bias of at least  $(N^3/4)\|T\|_{3,3}$ .

Properties 2. and 3. imply that in game  $G(T)$ , the ratio between the entangled and classical biases is at least

$$\frac{\beta^*(G)}{\beta(G)} \geq \frac{1}{4N^{3/2}} \frac{\|T\|_{3,3}}{\|T\|_{2,2,2}},$$

proving Proposition 18.

Let  $T$  be a  $N^2 \times N^2 \times N^2$  tensor. By replacing  $T$  by either  $(T + T^\dagger)/2$  or  $i(T - T^\dagger)/2$ , we may assume that  $T$ , when seen as an  $N^3 \times N^3$  matrix, is also Hermitian. One of these two possible choices necessarily results in a ratio of the  $\|\cdot\|_{3,3}$  norm to the  $\|\cdot\|_{2,2,2}$  norm that is at least half of what it was for  $T$ . In order to associate an XOR game to  $T$ , we first define (possibly complex) coefficients indexed by Pauli matrices  $P, Q, R \in \mathcal{P}_n$  as follows

$$M_{P,Q,R} := \widehat{T}(P, Q, R) = \sum_{(i,i'),(j,j'),(k,k') \in [N]^2} T_{(i,i'),(j,j'),(k,k')} P_{i,i'} Q_{j,j'} R_{k,k'}.$$

In order to obtain an XOR game  $G = G(T)$ , we take either the real or the imaginary part of the coefficients  $M_{P,Q,R}$  (whichever allows for the largest entangled bias), and normalize the resulting sequence according to its  $\ell_1$  norm (note that this normalization has no effect on the ratio of the biases that is considered in Proposition 18). This results in a game with  $N^2$  questions per player, indexed by the Pauli matrices. Since we are ultimately only concerned with the ratio  $\|T\|_{3,3}/\|T\|_{2,2,2}$ , without loss of generality we assume that the two transformations made above (making  $T$  Hermitian and such that the coefficients defined above are all real) resulted in the  $\|\cdot\|_{3,3,3}$  norm being divided by a factor at most 4, and the  $\|\cdot\|_{2,2}$  norm remaining unchanged.

The fact that property 1. above holds is clear, by definition. Next we prove that property 2. holds. Let  $\chi, v, \zeta : \mathcal{P}_n \rightarrow \{-1, 1\}$  be an optimal classical strategy. Define the matrices  $X = \sum_{P \in \mathcal{P}_n} \chi(P)P$ ,  $Y = \sum_{Q \in \mathcal{P}_n} v(Q)Q$  and  $Z = \sum_{R \in \mathcal{P}_n} \zeta(R)R$ . Then  $X, Y$  and  $Z$  are Hermitian, and

$$\|X\|_F^2 = \text{Tr}(X^\dagger X) = \sum_{P, P' \in \mathcal{P}_n} \chi(P)\chi(P')\text{Tr}(P^\dagger P') = N \sum_{P \in \mathcal{P}_n} \chi(P)^2 = N^3,$$

and the same holds for  $Y$  and  $Z$ . By the Cauchy-Schwarz inequality, the classical bias can be bounded as

$$\begin{aligned} \beta(M) &= \sum_{P, Q, R} \hat{T}(P, Q, R) \chi(P)v(Q)\zeta(R) \\ &= \sum_{P, Q, R \in \mathcal{P}_n} \langle T, X \otimes Y \otimes Z \rangle \chi(P)v(Q)\zeta(R) \\ &\leq \max_{X, Y, Z \in \mathcal{B}(\text{Herm}(N), N^{3/2})} \langle T, X \otimes Y \otimes Z \rangle \\ &\leq N^{9/2} \|T\|_{2,2,2}. \end{aligned}$$

Finally, we prove property 3. by exhibiting a good entangled strategy for  $G(T)$ . We simply let the observable corresponding to question  $P$  (resp.  $Q, R$ ) be the  $n$ -qubit Pauli matrix  $P$  (resp.  $Q, R$ ). Let  $|\Psi\rangle$  be a shared entangled state. The bias of the corresponding strategy is

$$\sum_{P, Q, R} \hat{T}(P, Q, R) \langle \Psi | P \otimes Q \otimes R | \Psi \rangle = N^3 \langle \Psi | T | \Psi \rangle = N^3 \|T\|_{3,3},$$

where for the last equality we chose  $|\Psi\rangle$  an eigenvector of  $T$  with largest eigenvalue.

**Remark.** *In our construction, the only properties of the Pauli matrices that we use is that they form a family of observables that is orthogonal with respect to the Hilbert-Schmidt inner product on  $\text{Herm}(N)$ . Any other such family would lead to a completely analogous construction (in which the player's observables in the entangled strategy are replaced by the corresponding elements).*

### 3.2 Constructing a good tensor $T$

In this section we prove Proposition 19 by giving a probabilistic argument for the existence of a tensor  $T$  with good spectral properties. Let  $N$  be an integer, and  $|g\rangle$  the (random)  $N^3$ -dimensional vector

$$|g\rangle := \sum_{i, j, k=1}^N g_{ijk} |i\rangle |j\rangle |k\rangle \sim \mathcal{N}(0, 1)^{N^3},$$

where the  $g_{ijk}$  are i.i.d.  $\mathcal{N}(0, 1)$  random variables. We define a tensor  $T$  depending on the  $g_{ijk}$ , and then prove bounds on the  $\|\cdot\|_{3,3}$  and  $\|\cdot\|_{2,2,2}$  norms of  $T$  that hold with high probability over the choice of the  $g_{ijk}$ . Let

$$T := \sum_{i \neq i', j \neq j', k \neq k'} g_{ijk} g_{i'j'k'} |i, j, k\rangle \langle i', j', k'|. \quad (7)$$

$T$  is a real  $N^3 \times N^3$  symmetric matrix that equals  $|g\rangle\langle g|$  with some coefficients zeroed out, including those on the diagonal. Hence  $T$  is very close to a rank 1 matrix and it should therefore be no surprise that its spectral norm is large, as we show in Section 3.2.1 below. More work is needed to upper bound the  $\|\cdot\|_{2,2,2}$  norm of  $T$ . In particular, we note that zeroing out the diagonal coefficients is essential to getting a good bound on  $\|T\|_{2,2,2}$ . While we show in Section 3.2.2 that with high probability over  $|g\rangle$  we have  $\|T\|_{2,2,2} = O(N \log^{5/2} N)$ , it is not hard to see that in expectation we already have  $\||g\rangle\langle g|\|_{2,2,2} = \Omega(N\sqrt{N})$  (indeed, simply choose  $X = Y = Z = I/\sqrt{N}$  in the definition of  $\|\cdot\|_{2,2,2}$ ). Zeroing out some entries of  $|g\rangle\langle g|$  approximately preserves the spectral norm, but decreases its norm as a trilinear operator by almost a factor  $\sqrt{N}$ .

**Remark.** *The same construction, with the normal random variables  $g_{ijk}$  replaced by i.i.d. Bernoulli random variables, can be used to obtain similar results.<sup>9</sup> Indeed, Lemma 20 below holds trivially in that case, and to obtain the analogue of Lemma 21 it suffices to replace the use of Corollary 13 and Corollary 15 in the proof of Lemma 24 by Corollary 14 and Theorem 16 respectively.*

### 3.2.1 A lower bound on the spectral norm

A lower-bound on the spectral norm of  $T$  as defined in (7) follows easily from the fact that it is, by definition, very close to a rank-1 matrix. We show the following.

**Lemma 20.** *For any  $\tau > 0$  and all large enough  $N$  it holds that*

$$\|T\|_{3,3} \geq N^3 - \tau N^2$$

*with probability at least  $1 - e^{-\Omega(\tau^2)}$ .*

*Proof.* Define  $|\Psi\rangle = N^{-3/2}|g\rangle$ . By Corollary 13, for any  $\delta > 0$  we have that

$$\Pr\left[\sum_{i,j,k} g_{ijk}^2 \leq (1 - \delta)N^3\right] \leq 2e^{-\delta^2 N^3 / (64e^2)}.$$

Provided this holds,

$$\||\Psi\rangle\|^2 = \frac{1}{N^3} \sum_{i,j,k} g_{ijk}^2 \leq 1 - \delta. \quad (8)$$

Another application of Corollary 13, together with a union bound, shows that the probability that there exists an  $i \in [N]$  such that  $\sum_{j,k} g_{ijk}^2 \geq (1 + \delta)N^2$  is at most  $2Ne^{-\delta^2 N^2 / (64e^2)}$ . Provided this holds,

$$\sum_i \left(\sum_{j,k} |g_{ijk}|^2\right)^2 \leq (1 + \delta)^2 N^5 \quad (9)$$

---

<sup>9</sup>We thank Ignacio Villanueva for asking this question.



and the same holds symmetrically for  $j$  or  $k$ . This lets us bound

$$\begin{aligned} \langle \Psi | T | \Psi \rangle &= \frac{1}{N^3} \sum_{i \neq i', j \neq j', k \neq k'} |g_{ijk}|^2 |g_{i'j'k'}|^2 \\ &\geq \frac{1}{N^3} \left( \left( \sum_{i,j,k} |g_{ijk}|^2 \right)^2 - \sum_i \left( \sum_{j,k} |g_{ijk}|^2 \right)^2 - \sum_j \left( \sum_{i,k} |g_{ijk}|^2 \right)^2 - \sum_k \left( \sum_{i,j} |g_{ijk}|^2 \right)^2 \right) \\ &\geq \frac{(1-\delta)^2 N^6 - 3(1+\delta)^2 N^5}{N^3} \geq (1-3\delta) N^3, \end{aligned}$$

where the second inequality uses (8) and (9), and the last holds for large enough  $N$ . Hence, using (8) once more,

$$\|T\|_{3,3} \geq \frac{\langle \Psi | T | \Psi \rangle}{\|\Psi\|^2} \geq (1-3\delta) N^3 (1-\delta)^{-1} \geq (1-6\delta) N^3$$

for small enough  $\delta$ . The claimed bound follows by setting  $\delta = \tau/(6N)$ .  $\square$

### 3.2.2 Upper-bounding $\|T\|_{2,2,2}$

In this section we give an upper bound for  $\|T\|_{2,2,2}$  that holds with good probability over the choice of  $T$ , where  $T$  is as in (7), a 3-tensor of dimensions  $N^2 \times N^2 \times N^2$ . Recall that

$$\|T\|_{2,2,2} = \max_{X,Y,Z \in \mathcal{B}(\text{Herm}(N))} |\langle T, X \otimes Y \otimes Z \rangle|.$$

We prove the following.

**Lemma 21.** *There exist universal constants  $d, D > 0$  such that for all large enough  $N$ , we have*

$$\|T\|_{2,2,2} \leq DN(\ln N)^{5/2}$$

with probability at least  $1 - e^{-dN}$  over the choice of  $|g\rangle$ .

We note that if  $T$  was a random tensor with entries i.i.d. standard normal, then a result by Nguyen et al. [NDT10] would show that  $\|T\|_{2,2,2} = O(N\sqrt{\log N})$  holds with high probability. However, the entries of our tensor  $T$  are not independent, and we need to prove a bound tailored to our specific setting.

Our first step consists in showing that the supremum in the definition of  $\|T\|_{2,2,2}$  can be restricted to a supremum over projector matrices, at the cost of the loss of a logarithmic factor in the bound.<sup>10</sup>

**Lemma 22.** *Let  $|g\rangle$  be a vector in  $\mathbb{R}^{N^3}$  and let  $T$  be the associated tensor, as in (7). Then*

$$\|T\|_{2,2,2} \leq 64(\ln N)^{3/2} \max |\langle g | X \otimes Y \otimes Z | g \rangle - \text{Tr}(X \otimes Y \otimes Z)|, \quad (10)$$

where the maximum is taken over all triples  $(X, Y, Z) \in \overline{\text{Proj}}(N)^3$ .

<sup>10</sup>We thank Gilles Pisier for suggesting the use of this decomposition.

*Proof.* Let  $X, Y, Z \in B(\text{Herm}(N))$  be traceless Hermitian matrices such that

$$\|T\|_{2,2,2} = \langle T, X \otimes Y \otimes Z \rangle = \langle g | X \otimes Y \otimes Z | g \rangle,$$

where the second equality follows from the definition of  $T$ . Decompose  $X, Y, Z$  as per Lemma 6, giving

$$X = \sum_x \alpha_x X_x, \quad Y = \sum_y \beta_y Y_y \quad \text{and} \quad Z = \sum_z \gamma_z Z_z,$$

where  $\|(\alpha_x)_x\|_1, \|(\beta_y)_y\|_1, \|(\gamma_z)_z\|_1 \leq 4\sqrt{\ln N}$  and  $X_x, Y_y, Z_z \in \overline{\text{Proj}}(N)$ . Note that

$$0 = \text{Tr}(X \otimes Y \otimes Z) = \sum_{x,y,z} \alpha_x \beta_y \gamma_z \text{Tr}(X_x \otimes Y_y \otimes Z_z).$$

By linearity and Hölder's inequality, we have

$$\begin{aligned} & \langle g | X \otimes Y \otimes Z | g \rangle - \text{Tr}(X \otimes Y \otimes Z) \\ &= \sum_{x,y,z} \alpha_x \beta_y \gamma_z \left( \langle g | X_x \otimes Y_y \otimes Z_z | g \rangle - \text{Tr}(X_x \otimes Y_y \otimes Z_z) \right) \\ &\leq 64(\ln N)^{3/2} \max_{x,y,z} \left| \langle g | X_x \otimes Y_y \otimes Z_z | g \rangle - \text{Tr}(X_x \otimes Y_y \otimes Z_z) \right|, \end{aligned}$$

proving the lemma. □

Our next step is to show that we may further restrict the maximum on the right-hand side of (10) to a maximum over projectors taken from the  $\varepsilon$ -net  $\mathcal{Z}_\varepsilon$  given in Definition 8.

**Lemma 23.** *Let  $|g\rangle$  be a vector in  $\mathbb{R}^{N^3}$ ,  $T$  the associated tensor and  $\varepsilon > 0$ . Then*

$$\|T\|_{2,2,2} \leq 64 (\ln N)^{3/2} \left( \max |\langle g | X \otimes Y \otimes Z | g \rangle - \text{Tr}(X \otimes Y \otimes Z)| + 3\varepsilon (N^{3/2} + \| |g\rangle \|_2^2) \right), \quad (11)$$

where the maximum is taken over all  $X \otimes Y \otimes Z \in \mathcal{Z}_\varepsilon$ .

*Proof.* Fix a triple  $(X, Y, Z) \in \overline{\text{Proj}}(N)^3$ . By Proposition 9, there exists an  $\tilde{X} \otimes \tilde{Y} \otimes \tilde{Z} \in \mathcal{Z}_\varepsilon$  such that

$$\|X \otimes Y \otimes Z - \tilde{X} \otimes \tilde{Y} \otimes \tilde{Z}\|_F \leq 3\varepsilon.$$

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned} |\langle g | X \otimes Y \otimes Z | g \rangle - \langle g | \tilde{X} \otimes \tilde{Y} \otimes \tilde{Z} | g \rangle| &= |\langle g | X \otimes Y \otimes Z - \tilde{X} \otimes \tilde{Y} \otimes \tilde{Z} | g \rangle| \\ &\leq \|X \otimes Y \otimes Z - \tilde{X} \otimes \tilde{Y} \otimes \tilde{Z}\|_F \| |g\rangle \|_F. \end{aligned}$$

Another application of the Cauchy-Schwarz inequality and the definition of the Frobenius norm give

$$\begin{aligned} |\text{Tr}(X \otimes Y \otimes Z - \tilde{X} \otimes \tilde{Y} \otimes \tilde{Z})| &= |\langle I, X \otimes Y \otimes Z - \tilde{X} \otimes \tilde{Y} \otimes \tilde{Z} \rangle| \\ &\leq N^{3/2} \|X \otimes Y \otimes Z - \tilde{X} \otimes \tilde{Y} \otimes \tilde{Z}\|_F. \end{aligned}$$

Hence the lemma follows from Lemma 22. □

We upper-bound the right-hand side of (11) by first showing that for any fixed triple  $(k, \ell, m) \in [N]^3$  and  $X \otimes Y \otimes Z \in \mathcal{Z}_\varepsilon^{(k, \ell, m)}$ , this quantity is bounded with high probability over the choice of  $|g\rangle$ . We conclude by applying a union bound over the net  $\mathcal{Z}_\varepsilon = \bigcup_{(k, \ell, m) \in [N]^3} \mathcal{Z}_\varepsilon^{(k, \ell, m)}$ .

**Lemma 24.** *There exist constants  $C, c > 0$  such the following holds. For any  $0 < \varepsilon \leq N^{-3}$  and  $\tau \geq CN \ln(1/\varepsilon)$ , the probability over the choice of  $|g\rangle$  that there exists an  $X \otimes Y \otimes Z \in \mathcal{Z}_\varepsilon$  such that*

$$|\langle g | X \otimes Y \otimes Z | g \rangle - \text{Tr}(X \otimes Y \otimes Z)| > \tau \quad (12)$$

is at most  $e^{-c\tau}$ .

*Proof.* Fix a triple  $(k, \ell, m) \in [N]^3$ , and assume that  $k \geq \max\{\ell, m\}$ , the other cases being reduced to this one by permutation of the indices. Since  $k + \ell + m \leq 3k$ , we have

$$|\mathcal{Z}_\varepsilon^{(k, \ell, m)}| \leq 8 \left(\frac{5}{\varepsilon}\right)^{(k+\ell+m)N} \leq e^{3kN \ln(5/\varepsilon) + 3}. \quad (13)$$

We distinguish two cases.

Case 1:  $\ell m > k$ . Fix an  $X \otimes Y \otimes Z \in \mathcal{Z}_\varepsilon^{(k, \ell, m)}$ . By definition of the nets  $\mathcal{Z}_\varepsilon^j$ ,

$$\|X \otimes Y \otimes Z\|_F \leq 1 \quad \text{and} \quad \|X \otimes Y \otimes Z\|_\infty = \frac{1}{\sqrt{k\ell m}}.$$

Hence, by Corollary 15 there exists a constant  $c' > 0$  such that for any  $\tau > 0$

$$\Pr_{|g\rangle} \left[ |\langle g | X \otimes Y \otimes Z | g \rangle - \text{Tr}(X \otimes Y \otimes Z)| \geq \tau \right] \leq e^{-c' \min\{\tau^2, \tau \sqrt{k\ell m}\}}. \quad (14)$$

Our assumption  $\ell m > k$  implies  $\sqrt{k\ell m} > k$ , hence the probability above is at most  $e^{-c' \min\{\tau^2, k\tau\}}$ . Using the bound (13) on the size of  $\mathcal{Z}_\varepsilon^{(k, \ell, m)}$ , by a union bound there exists a  $C' > 0$  such that for any  $\tau \geq C' N \ln(1/\varepsilon)$  the probability that there exists an  $X' \otimes Y' \otimes Z' \in \mathcal{Z}_\varepsilon^{(k, \ell, m)}$  such that

$$|\langle g | X' \otimes Y' \otimes Z' | g \rangle - \text{Tr}(X' \otimes Y' \otimes Z')| \geq \tau$$

is at most  $e^{-\Omega(\tau)}$ .

Case 2:  $k \geq \ell m$ . Fix an  $X \otimes Y \otimes Z \in \mathcal{Z}_\varepsilon^{(k, \ell, m)}$ . Since  $X, Y$  and  $Z$  are normalized projectors,

$$\text{Tr}(X \otimes Y \otimes Z) \leq \sqrt{k\ell m} \leq k \leq N.$$

Write the spectral decompositions of  $X, Y$  and  $Z$  as

$$X = \frac{1}{\sqrt{k}} \sum_p |x_p\rangle \langle x_p|, \quad Y = \frac{1}{\sqrt{\ell}} \sum_q |y_q\rangle \langle y_q| \quad \text{and} \quad Z = \frac{1}{\sqrt{m}} \sum_r |z_r\rangle \langle z_r|,$$

where the indices  $p, q, r$  run from 1 to at most  $k, \ell, m$ , respectively. For any unit  $|y\rangle, |z\rangle \in \mathbb{C}^N$ , define the  $N$ -dimensional vector  $|g(y, z)\rangle = (I \otimes |y\rangle \langle y| \otimes |z\rangle \langle z|)|g\rangle$ . By rotation invariance of the Gaussian distribution,  $|g(y, z)\rangle$  is distributed according to  $\mathcal{N}(0, 1)^N$ . Since  $|x_1\rangle, |x_2\rangle, \dots$  are pairwise

orthogonal, we have

$$\begin{aligned}
|\langle g|X \otimes Y \otimes Z|g\rangle| &= \frac{1}{\sqrt{k\ell m}} \sum_{p,q,r} |\langle x_p|g(y_q, z_r)\rangle|^2 \\
&\leq \sqrt{\frac{\ell m}{k}} \max_{|y\rangle, |z\rangle \in \mathbb{C}^N, \| |y\rangle \|, \| |z\rangle \| \leq 1} \| |g(y, z)\rangle \|_2^2 \\
&\leq \max_{|y\rangle, |z\rangle \in S_\varepsilon} \| |g(y, z)\rangle \|_2^2 + 4\varepsilon \| |g\rangle \|_2^2, \tag{15}
\end{aligned}$$

where for the last inequality we used that  $\sqrt{\ell m/k} \leq 1$  (which follows from our assumption  $k \geq \ell m$ ), and that for any unit  $y, \tilde{y}, z, \tilde{z}$ ,

$$\begin{aligned}
\left| \| |g(y, z)\rangle \|_2^2 - \| |g(\tilde{y}, \tilde{z})\rangle \|_2^2 \right| &= \left| (\langle g(y, z)| + \langle g(\tilde{y}, z)|) (|g(y, z)\rangle - |g(\tilde{y}, z)\rangle) \right. \\
&\quad \left. + (\langle g(\tilde{y}, z)| + \langle g(\tilde{y}, \tilde{z})|) (|g(\tilde{y}, z)\rangle - |g(\tilde{y}, \tilde{z})\rangle) \right| \\
&\leq \| |g(y + \tilde{y}, z)\rangle \|_2 \| |g(y - \tilde{y}, z)\rangle \|_2 + \| |g(\tilde{y}, z + \tilde{z})\rangle \|_2 \| |g(\tilde{y}, z - \tilde{z})\rangle \|_2 \\
&\leq 2 \| |g\rangle \|_2^2 (\| |y - \tilde{y}\rangle \|_2 + \| |z - \tilde{z}\rangle \|_2).
\end{aligned}$$

Applying Corollary 13, there exists a  $c'' > 0$  such that for any  $\tau > 0$  the maximum in (15) is greater than  $N + \tau$  with probability at most  $e^{-c'' \min\{\tau^2/N, \tau\}}$ . Since by Fact 5  $|S_\varepsilon| \leq e^{-2 \ln(1/\varepsilon)N}$ , a union bound shows that there exists a  $C'' > 0$  such that for all  $\tau \geq C''N \ln(1/\varepsilon)$  the bound

$$|\langle g|X \otimes Y \otimes Z|g\rangle| \leq \varepsilon(N^3 + N\tau) + \tau$$

holds with probability at least  $e^{-c'''\tau}$  over the choice of  $|g\rangle$ , for some  $c''' > 0$ . (Here we again used Corollary 13 to upper-bound  $\| |g\rangle \|_2^2 \leq N^3 + N\tau$  with probability at least  $1 - e^{-\Omega(\tau)}$ .)

The lemma follows for some  $c, C > 0$  by combining the two cases analyzed above and performing a union bound over all  $N^3$  triples  $(k, \ell, m)$ .  $\square$

We are now in a position to prove Lemma 21.

*of Lemma 21.* Let  $\varepsilon = N^{-3}$  and  $\tau = CN \ln(1/\varepsilon)$ , where  $C$  is the constant appearing in the statement of Lemma 24. That lemma shows that the bound

$$|\langle g|X \otimes Y \otimes Z|g\rangle - \text{Tr}(X \otimes Y \otimes Z)| \leq CN \ln(1/\varepsilon)$$

holds except with probability at most  $e^{-cCN \ln(1/\varepsilon)}$ . Moreover, by Corollary 13, there is a  $C' > 0$  such that

$$3\varepsilon \left( N^{3/2} + \| |g\rangle \|_2^2 \right) \leq 6\varepsilon N^{3/2},$$

except with probability at most  $1 - e^{-C'N}$ . Combining these two bounds with the estimate of Lemma 23 proves the lemma, provided  $d$  is chosen small enough and  $D$  large enough.  $\square$

## 4 Upper bounds on violations

### 4.1 Bounds in terms of the number of questions

In this section we prove Theorem 2, which we restate here for convenience.

**Theorem 2.** *For any 3-player XOR game  $G$  in which there are at most  $Q$  possible questions to the third player,*

$$\beta^*(G) \leq \sqrt{Q} K_G^{\mathbb{R}} \beta(G),$$

where  $K_G^{\mathbb{R}} < 1.783$  is the real Grothendieck constant.

The two main ingredients in the proof are a useful technique of Paulsen and Grothendieck's inequality. Paulsen's technique (see [Pau92, Proposition 2.10]) lets us "decouple" the third player from the other two players and turn his part of the entangled strategy into a classical one at a loss of a factor  $\sqrt{Q}$  in the overall bias.<sup>11</sup> Slightly more precisely, the proof goes as follows. By grouping the game tensor and the observables of the first two players together, the entangled bias takes the form

$$\beta^*(G) = \langle \psi | \sum_{k=1}^Q M_k \otimes C_k | \psi \rangle,$$

where the  $C_k$  are the third player's observables in an optimal entangled strategy. The decoupling technique relies on a collection of i.i.d.  $\{-1, 1\}$ -valued symmetrically distributed Bernoulli random variables  $\varepsilon_1, \dots, \varepsilon_Q$  which are used to split the above sum into two sums. Using the fact that  $\mathbb{E}[\varepsilon_k \varepsilon_\ell] = \delta_{k\ell}$ , the above expression can be written as

$$\mathbb{E} \left[ \left( \langle \psi | \sum_{k=1}^Q M_k \otimes (\varepsilon_k I) \right) \left( \sum_{\ell=1}^Q \varepsilon_\ell I \otimes C_\ell | \psi \rangle \right) \right].$$

After two applications of the Cauchy-Schwarz inequality, the third player's classical strategy will be a certain instantiation of the random variables  $\varepsilon_k$  appearing in the left brackets, while the factor  $\sqrt{Q}$  will come from the term between the right brackets. An application of Grothendieck's inequality will let us turn the first two players' entangled strategy into a classical one at a loss of an extra constant factor in the overall bias. We proceed with the formal proof of the theorem.

*of Theorem 2.* Suppose that the game  $G$  is defined by the probability distribution  $\pi$  and sign tensor  $M$ . Define the tensor  $T_{ijk} = \pi(ijk)M(ijk)$ . By setting some entries to zero we may assume without loss of generality that  $T$  has dimension  $Q \times Q \times Q$ . Fix an arbitrary constant  $\epsilon > 0$  and let  $|\psi\rangle$ ,  $A_i, B_j, C_k$  be a finite-dimensional state and  $\{-1, 1\}$ -valued observables such that<sup>12</sup>

$$\beta^*(G) \leq (1 + \epsilon) \sum_{i,j,k=1}^Q T_{ijk} \langle \psi | A_i \otimes B_j \otimes C_k | \psi \rangle.$$

<sup>11</sup>This technique is based on so-called Rademacher averaging, a well-know method in the field of Banach spaces.

<sup>12</sup>A standard approximation argument based on the Spectral Theorem shows that this is always possible.

Define for every  $k \in [Q]$  the matrix  $M_k = \sum_{i,j=1}^Q T_{ijk} A_i \otimes B_j$ . Let  $\varepsilon_1, \dots, \varepsilon_Q$  be i.i.d.  $\{-1, 1\}$ -valued symmetrically distributed Bernoulli random variables. Using the fact that  $\mathbb{E}[\varepsilon_k \varepsilon_\ell] = \delta_{k\ell}$  and the Cauchy-Schwarz inequality, the right-hand side of the above inequality can be written as and bounded by

$$\mathbb{E} \left[ \left( \langle \psi | \sum_{k=1}^Q M_k \otimes (\varepsilon_k I) \right) \left( \sum_{\ell=1}^Q \varepsilon_\ell I \otimes C_\ell | \psi \rangle \right) \right] \leq \mathbb{E} \left[ \left\| \langle \psi | \sum_{k=1}^Q M_k \otimes (\varepsilon_k I) \right\|_2 \left\| \sum_{\ell=1}^Q \varepsilon_\ell I \otimes C_\ell | \psi \rangle \right\|_2 \right].$$

Another application of Cauchy-Schwarz gives that the right-hand side is bounded from above by

$$\left( \mathbb{E} \left[ \left\| \langle \psi | \sum_{k=1}^Q M_k \otimes (\varepsilon_k I) \right\|_2^2 \right] \right)^{1/2} \left( \mathbb{E} \left[ \left\| \sum_{\ell=1}^Q \varepsilon_\ell I \otimes C_\ell | \psi \rangle \right\|_2^2 \right] \right)^{1/2}. \quad (16)$$

The fact that the matrices  $\varepsilon_\ell I \otimes C_\ell$  are unitary and  $|\psi\rangle$  is a unit vector shows that the above term on the right equals  $\sqrt{Q}$ . Since the matrices  $M_k \otimes (\varepsilon_k I)$  are Hermitian, the left term in (16) is at most

$$\max_{|\phi\rangle, \zeta: [Q] \rightarrow \{-1, 1\}} \langle \phi | \sum_{k=1}^Q M_k \otimes (\zeta(k) I) | \phi \rangle.$$

Expanding the definition of  $M_k$ , we have shown that

$$\beta^*(G) \leq (1 + \varepsilon) \sqrt{Q} \max_{|\phi\rangle, \zeta: [Q] \rightarrow \{-1, 1\}} \langle \phi | \sum_{i,j,k=1}^Q T_{ijk} A_i \otimes B_j \otimes (\zeta(k) I) | \phi \rangle. \quad (17)$$

The matrices  $\zeta(k) I$  may be interpreted as observables corresponding to single-outcome projective measurements. The outcome of such a measurement does not depend on the particular entangled state shared with the other players nor on their measurement outcomes. The entangled bias of the game  $G$  is thus at most  $(1 + \varepsilon) \sqrt{Q}$  times the bias achievable with strategies in which the third player uses a classical strategy. The maximum on the right-hand side of (17) thus equals<sup>13</sup>

$$\max_{|\phi'\rangle, \zeta: [Q] \rightarrow \{-1, 1\}} \langle \phi' | \sum_{i,j,k=1}^Q T_{ijk} A_i \otimes B_j \zeta(k) | \phi' \rangle.$$

Let  $|\phi'\rangle$  and  $\zeta : [Q] \rightarrow \{-1, 1\}$  be such that the maximum above is achieved. Define the  $Q$ -by- $Q$  matrix  $H_{ij} = \sum_{k=1}^Q T_{ijk} \zeta(k)$ . Rearranging terms gives that the above maximum equals  $\sum_{i,j=1}^Q H_{ij} \langle \phi' | A_i \otimes B_j | \phi' \rangle$ . Define the unit vectors  $x_i = A_i \otimes I | \phi' \rangle$  and  $y_j = I \otimes B_j | \phi' \rangle$ . Clearly we have  $\langle \phi' | A_i \otimes B_j | \phi' \rangle = \langle x_i, y_j \rangle$ . The result now follows by applying Grothendieck's inequality (6) and expanding the definition of  $H_{ij}$ .  $\square$

<sup>13</sup>Another way to see this is by writing  $\sum_{i,j,k=1}^Q T_{ijk} A_i \otimes B_j \otimes (\zeta(k) I) = (\sum_{i,j,k=1}^Q T_{ijk} A_i \otimes B_j \zeta(k)) \otimes I$  and using the facts that the operator norm is multiplicative under tensor products and the identity matrix has operator norm 1.

## 4.2 Bounds in terms of the Hilbert space dimension

In this section we give a proof of Theorem 3, which we restate for convenience.

**Theorem 3.** *Let  $G$  be a 3-player XOR game in which the maximal entangled bias  $\beta^*(G)$  is achieved by a strategy in which the third player's local dimension is  $d$ . Then*

$$\beta^*(G) \leq \sqrt{3d} (K_G^{\mathbb{C}})^{3/2} \beta(G),$$

where  $K_G^{\mathbb{C}} < 1.405$  is the complex Grothendieck constant.

As the bound in terms of the number of questions presented in the previous section, the proof of Theorem 3 relies on a decoupling technique, by which the third player is reduced to using a classical strategy, while only reducing the bias that the players achieve in the game by a factor depending on the local dimension of his share of the entangled state. We use the following version of the non-commutative Khinchine's inequality, proved with optimal constants in [HM07].

**Theorem 25** (Khinchine's inequality, Proposition 2.12 in [HM07]). *Let  $A_i$  be complex  $d \times d$  matrices, and  $\varepsilon_i$  i.i.d.  $\{-1, 1\}$  symmetrically distributed. Then there exists a matrix random variable  $\tilde{A}$  such that  $\mathbb{E}[\varepsilon_i \tilde{A}] = 0$  for every  $i$ , and for every possible joint value taken by the tuple of random variables  $(\varepsilon_1, \dots, \varepsilon_d, \tilde{A})$  it holds that*

$$\left\| \sum_i \varepsilon_i A_i + \tilde{A} \right\|_{\infty} \leq \sqrt{3} \max \left\{ \left\| \sum_i A_i A_i^{\dagger} \right\|_{\infty}^{1/2}, \left\| \sum_i A_i^{\dagger} A_i \right\|_{\infty}^{1/2} \right\}. \quad (18)$$

of Theorem 3. Suppose that the game  $G$  is defined by the probability distribution  $\pi$  and sign tensor  $M$ . Define the tensor  $T_{ijk} = \pi(ijk)M(ijk)$ . Fix an arbitrary constant  $\epsilon > 0$  and let  $|\psi\rangle$ ,  $A_i, B_j, C_k$  be a finite-dimensional state and  $\{-1, 1\}$ -valued observables, where  $C_k$  has dimension  $d \times d$  and  $A_i, B_j$  have (finite) dimension  $D \times D$ , such that

$$\beta^*(G) \leq (1 + \epsilon) \sum_{i,j,k=1}^Q T_{ijk} \langle \psi | A_i \otimes B_j \otimes C_k | \psi \rangle.$$

For each  $k$ , let  $M_k = \sum_{i,j} T_{ijk} A_i \otimes B_j$ . Let  $|\Psi\rangle = \sum_i \lambda_i |u_i\rangle |v_i\rangle$  be the Schmidt decomposition, where  $|u_i\rangle$  is a vector on the system held by the first two players, and  $|v_i\rangle$  is on the third player's. Assume without loss of generality that the  $|v_i\rangle$  span the local space of the third player. Letting  $M = \sum_k M_k \otimes C_k$ , the bias achieved by this strategy is  $\langle \Psi | M | \Psi \rangle \geq (1 + \epsilon)^{-1} \beta^*(G)$ . Decompose  $M$  as  $M = \sum_{i,j} E_{i,j} \otimes |v_i\rangle \langle v_j|$ , where for every  $(i, j) \in [d]^2$   $E_{i,j}$  is a  $D^2 \times D^2$  matrix on Alice and Bob's systems; by definition

$$E_{i,j} = \sum_k \langle v_i | C_k | v_j \rangle M_k.$$

Since  $M$  is Hermitian, we have  $E_{i,j} = (E_{j,i})^{\dagger}$ . We will need the following bound.

**Claim 26.** For every  $i \in [d]$ ,

$$\max \left\{ \left\| \sum_j E_{i,j} E_{i,j}^\dagger \right\|_\infty, \left\| \sum_j E_{i,j}^\dagger E_{i,j} \right\|_\infty \right\} \leq (K_G^C)^3 \beta(G)^2. \quad (19)$$

*Proof.* Let  $|\Phi\rangle$  be any vector. Then

$$\begin{aligned} \langle \Phi | \sum_j E_{i,j} (E_{i,j})^\dagger | \Phi \rangle &= \sum_{k,k'} \sum_j \langle \Phi | M_k M_{k'} | \Phi \rangle \langle v_i | C_k | v_j \rangle \overline{\langle v_i | C_{k'} | v_j \rangle} \\ &= \sum_{k,k'} \langle \Phi | M_k M_{k'} | \Phi \rangle \langle C_{k'}^i, C_k^i \rangle \\ &\leq K_G^C \sup_{c_k \in \{\pm 1\}} \sum_{k,k'} \langle \Phi | M_k M_{k'} | \Phi \rangle c_k c_{k'}, \end{aligned} \quad (20)$$

where in the second equality we let  $C_k^i$  be the  $i$ -th row of  $C_k$  (in the  $|v_j\rangle$  basis), which has norm 1 (since  $C_k$  as a matrix is an observable), and the last inequality is Grothendieck's inequality. But

$$\begin{aligned} \sum_{k,k'} \langle \Phi | M_k M_{k'} | \Phi \rangle c_k c_{k'} &\leq \left\| \sum_{k,k'} M_k M_{k'} c_k c_{k'} \right\|_\infty \\ &= \left\| \sum_k M_k c_k \right\|_\infty^2 \leq (K_G^C)^2 \beta(G), \end{aligned}$$

again by Grothendieck's inequality: here the first two players are entangled players in  $G$ , but the third is classical. Using  $E_{i,j} = E_{j,i}^\dagger$ , this proves (19).  $\square$

Let  $\varepsilon_j$  be i.i.d.  $\{\pm 1\}$ -valued standard Bernoulli random variables, and for every  $i \in [d]$  let  $\tilde{E}_i$  the matrix random variable promised by Theorem 25, and  $E_i := \sum_j \varepsilon_j E_{i,j} + \tilde{E}_i$ . Combining the estimate in Claim 26 with the bound (18) from Theorem 25, we get that

$$\max_i \|E_i\|_\infty \leq \sqrt{3} (K_G^C)^{3/2} \beta(G). \quad (21)$$

Let  $\varepsilon'_i$  be i.i.d.  $\{\pm 1\}$ -valued standard Bernoulli random variables independent from the  $\varepsilon_i$  such that for all  $(i, j)$ ,  $E[\varepsilon'_i \tilde{E}_j] = 0$ . Starting from the  $(E_i)$ , let  $\tilde{E}$  be the matrix random variable promised by Theorem 25, and let  $E := \sum_i \varepsilon'_i E_i + \tilde{E}$ . Using the triangle inequality, the bound (18) together with (21) leads to

$$\|E\|_\infty \leq \sqrt{3d} (K_G^C)^{3/2} \beta(G), \quad (22)$$

which is valid for all choices of  $\varepsilon_j$  and  $\varepsilon'_i$ . We may now write

$$\begin{aligned} \langle \Psi | M | \Psi \rangle &= \sum_{i,j} \lambda_i \lambda_j \langle u_i | E_{i,j} | u_j \rangle \\ &= E_{\varepsilon, \varepsilon'} \left[ \text{Tr} \left( E \cdot \left( \sum_{i,j} \varepsilon'_i \varepsilon_j \lambda_i \lambda_j |u_j\rangle \langle u_i| \right) \right) \right] \\ &\leq E_{\varepsilon, \varepsilon'} \left[ \|E\|_\infty \left\| \sum_{i,j} \varepsilon'_i \varepsilon_j \lambda_i \lambda_j |u_j\rangle \langle u_i| \right\|_1 \right], \end{aligned} \quad (23)$$



where for the second equality we used that  $E[\varepsilon'_i \tilde{E}] = 0$  for every  $i$ , and the last follows from Hölder's inequality. The norm  $\|E\|_\infty$  is bounded by (22), and to conclude it suffices to note that, since

$$\sum_{i,j} \varepsilon'_i \varepsilon_j \lambda_i \lambda_j |u_j\rangle \langle u_i| = \left( \sum_j \varepsilon_j \lambda_j |u_j\rangle \right) \left( \sum_i \varepsilon'_i \lambda_i \langle u_i| \right),$$

its trace norm is at most

$$\left\| \sum_j \varepsilon_j \lambda_j |u_j\rangle \right\| \left\| \sum_i \varepsilon'_i \lambda_i \langle u_i| \right\| \leq \left( \sum_j \lambda_j^2 \right)^{1/2} \left( \sum_i \lambda_i^2 \right)^{1/2} \leq 1.$$

□

## 5 Conclusion and open problems

We have described a probabilistic construction of a family of XOR games  $G = (G_N)$  in which players sharing entanglement may gain a large, unbounded advantage over the best classical, unentangled players. For any  $N = 2^n$  the game  $G_N$  has  $N^2$  questions per player, and is such that the ratio  $\beta^*(G)/\beta(G) = \Omega(\sqrt{N} \log^{-5/2} N)$ . Our results raise two immediate open questions. The first is whether this estimate is optimal: we could only prove an upper bound of  $O(N)$  on the largest possible ratio (for games, such as  $G_N$ , with at most  $N^2$  questions per player). The second is to give an explicit, deterministic construction of a family of games achieving a similar (or even weaker) ratio. Such a construction would be of great interest both to experimental physicists and to operator space theorists, no small feat!

In our results we measured the advantage of entangled players in a given XOR game  $G$  multiplicatively, as a function of the ratio  $\beta^*(G)/\beta(G)$ . Although this has become customary, if one is interested in experimental realizations it may not be the most appropriate way to measure the advantage gained by entanglement, as small biases may be hard to notice, however large the ratio between the entangled and unentangled biases. In the case of our specific construction, one may compute that  $\beta^*(G_N) = \Omega(N^{-3/2})$  and  $\beta(G_N) = O(N^{-2} \log^{5/2} N)$ : while the ratio of these two quantities is large, both are relatively close to 0 and may thus be difficult to differentiate through experiment. It is an interesting open problem to also obtain large separations as measured, say, by the difference  $\beta^*(G) - \beta(G)$ .

**Acknowledgements.** We thank Harry Buhrman, Carlos Palazuelos, David Perez-Garcia, Gilles Pisier, Oded Regev, Ignacio Villanueva and Ronald de Wolf for useful discussions. T.V. is grateful to CWI for hosting him while part of this work was done. J.B. thanks Universidad Complutense for hosting him while part of this work was done.

## References

- [ADR82] A. Aspect, J. Dalibard, and G. Roger. Experimental test of Bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.*, 49(25):1804–1807, 1982.

- [AGR81] A. Aspect, P. Grangier, and G. Roger. Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.*, 47(7):460–463, 1981.
- [BMMN11] M. Braverman, K. Makarychev, Y. Makarychev, and A. Naor. The Grothendieck constant is strictly smaller than Krivine's bound. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 453–462, oct. 2011.
- [BRSdW11] H. Buhrman, O. Regev, G. Scarpa, and R. de Wolf. Near-optimal and explicit Bell inequality violations. In *Proceedings of the 2011 IEEE 26th Annual Conference on Computational Complexity, CCC '11*, pages 157–166. IEEE Computer Society, 2011.
- [CHSH69] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, 1969.
- [CHTW04] R. Cleve, P. Høyer, B. Toner, and J. Watrous. Consequences and limits of nonlocal strategies. In *Proceedings of the 19th IEEE Conference on Computational Complexity (CCC 2004)*, pages 236–249, 2004.
- [CJPPG11] T. Cooney, M. Junge, C. Palazuelos, and D. Pérez-García. Rank-one quantum games. Technical report arXiv:1112.3563, 2011.
- [EPR35] A. Einstein, P. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 1935.
- [Gro53] A. Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. *Boletim Da Sociedade de Matemática de São Paulo*, 8:1, 1953.
- [Haa87] U. Haagerup. A new upper bound for the complex Grothendieck constant. *Israel Journal of Mathematics*, 60(2):199–224, 1987.
- [HM07] U. Haagerup and M. Musat. On the best constants in noncommutative Khintchine-type inequalities. *Journal of Functional Analysis*, 250(2):588 – 624, 2007.
- [HW71] D. L. Hanson and F. T. Wright. A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Statist.*, 3:1079–1083, Oct 1971.
- [JP11] M. Junge and C. Palazuelos. Large violation of Bell inequalities with low entanglement. *Communications in Mathematical Physics*, pages 1–52, 2011. 10.1007/s00220-011-1296-8.
- [JPPG<sup>+</sup>10] M. Junge, C. Palazuelos, D. Pérez-García, I. Villanueva, and M. Wolf. Unbounded violations of bipartite Bell inequalities via operator space theory. *Communications in Mathematical Physics*, 300:715–739, 2010. 10.1007/s00220-010-1125-5.
- [KRT10] J. Kempe, O. Regev, and B. Toner. Unique games with entangled provers are easy. *SIAM J. Comput.*, 39:3207–3229, July 2010.

- [Lat06] R. Latała. Estimates of moments and tails of Gaussian chaoses. *The Annals of Probability*, 34(6):2315–2331, November 2006.
- [Lou12] Elena R. Loubenets. Local quasi hidden variable modelling and violations of Bell-type inequalities by a multipartite quantum state. *Journal of Mathematical Physics*, 53(2):022201, 2012.
- [Mer90] N. D. Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65(15):1838–1840, Oct 1990.
- [NDT10] N.H. Nguyen, P. Drineas, and T.D. Tran. Tensor sparsification via a bound on the spectral norm of random tensors. *Technical report arXiv:1005.4732*, 2010.
- [Pal11] C. Palazuelos. Personal communication. 2011.
- [Pau92] V. I. Paulsen. Representations of function algebras, abstract operator spaces, and Banach space geometry. *Journal of Functional Analysis*, 109(1):113 – 129, 1992.
- [PGWP<sup>+</sup>08] D. Pérez-García, M. M. Wolf, C. Palazuelos, I. Villanueva, and M. Junge. Unbounded violation of tripartite Bell inequalities. *Communications in Mathematical Physics*, 279:455, 2008.
- [Pis99] G. Pisier. *The volume of convex bodies and Banach space geometry*. Cambridge University Press, 1999.
- [Pis12a] G. Pisier. Grothendieck’s theorem, past and present. *Bull. Amer. Math. Soc.*, 2012.
- [Pis12b] G. Pisier. Tripartite Bell inequality, random matrices and trilinear forms. Technical report arXiv:1203.2509, 2012.
- [Reg11] O. Regev. Bell violations through independent bases games. Technical report arXiv:1001.0576, 2011.
- [Sza82] S.J. Szarek. Nets of Grassmann manifold and orthogonal group. In *Proceedings of Banach Spaces Workshop*, pages 169–185. University of Iowa Press, 1982.
- [Tsi87] B. S. Tsirelson. Quantum analogues of the Bell inequalities. The case of two spatially separated domains. *J. Soviet Math.*, 36:557–570, 1987.
- [Ver10] Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. Available at <http://www-personal.umich.edu/~romanv/papers/>, 2010.
- [Zuk93] M. Zukowski. Bell theorem involving all settings of measuring apparatus. *Physics Letters A*, 177(4-5):290 – 296, 1993.