View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Caltech Authors - Main

arXiv:1711.10676v2 [quant-ph] 12 Jun 2018

ENTANGLEMENT IN NON-LOCAL GAMES AND THE HYPERLINEAR PROFILE OF GROUPS

WILLIAM SLOFSTRA* AND THOMAS VIDICK†

Abstract. We relate the amount of entanglement required to play linear-system non-local games near-optimally to the hyperlinear profile of finitely-presented groups. By calculating the hyperlinear profile of a certain group, we give an example of a finite non-local game for which the amount of entanglement required to play ϵ-optimally is at least $\Omega(1/\epsilon^k)$, for some $k > 0$. Since this function approaches infinity as ϵ approaches zero, this provides a quantitative version of a theorem of the first author.

1. Introduction

Let $\mathcal{G}$ be a two-player non-local game. Such a game is specified by finite sets of inputs X and Y and outputs A and B, associated with the first (Alice) and second (Bob) players respectively, a probability distribution π on input pairs $(x, y) \in X \times Y$, and a winning predicate $V \in \{0, 1\}^{A\times B\times X\times Y}$. A quantum strategy $\mathcal{S}$ for the game $\mathcal{G}$ is specified by Hilbert spaces H_A and H_B, a state $|\psi\rangle \in H_A \otimes H_B$, a positive operator-valued measure $\{A^a_x\}_{a\in A}$ on H_A for every $x \in X$, and a positive operator-valued measure $\{B^b_y\}_{b\in B}$ on H_B for every $y \in Y$. A strategy is finite-dimensional if at least one of $\dim(H_A)$ or $\dim(H_B)$ is finite; in this case we say that the strategy has dimension $\min\{\dim(H_A), \dim(H_B)\}$. The winning probability of strategy $\mathcal{S}$ in game $\mathcal{G}$ is defined as

$$\omega(\mathcal{G}; \mathcal{S}) = \sum_{(x,y)\in X\times Y} \pi(x, y) \sum_{(a,b)\in A\times B} V(a, b|x, y) \langle\psi| A^a_x \otimes B^b_y |\psi\rangle \ .$$

The quantum value $\omega^q(\mathcal{G})$ of $\mathcal{G}$ is the supremum of winning probabilities $\omega(\mathcal{G}; \mathcal{S})$ across all finite-dimensional quantum strategies $\mathcal{S}$. A basic question about $\mathcal{G}$ is: what is the minimum amount of entanglement $E(\mathcal{G}, \epsilon)$ used by a strategy that achieves winning probability at least $\omega^q(\mathcal{G}) - \epsilon$? Here we measure entanglement by the Schmidt rank. When lower bounds on $E(\mathcal{G}, \epsilon)$ are known, $\mathcal{G}$ can potentially be used to certify entangled states. As a result, many lower bound results exist [PGWP+08, JPPG+10, JP11, Slo11, Col16, CN16,

*Institute for Quantum Computing and Department of Pure Mathematics, University of Waterloo, Waterloo, Canada. email: weslofst@uwaterloo.ca.

†Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, USA. email: vidick@cms.caltech.edu.

CRSV17, JOP16, NV17, OV16]. In particular, it is known that $E(\mathcal{G}, \epsilon)$ can be arbitrarily large. For example, a result of Ostrev and the second author [OV16] states that for any $n \geq 1$, there is a two-player game $\mathcal{G}_n$ with input sets of size n and output sets of size 2, such that

$$E\Big(\mathcal{G}_n, O\Big(\frac{1}{n^{5/2}}\Big)\Big) \geq 2^{\Omega(\sqrt{n})} \ . \tag{1.1}$$

As with all known results of this type, the lower bound requires games of increasing size to get $E(\mathcal{G}_n, \epsilon) \to +\infty$. A recent result of the first author is that there is a fixed, constant-size non-local game $\mathcal{G}$ such that $E(\mathcal{G}, \epsilon) \to +\infty$ as $\epsilon \to 0$ [Slo17]. In other words, there is a game which cannot be played optimally using any finite-dimensional Hilbert space. The purpose of this paper is to prove the following quantitative version of this result:

Theorem 1.1. *Let $E(\mathcal{G}, \epsilon)$ be the smallest possible integer d such that there is a quantum strategy $\mathcal{S}$ of dimension d for $\mathcal{G}$ with success probability $\omega(\mathcal{G}; \mathcal{S}) \geq \omega^q(\mathcal{G}) - \epsilon$. Then there is a non-local game $\mathcal{G}$ and constants $C, C' > 0$ such that*

$$\frac{C}{\epsilon^{1/6}} \leq E(\mathcal{G}, \epsilon) \leq \frac{C'}{\epsilon^{1/2}}$$

for all $\epsilon \geq 0$.

The game $\mathcal{G}$ constructed in the proof of Theorem 1.1 is similar to the game constructed in [Slo17], and in particular is an example of a linear system non-local game with $\omega^q(\mathcal{G}) = 1$. Linear system games are a subclass of non-local games for which the existence of perfect quantum strategies is controlled by an associated finitely-presented group Γ, called the solution group. The proof of Theorem 1.1 is based on the observation that ϵ-perfect strategies for linear system games $\mathcal{G}$ correspond to approximate representations of the solution group Γ of $\mathcal{G}$. As a result, the function $E(\mathcal{G}, \epsilon)$ is linked to the dimension of approximate representations of the solution group Γ of $\mathcal{G}$. For sofic groups, the asymptotic dimension growth of sofic approximations is measured by the sofic profile of the group, a concept introduced by Cornulier [Cor13]. Theorem 1.1 can be thought of as a calculation of a "hyperlinear profile" of the solution group Γ, measuring the asymptotic dimension growth of unitary approximations to Γ.

Since hyperlinear profiles do not seem to have been studied heavily before, we explore some of their properties here. In particular, we give two related definitions of hyperlinear profile. The first, which we work with throughout the paper, is defined for finitely-presented groups, and is convenient for working with non-local games. The second matches Cornulier's definition of sofic profile, and in particular, is independent of the choice of presentation. Although we do not know if the two definitions are precisely the same, we show that they are the same under a natural equivalence relation.

In light of (1.1), we do not expect the inverse polynomial scaling from Theorem 1.1 to be optimal, and we hope that this initial lower bound opens the door to further results. We find it interesting that the theorem provides a finite test that may be executed on two spatially isolated quantum systems, such that the higher the success in the test, the larger the dimension that can be certified. The correspondence between near-optimal strategies for linear system games and approximate representations of solution groups may also be of independent interest. In particular, this correspondence implies that any ϵ-optimal strategy can be turned into an $O(\epsilon^{1/2})$-optimal strategy with a maximally entangled state (see Remark 5.7).

The rest of the paper is organized as follows. In Sections 2 and 3, we introduce our first definition of hyperlinear profile, and prove our main lower bound. In Section 4, we recall the notion of a solution group of a linear system game, and explain how hyperlinear profile is related to entanglement for strategies with maximally entangled states. In section 5, we show that hyperlinear profile is related to entanglement for strategies with general states. In Section 6, we use the embedding theorem of [Slo17] and the results of Section 3 to prove explicit bounds on the hyperlinear profile of the solution group, finishing the proof of Theorem 1.1. Finally, in Section 7 we give the second definition of hyperlinear profile, and compare the two definitions.

1.1. **Notation.** We use the following notation throughout the paper. $\mathcal{F}(S)$ is the free group generated by S, $\mathcal{U}(\mathbb{C}^d)$ is the unitary group of $\mathbb{C}^d$, and $M_d(\mathbb{C})$ is the set of $d \times d$ matrices. We use the following norms on $M_d(\mathbb{C})$: the operator norm $\|\cdot\|_{op}$, the Frobenius norm $\|\cdot\|_F$ defined by $\|A\|_F = \sqrt{\mathrm{Tr}(A^*A)}$, and the normalized Frobenius norm $\|\cdot\|_f = \|\cdot\|_F / \sqrt{d}$. Given a positive semidefinite matrix ρ, we also let $\|\cdot\|_\rho$ denote the seminorm defined by $\|A\|_\rho = \sqrt{\mathrm{Tr}(A^*A\rho)}$. Note that $\|\cdot\|_F = \|\cdot\|_{\mathbb{1}}$, $\|\cdot\|_f = \|\cdot\|_{\mathbb{1}/d}$, and $\|A\|_\rho = \left\|A\rho^{1/2}\right\|_F$. We use $\widetilde{\mathrm{tr}}$ for the linear functional $\mathrm{Tr}(\cdot)/d$ on $M_d(\mathbb{C})$.

1.2. **Acknowledgements.** We are indebted to Narutaka Ozawa for suggesting the use of the Connes embedding trick and the beautiful line of argument now incorporated in Section 5; this lead to a substantial improvement in our results. The first author also thanks Martino Lupini for helpful discussions.

The second author is supported by NSF CAREER Grant CCF-1553477, AFOSR YIP award number FA9550-16-1-0495, a CIFAR Azrieli Global Scholar award, and the IQIM, an NSF Physics Frontiers Center (NSF Grant PHY-1125565) with support of the Gordon and Betty Moore Foundation (GBMF-12500028).

2. Hyperlinear profile of finitely-presented groups

In this section, we state our first definition of hyperlinear profile, along with some basic properties. The starting point is the following definition from, e.g., [Slo17] or [HS17].

Definition 2.1 ([Slo17])**.** *An ϵ-representation of a finitely-presented group $G = \langle S : R\rangle$ is a homomorphism $\phi : \mathcal{F}(S) \to \mathcal{U}(\mathbb{C}^d)$ from the free group $\mathcal{F}(S)$ generated by S to the unitary group $\mathcal{U}(\mathbb{C}^d)$, such that*

$$\|\phi(r) - \mathbb{1}\|_f \leq \epsilon$$

for all $r \in R$.

An element $g \in G$ is non-trivial in approximate representations *if there is some representative $w \in \mathcal{F}(S)$ of g and $\delta > 0$ such that for all $\epsilon > 0$, there is an ϵ-representation ϕ with $\|\phi(w) - \mathbb{1}\|_f \geq \delta$.*

Suppose $\langle S : R\rangle$ and $\langle S' : R'\rangle$ are two presentations for a group G. While the set of ϵ-representations depends on the choice of presentation, any ϵ-representation with respect to $\langle S : R\rangle$ is an $O(\epsilon)$-representation with respect to $\langle S' : R'\rangle$ (with the constant depending on the isomorphism between the two presentations), and vice-versa. Similarly, whether g is non-trivial in approximate representations is independent of the choice of presentation. The dependence on δ is also somewhat arbitrary, due to the following well-known consequence of the tensor-power trick:

Lemma 2.2. *Suppose $g_1, \ldots, g_n \in G = \langle S : R\rangle$ are all non-trivial in approximate representations. Let $w_i \in \mathcal{F}(S)$ be a representative of g_i for each $1 \leq i \leq n$, and choose $\delta \in (0, \sqrt{2})$. Then for every $\epsilon > 0$, there is an ϵ-representation ϕ such that $\|\phi(w_i) - \mathbb{1}\|_f \geq \delta$ for all $1 \leq i \leq n$.*

A finitely-presented group is said to be *hyperlinear* if every non-trivial element is non-trivial in approximate representations. It is not known if there is a group which is not hyperlinear, and deciding this is a major problem in the field. For comparison, if every non-trivial element is non-trivial in finite-dimensional representations, then the group is said to be *residually-finite*, and there are examples of finitely-presented, hyperlinear, but non-residually-finite groups. Suppose G is such a group, so that there is $w \in \mathcal{F}(S)$ representing an element which is trivial in all finite-dimensional representations, but non-trivial in approximate representations. If we fix $\delta > 0$, then the dimension of ϵ-representations ϕ with $\|\phi(w) - \mathbb{1}\|_f \geq \delta$ must increase as $\epsilon \to 0$. The *hyperlinear profile* of G is a collection of functions which measure these growth rates.

Definition 2.3. *Let $G = \langle S : R\rangle$ be a finitely-presented group, and let T be a finite subset of $\mathcal{F}(S)$. The* (hyperlinear) profile *of T is the function*

$\mathrm{hlp}(T) : \mathbb{R}_{>0} \times \mathbb{R}_{>0} \to \mathbb{N} \cup \{+\infty\}$ *such that* $\mathrm{hlp}(T; \delta, \epsilon)$ *is the smallest integer* d *for which there is an* ϵ*-representation* ϕ *of dimension* d *with*

$$\|\phi(w) - \mathbb{1}\|_f \geq \delta \quad \text{for all } w \in T,$$

or $+\infty$ *if no such* d *exists.*

If $T = \{w\}$*, then we write* $\mathrm{hlp}(w; \delta, \epsilon)$ *for* $\mathrm{hlp}(T; \delta, \epsilon)$*. The hyperlinear profile of* G *is the collection of functions* $\mathrm{hlp}(T)$*, where* T *is a finite subset of* $\mathcal{F}(S)$ *not containing any element with trivial image in* G*.*

If T does contain an element with trivial image, then $\mathrm{hlp}(T; \delta, \epsilon)$ will always be infinite for small enough ϵ. It is clear that $\mathrm{hlp}(T)$ is non-decreasing in δ and non-increasing in ϵ. Another easy property of the hyperlinear profile is that it is non-decreasing under homomorphisms, in the following sense:

Lemma 2.4. *Let* $G_i = \langle S_i : R_i \rangle$*,* $i = 1, 2$*, be two finitely-presented groups, and suppose* $\phi : \mathcal{F}(S_1) \to \mathcal{F}(S_2)$ *is a homomorphism which descends to a homomorphism* $G_1 \to G_2$*. Then there is a constant* $C \geq 1$ *such that*

$$\mathrm{hlp}(T, \delta, C\epsilon) \leq \mathrm{hlp}(\phi(T), \delta, \epsilon) \ ,$$

for any finite set $T \subset \mathcal{F}(S_1)$ *and* $\epsilon, \delta > 0$*.*

Note that while the elements $\phi(R_1)$ are trivial in G_2, they do not necessarily appear in R_2. The constant C in Lemma 2.4 depends on how many times the relations from R_2 must be applied to show that the elements $\phi(R_1)$ are trivial.

We are mainly interested in the asymptotic behaviour of $\mathrm{hlp}(T)$ as $\epsilon \to 0$ with δ fixed. When ϕ induces an isomorphism between G_1 and G_2, Lemma 2.4 shows that the asymptotics of the functions $\mathrm{hlp}(T)$ are somewhat independent of the choice of presentation. We can say the same thing about the choice of representatives in $\mathcal{F}(S)$:

Lemma 2.5. *Suppose* $w_0, w_1 \in \mathcal{F}(S)$ *represent the same element in* $G = \langle S : R \rangle$*. Then there is a constant* $C \geq 1$ *such that*

$$\mathrm{hlp}(w_0, \delta - C\epsilon, \epsilon) \leq \mathrm{hlp}(w_1, \delta, \epsilon)$$

for all $\epsilon > 0$ *and* $\delta > C\epsilon$*.*

Although we won't use it, the following lemma puts some limits on how much the choice of δ can affect the asymptotics:

Lemma 2.6. *Let* $G = \langle S : R \rangle$*, and* $0 < \delta < \delta' < \sqrt{2}$*. Then there is a constant* $k \geq 1$ *such that*

$$\mathrm{hlp}(T, \delta, \epsilon) \leq \mathrm{hlp}(T, \delta', \epsilon) \leq k\, \mathrm{hlp}(T, \delta, \epsilon/k)^k$$

for all $\epsilon > 0$ *and finite subsets* $T \subset \mathcal{F}(S)$*.*

For instance, the choice of δ within the range $(0, \sqrt{2})$ does not affect whether hlp is polynomial or exponential in $1/\epsilon$.

Proof. We only need to prove the second inequality. The proof is similar to Lemma 2.2, in that we can use the tensor power trick. Suppose ϕ is an ϵ-representation of dimension d with

$$\|\phi(w) - \mathbb{1}\|_f \geq \delta$$

for all $w \in T$. Let ψ be the direct sum of ϕ with its complex conjugate, and with $2d$ copies of the trivial representation. Then ψ is an $\epsilon/2$-representation of dimension $4d$, such that $\widetilde{\mathrm{tr}}(\psi(w))$ is real and non-negative for all $w \in \mathcal{F}(w)$, and $\|\psi(w) - \mathbb{1}\|_f \geq \delta/2$ for all $w \in T$. Since $\|\psi(w) - \mathbb{1}\|_f^2 = 2 - 2\widetilde{\mathrm{tr}}(\psi(w))$, we have that $\widetilde{\mathrm{tr}}(\psi(w)) \leq 1 - \delta^2/8$. Hence

$$\left\|\psi(w)^{\otimes n} - \mathbb{1}\right\|_f^2 = 2 - 2\widetilde{\mathrm{tr}}(\psi(w))^n \geq 2 - 2\left(1 - \frac{\delta^2}{8}\right)^n .$$

By choosing n large enough, we can make the right-hand side larger than $(\delta')^2$. Since $\psi^{\otimes n}$ is a $(n\epsilon/2)$-representation of dimension $(4d)^n$, we conclude that

$$\mathrm{hlp}\left(T, \delta', \frac{n\epsilon}{2}\right) \leq 4^n \,\mathrm{hlp}(T, \delta, \epsilon)^n,$$

and the lemma follows. □

Note that we have not restricted Definition 2.3 to hyperlinear groups. In terms of hyperlinear profile, a finitely-presented group $G = \langle S : R\rangle$ is hyperlinear if and only if $\mathrm{hlp}(T, \delta, \epsilon) < +\infty$ for all finite subsets $T \subset \mathcal{F}(S)$ not containing any elements with trivial image in G, and real numbers $\delta \in (0, \sqrt{2})$, $\epsilon > 0$. Similarly, G is residually finite if and only if $\mathrm{hlp}(T, \delta, \epsilon)$ is bounded as $\epsilon \to 0$ for all finite subsets $T \in \mathcal{F}(S)$ (again, not containing any elements with trivial image in G) and $\delta \in (0, \sqrt{2})$.

Although we can choose any δ in $(0, \sqrt{2})$ when measuring $\mathrm{hlp}(T, \delta, \epsilon)$, it can make sense to take δ to be greater than or equal to $\sqrt{2}$. The largest value of $\|U - \mathbb{1}\|_f$ when U is a unitary is 2, and this is achieved when $U = -\mathbb{1}$. Thus, $\mathrm{hlp}(w, 2, \epsilon)$ measures the growth rate of ϵ-representations ϕ where $\phi(w) = -\mathbb{1}$. The following proposition shows that when $w \in \mathcal{F}(S)$ represents a central involution in G (which will be the case in the main example of the next section, and also when working with linear system games), $\delta = 2$ is a natural choice.

Proposition 2.7. *Suppose $g \in G$ is a central involution, and $G = \langle S : R\rangle$ is a presentation such that g has a representative $w \in S$. Let $0 < \delta \leq 2$. Then there is a constant $C \geq 1$ (depending on δ) such that*

$$\mathrm{hlp}(w, 2, C\epsilon) \leq \mathrm{hlp}(w, \delta, \epsilon) \leq \mathrm{hlp}(w, 2, \epsilon)$$

for all $\epsilon > 0$.

The proof of Proposition 2.7 relies on some simple stability properties.

Lemma 2.8.

(a) *Suppose X is a matrix with $\|X\|_{op} \leq 1$. Then there is a unitary matrix U with $\|X - U\|_f \leq \|X^*X - \mathbb{1}\|_f$, and in fact U can be any unitary in a polar decomposition $X = UP$ of X.*

(b) *If X, Y are unitary matrices and $X^2 = \mathbb{1}$, then there is a unitary matrix Z such that $ZX = XZ$ and*

$$\|Y - Z\|_f \leq \|XY - YX\|_f .$$

(c) *If X is any normal matrix, there is a self-adjoint matrix Z with $Z^2 = \mathbb{1}$ such that*

$$\|X - Z\|_f \leq 2 \left\|X^2 - \mathbb{1}\right\|_f .$$

We expect that Lemma 2.8 is well-known to experts. We give the proof of parts (a) and (b) for the convenience of the reader. A proof of part (c) can be found in [Slo17, Lemma 3.6].

Proof of Lemma 2.8. For part (a), let $X = UP$ be a polar decomposition of X, where U is unitary. Then

$$\|X - U\|_f = \|P - \mathbb{1}\|_f = \left\|\sqrt{X^*X} - \mathbb{1}\right\|_f \leq \|X^*X - \mathbb{1}\|_f ,$$

where the last inequality comes from the fact that $1 - z \leq 1 - z^2$ for all $z \in [0, 1]$.

For part (b), let $Z_0 = \frac{1}{2}(Y + XYX)$. Then Z_0 and X commute, and $\|Y - Z_0\|_f$ and $\|Z_0^* Z_0 - \mathbb{1}\|_f$ are both at most $\frac{1}{2}\|XY - YX\|_f$. So part (b) follows from applying part (a) to Z_0. □

We also need that small perturbations of ϵ-representations remain $O(\epsilon)$-representations.

Lemma 2.9 ([Slo17], Lemma 2.3)**.** *Let ϕ be a d-dimensional ϵ-representation of the finitely-presented group $G = \langle S : R \rangle$. If $\psi : \mathcal{F}(S) \to \mathcal{U}(\mathbb{C}^d)$ is a homomorphism such that*

$$\|\psi(s) - \phi(s)\|_f \leq \epsilon'$$

for all $s \in S$, then ψ is an $(\epsilon + O(\epsilon'))$-representation, where the constant depends on the length of the relations in R.

Proof of Proposition 2.7. We only need to prove the first inequality. Suppose ϕ is a d-dimensional ϵ-representation of G with $\|\phi(w) - \mathbb{1}\|_f \geq \delta$. Then by Lemmas 2.8 and 2.9, there is an $O(\epsilon)$-representation ψ such that $\psi(w)$ is a central involution in the group generated by $\psi(S)$, and $\|\phi(s) - \psi(s)\| \leq O(\epsilon)$

for all $s \in S$. If ϵ is small enough such that $\|\phi(w) - \psi(w)\| \leq \delta/2$, we then have $\|\psi(w) - \mathbb{1}\|_f \geq \delta/2$. This inequality implies that the dimension of the (-1)-eigenspace of $\psi(w)$ is bounded below by $\Omega(d\delta^2)$. If P is the projection on the (-1)-eigenspace of $\psi(w)$, then it is not hard to see that $P\psi P$ is an $O(\epsilon/\delta)$-representation of G (see for instance the proof of Lemma 3.9 of [Slo17]). Since δ is fixed, there is a constant C such that for every d-dimensional ϵ-representation, there is a $C\epsilon$-representation ψ of dimension at most d with $\psi(w) = -\mathbb{1}$. □

3. Bounds on hyperlinear profile for a specific group

The question of whether there is a non-hyperlinear group is notoriously difficult, and it seems reasonable to look for related questions which might be more approachable. In this context, it seems natural to seek examples of groups for which the hyperlinear profile grows particularly fast. We note that very little seems to be known about this question. In the case of the sofic profile, any non-residually-finite group has a sofic profile at least as large as $1/\epsilon$, essentially because the smallest positive value of the normalized Hamming metric on the symmetric group S_n is $1/n$ [Cor13, Fact 3.8]. In contrast, for the hyperlinear profile, it is not clear that there is even a "default" lower bound of this form. The only explicit lower bound we are aware of prior to our work is the following example due to Tobias Fritz [Fri13]:

Example 3.1. *Let G be the Baumslag-Solitar group* $\mathrm{BS}(2,3)$*, so*

$$G = \langle u, v : v^{-1}u^2v = u^3 \rangle.$$

It is shown in [Fri13] *that there is a constant $C > 0$ such that if $\phi : \mathcal{F}(u,v) \to \mathcal{U}(\mathbb{C}^d)$ is a homomorphism with*

$$\left\|\phi(v)^{-1}\phi(u)^2\phi(v) - \phi(u)^3\right\|_{op} \leq \epsilon,$$

for some $\epsilon > 0$, then

$$\|\phi(w) - \mathbb{1}\|_{op} \leq e^{Cd^2}\epsilon,$$

where $w = uv^{-1}uvu^{-1}v^{-1}u^{-1}v$. Note that $w \neq e$ in G. If ϕ is an ϵ-representation of G then

$$\left\|\phi(v)^{-1}\phi(u)^2\phi(v) - \phi(u)^3\right\|_{op} \leq \sqrt{d}\left\|\phi(v)^{-1}\phi(u)^2\phi(v) - \phi(u)^3\right\|_f \leq \sqrt{d}\epsilon.$$

Consequently

$$\|\phi(w) - \mathbb{1}\|_f \leq \|\phi(w) - \mathbb{1}\|_{op} \leq e^{Cd^2}\sqrt{d}\epsilon,$$

from which by rearranging terms we get the lower bound

$$\mathrm{hlp}(w, \delta, \epsilon) = \Omega\big(\sqrt{\ln(\delta/\epsilon)}\big)$$

for any $\delta > 0$.

The main result of this section, and the key result of this paper, is that there is a group with a hyperlinear profile somewhere between $1/\epsilon^{2/3}$ and $1/\epsilon$, up to constants.

Proposition 3.2. *Define*

$$\begin{aligned} K = \langle a, b, c, x, y \ : & \ xyx^{-1} = y^2, xcx^{-1} = c, \\ & \ yay^{-1} = b, yby^{-1} = a, \\ & \ c = ab, a^2 = b^2 = c^2 = e \rangle. \end{aligned}$$

For any $0 < \delta \leq 2$ there exists $C, C' > 0$ such that

$$\frac{C}{\epsilon^{\frac{2}{3}}} \leq \mathrm{hlp}(c, \delta, \epsilon) \leq \frac{C'}{\epsilon} \tag{3.1}$$

for all $\epsilon > 0$.

For the remainder of this section, we let K denote the finitely-presented group defined in Proposition 3.2. The group is similar to the group defined in [Slo17, Section 5]. In particular, K is sofic, and hence hyperlinear. The element c is a central involution which is trivial in all finite-dimensional representations, so K is non-residually finite. Because we have chosen a presentation for K in which c is a generator, Proposition 2.7 implies that we only need to prove Proposition 3.2 when $\delta = 2$. In fact, it is not necessary to make c a generator—we could replace c with ab, the relation $c = ab$ with $[a, b] = e$, and remove c from the presentation—but we keep the slightly redundant presentation for simplicity.

The proof of Proposition 3.2 is split into several steps. For both the upper and lower bound, it is helpful to look at the subgroup

$$K_0 = \langle y, a, b, c : yay^{-1} = b, yby^{-1} = a, c = ab, a^2 = b^2 = c^2 = e \rangle$$

of K. Abstractly, K_0 is the semidirect product $\mathbb{Z} \ltimes \mathbb{Z}_2 \times \mathbb{Z}_2$ of $\mathbb{Z}_2 \times \mathbb{Z}_2$ by the automorphism switching the order of the factors. Given $U \in \mathcal{U}(\mathbb{C}^d)$, we can define a $2d$-dimensional representation ϕ of K_0 by

$$\phi(a) = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \phi(b) = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \text{ and } \phi(y) = \begin{pmatrix} 0 & \mathbb{1} \\ U & 0 \end{pmatrix}. \tag{3.2}$$

The first lemma needed for the lower bound in Proposition 3.2 is that every d-dimensional ϵ-representation of K can be turned into a $2d$-dimensional $O(\epsilon)$-representation which restricts to a representation of K_0 that satisfies (3.2).

Lemma 3.3. *Let ϕ be a d-dimensional ϵ-representation of K such that $\phi(c) = -\mathbb{1}$. Then there is a $2d$-dimensional $O(\epsilon)$-representation ψ of K_0 such that*

$$\psi(a) = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \psi(b) = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \quad \text{and} \quad \psi(y) = \begin{pmatrix} 0 & \mathbb{1} \\ U & 0 \end{pmatrix},$$

where the blocks are of size $d \times d$, and $U \in \mathcal{U}(\mathbb{C}^d)$.

Proof. Since $\|\phi(a)^2 - \mathbb{1}\|_f \leq \epsilon$, part (c) of Lemma 2.8 implies that there is a unitary matrix Z such that $\|\phi(a) - Z\|_f \leq 2\epsilon$ and $Z^2 = \mathbb{1}$. By Lemma 2.9, replacing $\phi(a)$ with Z and $\phi(b)$ with $-Z$ yields an $O(\epsilon)$-representation α with $\alpha(a)^2 = \alpha(b)^2 = \mathbb{1}$ and $\alpha(a)\alpha(b) = -\mathbb{1} = \alpha(c)$. Let $\beta = \alpha \oplus \widetilde{\alpha}$, where $\widetilde{\alpha}$ is the $O(\epsilon)$-representation defined from α by switching a and b, so

$$\widetilde{\alpha}(x) = \alpha(x), \widetilde{\alpha}(y) = \alpha(y), \widetilde{\alpha}(a) = \alpha(b), \widetilde{\alpha}(b) = \alpha(a), \text{ and } \widetilde{\alpha}(c) = -\mathbb{1}.$$

Then $\beta(a)^2 = \beta(b)^2 = \mathbb{1}$, $\beta(a)\beta(b) = -\mathbb{1} = \beta(c)$, and $\mathrm{tr}(\beta(a)) = 0$, so there is a choice of basis such that

$$\beta(a) = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \quad \text{and} \quad \beta(b) = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix},$$

where the blocks are of size $d \times d$.

Claim 3.4. *Let*

$$B = \begin{pmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{pmatrix}$$

be a $2d \times 2d$ unitary matrix such that $\|B\beta(a) - \beta(b)B\|_f \leq \epsilon'$, for some $\epsilon' \geq 0$. Then there are $d \times d$ unitaries V_2 and V_3 such that

$$\left\| B - \begin{pmatrix} 0 & V_2 \\ V_3 & 0 \end{pmatrix} \right\|_f = O(\epsilon').$$

Proof. Let

$$Y = \begin{pmatrix} 0 & Y_2 \\ Y_3 & 0 \end{pmatrix}.$$

The assumption made in the claim implies that $\|Y - B\|_f \leq \epsilon'/2$, and since B is assumed unitary, $\|Y^*Y - \mathbb{1}\|_f \leq O(\epsilon')$ (to avoid having to add $(\epsilon')^2$ to the right side of this inequality, note that we can assume $\epsilon' \leq 2$). It follows that $\|Y_i^*Y_i - \mathbb{1}\|_f \leq O(\epsilon')$ for $i = 2, 3$. Since B is unitary, $\|Y_i\|_{op} \leq 1$ for $i = 2, 3$, and hence by part (a) of Lemma 2.8 there are unitaries V_i with $\|Y_i - V_i\|_f \leq O(\epsilon')$, $i = 2, 3$. □

Let

$$\beta(y) = \begin{pmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{pmatrix}.$$

Applying Claim 3.4 to $B = \beta(y)$, there exists unitaries V_2 and V_3 such that

$$\left\| B - \begin{pmatrix} 0 & V_2 \\ V_3 & 0 \end{pmatrix} \right\|_f = O(\epsilon).$$

Applying Lemma 2.9 once again, we conclude that the homomorphism $\gamma : \mathcal{F}(\{a,b,c,x,y\}) \to \mathcal{U}(\mathbb{C}^{2d})$ defined by

$$\gamma(y) = \begin{pmatrix} 0 & V_2 \\ V_3 & 0 \end{pmatrix}, \gamma(s) = \beta(s) \text{ for } s \in \{a,b,c,x\}$$

is an $O(\epsilon)$-representation. Then

$$\psi = \begin{pmatrix} V_2^* & 0 \\ 0 & \mathbb{1} \end{pmatrix} \cdot \gamma \cdot \begin{pmatrix} V_2 & 0 \\ 0 & \mathbb{1} \end{pmatrix}$$

satisfies the conditions of the lemma with $U = V_3 V_2$. □

With Lemma 3.3, we can prove the lower bound in (3.1).

Lemma 3.5. *There is a constant C such that*

$$\mathrm{hlp}(c,2,\epsilon) \geq \frac{C}{\epsilon^{\frac{2}{3}}}, \tag{3.3}$$

for all $\epsilon > 0$.

Proof. Let ϕ be a d-dimensional ϵ-representation of K such that $\|\phi(c) - \mathbb{1}\|_f \geq 2$, i.e. $\phi(c) = -\mathbb{1}$. Applying Lemma 3.3 we deduce the existence of a unitary $U \in \mathrm{U}(\mathbb{C}^{2d})$, and unitaries $X = \psi(x), Y = \psi(y) \in \mathrm{U}(\mathbb{C}^{2d})$ such that

$$Y = \begin{pmatrix} 0 & \mathbb{1} \\ U & 0 \end{pmatrix} \qquad \text{and} \qquad \left\|XYX^\dagger - Y^2\right\|_{op} \leq C_1 \sqrt{d}\, \epsilon, \tag{3.4}$$

for some constant $C_1 > 0$ (the second condition uses $\|\cdot\|_{op} \leq \sqrt{d}\|\cdot\|_f$). Let $\epsilon_1 = C_1\sqrt{d}\epsilon$. We show that satisfying both conditions requires $d \geq C_2\epsilon_1^{-1}$ for some constant $C_2 > 0$; this will prove the lemma.

Let $\Theta \subseteq \mathbb{R}$ be such that the elements ($\theta \bmod 1$) are pairwise distinct for $\theta \in \Theta$, and the set of eigenvalues of U is $\Lambda = \{e^{2i\pi\theta}, \theta \in \Theta\}$. We prove a lower bound on d by showing that the set Θ must be large.

For $a, b \in \mathbb{R}$, let $d(a,b)$ measure the arc-length distance between $x = e^{2i\pi a}$ and $y = e^{2i\pi b}$. We first prove a simple lemma that will facilitate manipulation of angles.

Claim 3.6. *Let $a, b \in \mathbb{R}$, and set $x = e^{2i\pi a}$ and $y = e^{2i\pi b}$. Then $d(a,b) \leq \pi|x-y|$. Moreover, for any square root $x' = e^{2i\pi a'}$ of x there is a square root $y' = 2^{2i\pi b'}$ of y such that $d(a',b') = \frac{1}{2}d(a,b)$.*

Proof. The statement of the lemma is invariant under rotation, so we may assume $a = 0$, $x = 1$, and $b \in (-1/2, 1/2]$. If $|x-y| \geq 1$ the bound on $d(a,b)$ is trivial, so for the first part of the lemma we can assume in particular that $b \in [-1/4, 1/4)$. Then

$$d(b,0) = 2\pi|b|$$

$$\leq \pi|\sin(2\pi b)|$$
$$\leq \pi|1-y|.$$

For the moreover part of the lemma, note that the square roots $x' = e^{2i\pi a'}$ of x are such that $a' \in \frac{1}{2}\mathbb{Z}$. For any such square root, $y' = e^{2i\pi b'}$ for $b' = a' + \frac{1}{2}b$ is a square root of y such that $d(b', a') = d(b/2, 0) = d(b, 0)/2$. □

The next claim shows that as soon as $\Theta \neq \emptyset$ it must contain many distinct values.

Claim 3.7. *Let $\theta \in \Theta$. For every integer $\ell \geq 0$ there exists angles*

$$\left\{\theta_j^{(\ell)},\, j \in \{1, \ldots, 2^\ell\}\right\} \subseteq \Theta$$

such that for each $j \in \{1, \ldots, 2^\ell\}$,

$$d\big(\theta_j^{(\ell)},\, 2^{-\ell}\theta + j2^{-\ell}\big) \;\leq\; 2\pi\epsilon_1. \tag{3.5}$$

Proof. If $\epsilon_1 \geq 1$ the claim is trivial, so assume $\epsilon_1 < 1$. We prove the claim by induction on ℓ. The base case $\ell = 0$ is obtained by setting $\theta_j^{(0)} = \theta$. Assume the statement of the claim true for some $\ell \geq 0$. Let $j \in \{1, \ldots, 2^\ell\}$. By assumption $d(\theta_j^{(\ell)}, 2^{-\ell}(\theta + j)) \leq 2\pi\epsilon_1$, which using Claim 3.6 implies that $\lambda_j^{(\ell)} = e^{2i\pi\theta_j^{(\ell)}}$ has square roots $y_1 = e^{2i\pi\varphi_1}$ and $y_2 = e^{2i\pi\varphi_2}$ such that

$$\max\left\{d\big(\varphi_1, 2^{-(\ell+1)}(\theta + j)\big),\, d\big(\varphi_2, 2^{-(\ell+1)}(\theta + j + 2^\ell)\big)\right\} \;\leq\; \pi\epsilon_1. \tag{3.6}$$

Since $\theta_j^{(\ell)} \in \Theta$, $\lambda_j^{(\ell)}$ is an eigenvalue of U, which using the first equation in (3.4) implies that both y_1 and y_2 are eigenvalues of Y. Using the second equation in (3.4) there must exist eigenvalues $\lambda_{2j-1}^{(\ell+1)} = e^{2i\pi\theta_{2j-1}^{(\ell+1)}}$ and $\lambda_{2j}^{(\ell+1)} = e^{2i\pi\theta_{2j}^{(\ell+1)}}$ of Y^2 such that

$$\max\left\{\left|\lambda_{2j-1}^{(\ell+1)} - y_1\right|,\, \left|\lambda_{2j}^{(\ell+1)} - y_2\right|\right\} \;\leq\; \epsilon_1.$$

Using Claim 3.6, this implies

$$\max\left\{d(\theta_{2j-1}^{(\ell+1)}, \varphi_1),\, d(\theta_{2j}^{(\ell+1)}, \varphi_2)\right\} \;\leq\; \pi\epsilon_1.$$

Using the triangle inequality and (3.6),

$$\max\left\{d\big(\theta_{2j}^{(\ell+1)}, 2^{-(\ell+1)}(\theta + j)\big),\, d\big(\theta_{2j-1}^{(\ell+1)}, 2^{-(\ell+1)}(\theta + j + 2^\ell)\big)\right\} \leq \pi\epsilon_1 + \pi\epsilon_1,$$

which (after relabeling the $\{\theta_j^{(\ell+1)},\, j \in \{1, \ldots, 2^{\ell+1}\}\}$) completes the induction. □

As long as $2\pi\epsilon_1 < 2^{-(\ell+1)}\pi$ the bound (3.5) implies that all $\theta_j^{(\ell)}$ are distinct modulo 1, and hence that $d \geq 2^\ell$. If we choose ℓ so that $2^{-(\ell+2)} \leq 2\epsilon_1 \leq 2^{-(\ell+1)}$, then $d \geq 2^\ell > 1/(8\epsilon_1)$, proving the lemma. □

Finally we prove the key lemma for the upper bound.

Lemma 3.8. *Let $d = 2^\ell$, where $\ell \geq 2$. Then there is a homomorphism*

$$\phi : \mathcal{F}(\{a, b, c, x, y\}) \to \mathcal{U}(\mathbb{C}^d)$$

such that $\phi(c) = -\mathbb{1}$, and

$$\|\phi(r) - \mathbb{1}\|_{op} \leq 2^{-\ell}$$

for all defining relations r of K.

Proof. Let $d = 2^\ell$ be a power of two. Let $\{f_i,\, i \in \{1, \ldots, d\}\}$ denote the canonical basis of $\mathbb{C}^d$. Let

$$A = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad B = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}$$

in this basis, where each block has size $d/2$. Let

$$Y = \sum_{j=1}^{d/2} f_j {f_{\frac{d}{2}+j}}^* + \sum_{j=1}^{d/2} e^{2i\pi\theta_j} f_{\frac{d}{2}+j} {f_j}^*,$$

where for $j \in \{1, \ldots, d/2\}$, $\theta_j = j2^{-(\ell-1)}$. Then Y is unitary and satisfies $YAY^\dagger = B$ and $YBY^\dagger = A$. For $j \in \{1, \ldots, d/2\}$, let

$$u_j = \frac{f_j + e^{2i\pi\frac{\theta_j}{2}} f_{\frac{d}{2}+j}}{\sqrt{2}}, \qquad v_j = \frac{f_j - e^{2i\pi\frac{\theta_j}{2}} f_{\frac{d}{2}+j}}{\sqrt{2}}.$$

We can diagonalize Y as

$$Y = \sum_{j=1}^{d/2} e^{2i\pi\frac{\theta_j}{2}} u_j u_j^* + \sum_{j=1}^{d/2} e^{2i\pi\left(\frac{\theta_j}{2}+\frac{1}{2}\right)} v_j v_j^*,$$

and verify that $Y^2 = \sum_{j=1}^{d/2} e^{2i\pi\theta_j}(f_j f_j^* + f_{\frac{d}{2}+j} f_{\frac{d}{2}+j}^*)$. Let X be the unitary defined as follows: if j is even, $Xu_j = f_{\frac{j}{2}}$ and $Xv_j = f_{\frac{d}{4}+\frac{j}{2}}$, while if j is odd, $Xu_j = f_{\frac{d}{2}+\frac{j+1}{2}}$ and $Xv_j = f_{\frac{3d}{4}+\frac{j+1}{2}}$. We verify that X maps eigenvectors of Y to eigenvectors of Y^2 in a way such that the associated eigenvalues always differ by at most $2^{-\ell}$.

For the case of even j, the eigenvalue of Y associated to u_j is the same as the eigenvalue of Y^2 associated to $f_{\frac{j}{2}}$. The eigenvalue of Y associated to v_j is $e^{2i\pi(\frac{j2^{-(\ell-1)}}{2}+\frac{1}{2})}$, while the eigenvalue of Y^2 associated to $f_{\frac{d}{4}+\frac{j}{2}}$ is $e^{2i\pi(\frac{d}{4}+\frac{j}{2})2^{-(\ell-1)}}$, which given $d = 2^\ell$ are identical.

For the case of odd j, the difference between the eigenvalue of Y associated to v_j and the eigenvalue of Y^2 associated to $f_{\frac{3d}{4}+\frac{j+1}{2}}$ is

$$\left|\left(\frac{\theta_j}{2} + \frac{1}{2}\right) - \theta_{\frac{3d}{4}+\frac{j+1}{2}} \mod 1\right|$$

$$= \left| \left(j2^{-\ell} + \frac{1}{2} \right) - \left(\frac{j+1}{2} 2^{-(\ell-1)} + \frac{3 \cdot 2^{\ell}}{4} 2^{-(\ell-1)} \right) \mod 1 \right|$$
$$= 2^{-\ell}.$$

A similar calculation holds for the remaining case. Thus $\|XYX^\dagger - Y^2\| \leq 2^{-\ell}$. The proof of the lemma is concluded by defining a homomorphism $\phi : \mathcal{F}(\{a, b, x, y\}) \to \mathcal{U}(\mathbb{C}^d)$ by $\phi(a) = A$, $\phi(b) = B$, $\phi(x) = X$ and $\phi(y) = Y$. □

Proof of Proposition 3.2. Using Proposition 2.7, the lower bound is shown in Lemma 3.5. The upper bound follows directly from Lemma 3.8 by noting that $\|\cdot\|_f \leq \|\cdot\|_{op}$. □

4. Linear system games

We now turn to the connection between hyperlinear profile and linear system games. Let $Ax = b$ be an $m \times n$ linear system over $\mathbb{Z}_2$. For every $1 \leq i \leq m$, let V_i be set of indices of variables which appear in equation i, i.e. $V_i = \{1 \leq j \leq n : A_{ij} \neq 0\}$. The linear system non-local game associated to $Ax = b$ is played as follows [CM14]:

(1) The first player (Alice) is given an integer i, chosen uniformly at random from $\{1, \ldots, m\}$, and must reply with a vector $y \in \mathbb{Z}_2^{V_i}$ satisfying $\sum_{k \in V_i} y_k = b_i$. The output is interpreted as an assignment to the variables in equation i.

(2) The second player (Bob) is given an integer j, chosen uniformly at random from $\{1, \ldots, n\}$, and must reply with $x_j \in \mathbb{Z}_2$, interpreted as an assignment to the variable x_j.

(3) The two players win if either $j \notin V_i$, or $y_j = x_j$.

A quantum strategy $\mathcal{S}$ for the game (presented in terms of observables) consists of

(a) two Hilbert spaces H_A and H_B, along with a unit vector $|\psi\rangle \in H_A \otimes H_B$,

(b) a collection of $\{\pm 1\}$-valued observables Y_{ij}, $1 \leq i \leq m$, $j \in V_i$, on H_A such that:

  (i) $\prod_{j \in V_i} Y_{ij} = (-\mathbb{1})^{b_i}$ for all $1 \leq i \leq m$, and

  (ii) $Y_{ij} Y_{ij'} = Y_{ij'} Y_{ij}$ for all $1 \leq i \leq m$, $j, j' \in V_i$,

  and

(c) a collection of $\{\pm 1\}$-valued observables X_j, $1 \leq j \leq n$, on H_B.

As mentioned in the introduction, a quantum strategy for a non-local game is usually presented in terms of POVMs, rather than observables. A strategy as above, presented in terms of observables, is equivalent to a strategy in the usual form, in which the measurements are projective [CM14, Slo17]. Indeed, if $\{A_i^y\}_{y\in\mathbb{Z}_2^{V_i}}$, $\{B_j^c\}_{c\in\mathbb{Z}_2}$, $|\psi\rangle$ is a strategy for a linear system game in which $\{A_i^y\}_y$ and $\{B_j^c\}_c$ are projective measurements, then we can express this strategy in terms of observables by setting $Y_{ij} = \sum_y (-1)^{y_j} A_i^y$ for each $1 \le i \le m$, $j \in V_i$, and $X_j = B_j^0 - B_j^1$. In this way, the observables Y_{ij} and X_j correspond to Alice's and Bob's assignments to variable x_j.

Using Naimark dilation, any strategy with POVMs can be turned into a strategy with projective measurements. This multiplies the dimension of the Hilbert space by the number of outputs.

Lemma 4.1. *Let $\{A_i^a\}_{a=1}^k$ and $\{B_j^b\}_{b=1}^\ell$ be two families of POVMs, on Hilbert spaces H_A and H_B respectively, and let $|\psi\rangle$ be a state in $H_A \otimes H_B$. Then there are families of projective measurements $\{\widehat{A}_i^a\}_{a=1}^k$ and $\{\widehat{B}_j^b\}_{b=1}^\ell$, on Hilbert spaces $\widehat{H}_A$ and $\widehat{H}_B$ respectively, and a state $|\widehat{\psi}\rangle \in \widehat{H}_A \otimes \widehat{H}_B$, such that*

$$\langle\psi| A_i^a \otimes B_j^b |\psi\rangle = \langle\widehat{\psi}| \widehat{A}_i^a \otimes \widehat{B}_j^b |\widehat{\psi}\rangle$$

for all i, j, a, b. Furthermore, the Hilbert spaces $\widehat{H}_A$ and $\widehat{H}_B$ can be chosen so that $\dim \widehat{H}_A = k \dim H_A$ and $\dim \widehat{H}_B = \ell \dim H_B$.

As a result, it suffices to prove Theorem 1.1 for strategies with projective measurements. Thus for the remainder of the paper, we assume that all strategies are projective, and presented in terms of observables.

Given a strategy $\mathcal{S}$ for a non-local game presented in terms of observables as above, if $j \in V_i$, then Alice and Bob win with probability $p_{ij}(\mathcal{S}) = (b_{ij} + 1)/2$ on inputs (i, j), where

$$b_{ij}(\mathcal{S}) := \langle\psi| Y_{ij} \otimes X_j |\psi\rangle .$$

The quantity $b_{ij}(\mathcal{S})$ is called the *success bias on inputs* (i, j). The strategy $\mathcal{S}$ is *perfect* if $p_{ij}(\mathcal{S}) = 1$ for all inputs (i, j), and *ϵ-perfect* if $p_{ij}(\mathcal{S}) \ge 1-\epsilon$ for all inputs (i, j). Note that a strategy is perfect if and only if $\omega(\mathcal{G}; \mathcal{S}) = 1$, but it is not true that a strategy is ϵ-perfect if and only if $\omega(\mathcal{G}; \mathcal{S}) = 1 - \epsilon$. It is not hard to show that the two are related up to a factor depending on the size of the game:

Lemma 4.2. *Let $\mathcal{G}$ be the linear system game associated to the $m \times n$ linear system $Ax = b$. If $\mathcal{S}$ is an ϵ-perfect strategy, then $\omega(\mathcal{G}; \mathcal{S}) \ge 1-\epsilon$. On the other hand, if $\mathcal{S}$ is a strategy such that $\omega(\mathcal{G}; \mathcal{S}) \ge 1 - \epsilon$, then $\mathcal{S}$ is $(nm\epsilon)$-perfect.*

Proof. The game $\mathcal{G}$ has nm pairs of inputs in total, each chosen with probability $1/(nm)$. Therefore success probability $1-\epsilon$ in the game implies that each question must lead to a valid answer with probability at least $1-nm\epsilon$. □

The *solution group* [CLS16] associated to $Ax = b$ is the finitely-presented group $\Gamma(A, b)$ generated by indeterminates $x_1, \ldots, x_n$ and J, and satisfying the relations

(a) $J^2 = e$, and $[x_j, J] = e$ for all $1 \leq j \leq n$,

(b) $x_j^2 = e$ for all $1 \leq j \leq n$,

(c) $\prod_{j \in V_i} x_j = J^{b_i}$ for all $1 \leq i \leq m$, and

(d) $x_j x_{j'} = x_{j'} x_j$ for all $j, j' \in V_i$, $1 \leq i \leq m$.

Let $\mathcal{G}$ be the linear system non-local game associated to the linear system $Ax = b$. Several criteria relating representations of $\Gamma(A, b)$ to the existence of perfect strategies for $\mathcal{G}$ have been given in [CM14, CLS16, Slo17], including:

Proposition 4.3 (Proposition 3.4, [Slo17])**.** *If J is non-trivial in approximate representations of $\Gamma(A, b)$, then $\omega^q(\mathcal{G}) = 1$.*

For this paper, we need a quantitative version of Proposition 4.3, and also its converse. We start by giving the quantitative version, together with a converse for maximally entangled states; although this is not necessary for the proof of Theorem 1.1, it is easier to establish, and gives a tighter connection with hyperlinear profile. The converse for general states will be given in the next section.

Proposition 4.4. *Let $\mathcal{G}$ be the linear system non-local game associated to the $m \times n$ linear system $Ax = b$.*

(a) Given a d-dimensional ϵ-representation ϕ of $\Gamma(A, b)$ with $\phi(J) = -\mathbb{1}$, we can construct an $O(\epsilon^2)$-perfect strategy for $\mathcal{G}$ which uses the maximally entangled state $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$.

(b) If $|\psi\rangle, \{Y_{ij}\}, \{X_j\}$ is an ϵ-perfect strategy for $\mathcal{G}$, where $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ is maximally entangled, then the function

$$\phi : \mathcal{F}(x_1, \ldots, x_n, J) \to \mathcal{U}(\mathbb{C}^d)$$

sending $x_j \mapsto X_j$ and $J \mapsto -\mathbb{1}$ is an $O(\sqrt{\epsilon})$-representation of $\Gamma(A, b)$.

Proof. For part (a), see the proof of Proposition 3.4 in [Slo17] (the key idea is that $\mathbb{Z}_2^k$ is a stable group for any k).

For part (b), observe that if $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ is maximally entangled, we have that

$$\langle\psi| A \otimes B |\psi\rangle = \widetilde{\mathrm{tr}}(A^T B),$$

where the transpose is taken with respect to the Schmidt basis for $|\psi\rangle$. Given an ϵ-perfect strategy where $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ is maximally entangled, let $Z_{ij} = \overline{Y_{ij}}$, the entry-wise complex conjugate of Y_{ij} with respect to the Schmidt basis of $|\psi\rangle$. Then

$$\widetilde{\mathrm{tr}}(Z_{ij}^* X_j) = \widetilde{\mathrm{tr}}(Y_{ij}^T X_j) = \langle\psi| Y_{ij} \otimes X_j |\psi\rangle = b_{ij}(\mathcal{S}) \geq 1 - 2\epsilon.$$

Hence $\|Z_{ij} - X_j\|_f^2 = 2 - 2\widetilde{\mathrm{tr}}(Z_{ij}^* X_j) \leq O(\epsilon)$. For any $1 \leq i \leq m$, we have that

$$\prod_{j \in V_i} Z_{ij} = (-\mathbb{1})^{b_i}, \text{ and } Z_{ij} Z_{ij'} = Z_{ij'} Z_{ij} \text{ for all } j, j' \in V_i.$$

Hence

$$\Big\| \prod_{j \in V_i} X_j - (-\mathbb{1})^{b_i} \Big\|_f \leq O(\sqrt{\epsilon})$$

(where the constant depends on the size of V_i), and

$$\|X_j X_{j'} - X_{j'} X_j\|_f \leq O(\sqrt{\epsilon}) \text{ for all } j, j' \in V_i.$$

Hence the homomorphism ϕ sending $x_j \mapsto X_j$ and $J \mapsto -\mathbb{1}$ is an $O(\sqrt{\epsilon})$-representation of $\Gamma(A, b)$. □

Proposition 4.4 implies that, for strategies using maximally entangled states, the Hilbert space dimension of near-perfect strategies is given by the hyperlinear profile of the generator J in the solution group. To make this precise, we make the following definition.

Definition 4.5. *Given a non-local game $\mathcal{G}$, let $\omega^{me}(\mathcal{G})$ be the supremum of winning probabilities $\omega(\mathcal{G}; \mathcal{S})$ across all finite-dimensional quantum strategies $\mathcal{S}$ using a maximally entangled state. Let $E^{me}(\mathcal{G}, \epsilon)$ be the smallest possible integer d such that there is a quantum strategy $\mathcal{S}$ for $\mathcal{G}$ with*

$$\omega(\mathcal{G}; \mathcal{S}) \geq \omega^{me}(\mathcal{G}) - \epsilon,$$

where $\mathcal{S}$ uses a maximally entangled state in $\mathbb{C}^d \otimes \mathbb{C}^d$.

Lemma 4.2 and Proposition 4.4 imply:

Corollary 4.6. *Let $\mathcal{G}$ be the linear system non-local game associated to a linear system $Ax = b$, and let J be the generator $J \in \Gamma(A, b)$. If $\omega^{me}(\mathcal{G}) = 1$, then there are constants $C \geq C' > 0$ (depending on $\mathcal{G}$) such that*

$$\mathrm{hlp}(J, 2, C\sqrt{\epsilon}) \leq E^{me}(\mathcal{G}, \epsilon) \leq \mathrm{hlp}(J, 2, C'\sqrt{\epsilon}) \ .$$

Note that, by Proposition 2.7, the asymptotics of $\mathrm{hlp}(J, 2, \epsilon)$ are determined by $\mathrm{hlp}(J, \delta, \epsilon)$ for any $\delta \in (0, 2]$.

5. Near-optimal strategies with general states

In this section we prove a version of part (b) of Proposition 4.4 for strategies with general states. Recall that any Hermitian matrix X has a unique polar decomposition $X = U\Sigma$, where Σ is positive-semidefinite and U is a unitary operator on $\operatorname{Im}\Sigma$. We write $|X|$ for the positive-semidefinite part Σ, and $X|X|^{-1}$ for the unitary part U.

Theorem 5.1. *Let $\mathcal{G}$ be the linear system non-local game associated to the $m \times n$ linear system $Ax = b$, and suppose A has no non-zero columns. If $|\psi\rangle, \{Y_{ij}\}, \{X_j\}$ is an ϵ-perfect strategy for $\mathcal{G}$, where $|\psi\rangle \in H_A \otimes H_B$, then there is a projection P on H_B such that*

$$\phi : \mathcal{F}(x_1, \ldots, x_n, J) \to \mathcal{U}(\operatorname{Im} P)$$

sending $x_j \mapsto (PX_jP)|PX_jP|^{-1}$ and $J \mapsto -P$ is an $O(\epsilon^{1/4})$-representation of $\Gamma(A, b)$ on $\operatorname{Im} P$.

If the jth column of A is zero, then the only relations containing x_j in $\Gamma(A, b)$ are $x_j^2 = e$ and $[x_j, J] = e$. So if we set $\phi(x_j)$ to an arbitrary unitary when the jth row of A is zero, then Theorem 5.1 holds for any matrix A. Hence, the following corollary is an immediate consequence of Theorem 5.1 and part (a) of Proposition 4.4:

Corollary 5.2. *Let $\mathcal{G}$ be the linear system non-local game associated to a linear system $Ax = b$, and let J be the generator $J \in \Gamma(A, b)$. If $\omega^q(\mathcal{G}) = 1$, then there are constants $C \geq C' > 0$ (depending on $\mathcal{G}$) such that*

$$\operatorname{hlp}(J, 2, C\epsilon^{1/4}) \leq E(\mathcal{G}, \epsilon) \leq \operatorname{hlp}(J, 2, C'\sqrt{\epsilon}).$$

In particular, if $\omega^q(\mathcal{G}) = 1$ then $E(\mathcal{G}, \epsilon)$ is finite for all $\epsilon > 0$, so we can strengthen Proposition 4.3:

Corollary 5.3. *Let $\mathcal{G}$ be the linear system game associated to a linear system $Ax = b$. Then $\omega^q(\mathcal{G}) = 1$ if and only if J is non-trivial in approximate representations of $\Gamma(A, b)$.*

The proof of Theorem 5.1 has two ingredients. The first is a state-dependent version of part (b) of Proposition 4.4.

Proposition 5.4. *Let $\mathcal{G}$ be the linear system non-local game associated to the $m \times n$ linear system $Ax = b$, and let $V_i = \{j : A_{ij} \neq 0\}$. Suppose that A has no zero columns. If $|\psi\rangle$, $\{Y_{ij}\}$, $\{X_j\}$ is a finite-dimensional ϵ-perfect strategy for $\mathcal{G}$, where $|\psi\rangle \in H_A \otimes H_B$, and ρ is the reduced density matrix of $|\psi\rangle$ on H_B, then*

(a) $\left\|X_j\rho^{1/2} - \rho^{1/2}X_j\right\|_F \leq O(\sqrt{\epsilon})$ for all $1 \leq j \leq n$,

(b) $\left\|\prod_{j\in V_i} X_j - (-\mathbb{1})^{b_i}\right\|_\rho \leq O(\sqrt{\epsilon})$ *for all* $1 \leq i \leq m$*, where the constant depends on the size of* V_i*, and*

(c) $\|X_j X_k - X_k X_j\|_\rho \leq O(\sqrt{\epsilon})$ *for all* $1 \leq i \leq m$*,* $j, k \in V_i$*.*

Proof. Without loss of generality, we can pick a basis $\{|t\rangle\}$ of H_A such that $|\psi\rangle = \sum_t |t\rangle \lambda |t\rangle$, where $\lambda = \rho^{1/2}$. For any matrices B and C,

$$\langle\psi| B^T \otimes C |\psi\rangle = \operatorname{tr}(B\lambda C\lambda),$$

where the transpose is taken with respect to the basis $\{|i\rangle\}$. Suppose that B and C are unitary, and that

$$\langle\psi| B^T \otimes C |\psi\rangle \geq 1 - O(\epsilon).$$

We claim that

$$\|C\lambda - \lambda C\|_F \leq O(\sqrt{\epsilon}), \tag{5.1}$$

and that

$$\|C\lambda - \lambda B^*\|_F \leq O(\sqrt{\epsilon}). \tag{5.2}$$

To prove equation (5.1), observe that $\operatorname{tr}(B\lambda C\lambda)$ can be thought of as the Frobenius inner product of $\lambda^{1/2}B^*\lambda^{1/2}$ and $\lambda^{1/2}C\lambda^{1/2}$. Applying the Cauchy-Schwarz inequality, we get that

$$1 - O(\epsilon) \leq \operatorname{tr}(B\lambda C\lambda) \leq \left\|\lambda^{1/2}B^*\lambda^{1/2}\right\|_F \left\|\lambda^{1/2}C\lambda^{1/2}\right\|_F \leq \left\|\lambda^{1/2}C\lambda^{1/2}\right\|_F,$$

where the last inequality uses the fact that

$$\left\|\lambda^{1/2}B^*\lambda^{1/2}\right\|_F^2 = \operatorname{tr}(B\lambda B^*\lambda) \leq \|B\lambda B^*\|_F \|\lambda\|_F = 1.$$

Since we can assume that $\epsilon \leq 1$ (so that $\epsilon^2 = O(\epsilon)$), we get that

$$\|C\lambda - \lambda C\|_F^2 = 2 - 2\operatorname{tr}(C^*\lambda C\lambda) = 2 - 2\left\|\lambda^{1/2}C\lambda^{1/2}\right\|_F^2 \leq O(\epsilon).$$

Equation (5.2) follows immediately from the fact that λB^* and $C\lambda$ are unit vectors in the Frobenius norm.

Now for part (a), if $j \in V_i$ then by assumption

$$\langle\psi| Y_{ij} \otimes X_j |\psi\rangle \geq 1 - O(\epsilon). \tag{5.3}$$

Since A has no zero columns, every j appears in some V_i, so we conclude from equation (5.1) that

$$\|X_j\lambda - \lambda X_j\|_F \leq O(\sqrt{\epsilon}).$$

For part (b), fix $1 \leq i \leq m$. Using equations (5.3) and (5.2), as well as the unitary invariance of the Frobenius norm, we get that

$$\left\|\prod_{j\in V_i} X_j\lambda - \lambda \prod_{j\in V_i} \overline{Y_{ij}}\right\|_F \leq O(\sqrt{\epsilon}),$$

where the constant depends on the size of V_i. Since $\prod_{j\in V_i}\overline{Y_{ij}} = (-\mathbb{1})^{b_i}$, we get that

$$\Big\| \prod_{j\in V_i} X_j - (-\mathbb{1})^{b_i} \Big\|_\rho = \Big\| \Big(\prod_{j\in V_i} X_j - (-\mathbb{1})^{b_i} \Big)\lambda \Big\|_F \leq O(\sqrt{\epsilon}).$$

The same argument, using the fact that Y_{ij} and Y_{ik} commute, can be used for part (c). □

Let χ_I denote the indicator function of a subset $I \subseteq \mathbb{R}$, and for $a \in \mathbb{R}$ let $\chi_{\geq a} := \chi_{[a,+\infty)}$. Note that if λ is a self-adjoint operator (and I is measurable), then $\chi_I(\lambda)$ is a projection. The second ingredient in the proof of Theorem 5.1 is the following version of Connes' joint distribution trick:

Lemma 5.5 ([Con76], Lemma 1.2.6)**.** *Let λ, λ' be positive semidefinite operators on a finite-dimensional Hilbert space. Then*

$$\int_0^{+\infty} \left\|\chi_{\geq\sqrt{a}}(\lambda) - \chi_{\geq\sqrt{a}}(\lambda')\right\|_F^2 da \leq \|\lambda - \lambda'\|_F \, \|\lambda + \lambda'\|_F \, .$$

Connes' argument applies more generally to any semifinite von Neumann algebra with a normal semifinite faithful trace. For the convenience of the reader, we give a self-contained proof (following the original) for the finite-dimensional case.

Proof. The proof is based on the following trick due to Connes: for any positive semidefinite operators λ, λ' on a finite-dimensional Hilbert space, there is a discrete measure ν on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ such that

$$\|f(\lambda) - g(\lambda')\|_F^2 = \int_{(x,y)} |f(x) - g(y)|^2 \, d\nu \tag{5.4}$$

for any pair of functions $f, g : \mathbb{R}_+ \to \mathbb{R}_+$. Indeed, if we write the spectral decomposition $\lambda = \sum_i \lambda_i \, |u_i\rangle\langle u_i|$, $\lambda' = \sum_j \lambda'_i \, |v_j\rangle\langle v_j|$, and set $\nu = \sum_{i,j} \delta_{(\lambda_i,\mu_j)} |\langle u_i | v_j\rangle|^2$, then Equation (5.4) follows by direct calculation.

Now for non-negative real numbers x, y,

$$\int_0^{+\infty} \left|\chi_{\geq\sqrt{a}}(x) - \chi_{\geq\sqrt{a}}(y)\right|^2 da = |x^2 - y^2| = |x - y||x + y| \, . \tag{5.5}$$

Thus, using (5.4) followed by Fubini's theorem,

$$\begin{aligned}
\int_0^{+\infty} \left\|\chi_{\geq\sqrt{a}}(\lambda) - \chi_{\geq\sqrt{a}}(\lambda')\right\|_F^2 da &= \int_{(x,y)} \int_0^{+\infty} \left|\chi_{\geq\sqrt{a}}(x) - \chi_{\geq\sqrt{a}}(y)\right|^2 da \, d\nu \\
&= \int_{(x,y)} |x - y||x + y| d\nu \\
&\leq \left(\int_{(x,y)} |x - y|^2 d\nu \right)^{1/2} \left(\int_{(x,y)} |x + y|^2 d\nu \right)^{1/2}
\end{aligned}$$

$$= \|\lambda - \lambda'\|_F \|\lambda + \lambda'\|_F \ ,$$

where the last equality again uses (5.4). □

We need one other easy lemma:

Lemma 5.6. *Let λ be a compact positive semidefinite operator on a Hilbert space. Then*

$$\int_0^{+\infty} \chi_{\geq\sqrt{a}}(\lambda) da = \lambda^2.$$

Proof. It suffices to prove the lemma for $\lambda = t\,|v\rangle\,\langle v|$, where $|v\rangle$ is a unit vector, and $t \geq 0$. But then

$$\chi_{\geq\sqrt{a}}(\lambda) = \begin{cases} |v\rangle\,\langle v| & a \leq t^2 \\ 0 & a > t^2 \end{cases},$$

so

$$\int_0^{+\infty} \chi_{\sqrt{a}}(\lambda) da = \int_0^{t^2} |v\rangle\,\langle v|\, da = t^2\,|v\rangle\,\langle v| = \lambda^2.$$

□

Proof of Theorem 5.1. Let $V_i = \{j : A_{ij} \neq 0\}$, and suppose $|\psi\rangle$, $\{Y_{ij}\}$, $\{X_j\}$ is an ϵ-perfect strategy for $\mathcal{G}$. Let ρ be the reduced density matrix of $|\psi\rangle$ on H_B, and let $\lambda = \rho^{1/2}$. By Lemma 5.6, for any Hermitian W on H_B,

$$\int_0^{+\infty} \mathrm{tr}(W\chi_{\geq\sqrt{a}}(\lambda)) da = \mathrm{tr}(W\rho)\ . \tag{5.6}$$

Then, if R is one of the relations $\prod_{j\in V_i} X_j - (-\mathbb{1})^{b_i}$, $1 \leq i \leq m$, or $X_jX_k - X_kX_j$, $j,k \in V_i$,

$$\int_0^{+\infty} \left\|R\,\chi_{\geq\sqrt{a}}(\lambda)\right\|_F^2 da = \|R\|_\rho^2 \leq O(\epsilon), \tag{5.7}$$

where the equality is by (5.6), using that $\chi_{\geq\sqrt{a}}(\lambda)$ is a projection, and the inequality follows from parts (b) and (c) of Proposition 5.4. Finally, Lemma 5.5 and part (a) of Proposition 5.4 imply that

$$\begin{aligned} \int_0^{+\infty} &\left\|\chi_{\geq\sqrt{a}}(\lambda) - X_j^*\chi_{\geq\sqrt{a}}(\lambda)X_j\right\|_F^2 da \\ &\leq \left\|\lambda - X_j^*\chi_{\geq\sqrt{a}}(\lambda)X_j\right\|_F \left\|\lambda + X_j^*\chi_{\geq\sqrt{a}}(\lambda)X_j\right\|_F \leq O(\sqrt{\epsilon}) \end{aligned} \tag{5.8}$$

for all $1 \leq j \leq n$. Putting equations (5.6), (5.7), and (5.8) together, and using the fact that $\epsilon \leq \epsilon^2$ for ϵ small enough,

$$\int_0^{+\infty} \sum_{j=1}^n \left\|\chi_{\geq\sqrt{a}}(\lambda) - X_j^*\chi_{\geq\sqrt{a}}(\lambda)X_j\right\|_F^2 + \sum_i \left\|\Big(\prod_{j\in V_i} X_j - (-\mathbb{1})^{b_i}\Big)\chi_{\geq\sqrt{a}}(\lambda)\right\|_F^2$$

$$+\sum_i \sum_{j\neq k\in V_i} \left\|(X_jX_k - X_kX_j)\,\chi_{\geq\sqrt{a}}(\lambda)\right\|_F^2 \quad da \leq O(\epsilon^{1/2})\int_0^{+\infty} \operatorname{tr}(\chi_{\geq\sqrt{a}}(\lambda))da.$$

All the integrands in the above equation are zero if $a > \|\lambda\|_{op}^2$, so it follows that we can find $0 \leq a_0 \leq \|\lambda\|_{op}^2$ such that if $P := \chi_{\geq\sqrt{a_0}}(\lambda)$, then

$$\sum_{j=1}^n \left\|P - X_j^*PX\right\|_F^2 + \sum_i \left\|\Big(\prod_{j\in V_i} X_j - (-\mathbb{1})^{b_i}\Big)P\right\|_F^2 + \sum_i \sum_{j\neq k\in V_i} \|(X_jX_k - X_kX_j)\,P\|_F^2 \leq O(\epsilon^{1/2})\operatorname{tr}(P). \tag{5.9}$$

Since all summands on the left are positive, we conclude that all summands are bounded by $O(\epsilon^{1/2})\operatorname{tr}(P)$ (where the constant depends only on m and n). In addition, $a \leq \|\lambda\|_{op}^2$ implies that $P \neq 0$. Let $\widetilde{X}_j := (PX_jP)|PX_jP|^{-1}$ for $j = 1,\ldots,n$. We want to show that the homomorphism ϕ from $\mathcal{F}(S)$ to $\mathcal{U}(\operatorname{Im} P)$ sending $x_j \mapsto \widetilde{X}_j$ is an $O(\epsilon^{1/4})$-representation. To start, we need to show that PX_jP is almost unitary, so $\widetilde{X}_j$ is close to PX_jP. Note that if $BB^* \leq \mathbb{1}$ (or equivalently, $B^*B \leq \mathbb{1}$) then $\|AB\|_F \leq \|A\|_F$ and similarly $\|BA\|_F \leq \|A\|_F$ for any matrix A. As a result,

$$\|X_jP - PX_jP\|_F = \|(X_jP - PX_j)P\|_F \leq \|X_jP - PX_j\|_F \leq O(\epsilon^{1/4})\operatorname{tr}(P)^{1/2},$$

and hence

$$\begin{aligned}\left\|(PX_jP)^2 - P\right\|_F &= \|PX_j(PX_jP - X_jP)\|_F \\ &\leq \|PX_jP - X_jP\|_F \\ &\leq O(\epsilon^{1/4})\operatorname{tr}(P)^{1/2}.\end{aligned}$$

By Lemma 2.8, part (a),

$$\left\|\widetilde{X}_j - PX_jP\right\|_F \leq O(\epsilon^{1/4})\operatorname{tr}(P)^{1/2},$$

and we also have

$$\left\|\widetilde{X}_j - X_jP\right\|_F \leq O(\epsilon^{1/4})\operatorname{tr}(P)^{1/2}.$$

To finish the proof, we claim that

$$\left\|X_{i_1}\cdots X_{i_k}P - \widetilde{X}_{i_1}\cdots\widetilde{X}_{i_k}\right\|_F \leq O(\epsilon^{1/4})\operatorname{tr}(P)^{1/2}$$

for any $1 \leq i_1,\ldots,i_k \leq n$, where the constant depends on k. Indeed, for any $1 \leq j \leq n$, if $B = \widetilde{X}_{i_{j+1}}\cdots\widetilde{X}_{i_k}P$ (P is included at the end for the case $j = n$) then $B^*B = P \leq \mathbb{1}$, thus $BB^* \leq \mathbb{1}$ and

$$\begin{aligned}&\left\|X_{i_1}\cdots X_{i_j}\widetilde{X}_{i_{j+1}}\cdots\widetilde{X}_{i_k}P - X_{i_1}\cdots X_{i_{j-1}}\widetilde{X}_{i_j}\cdots\widetilde{X}_{i_k}P\right\|_F \\ &\qquad = \left\|\big(X_{i_j}P - \widetilde{X}_{i_j}\big)\widetilde{X}_{i_{j+1}}\cdots\widetilde{X}_{i_k}P\right\|_F\end{aligned}$$

$$\leq \left\| X_{i_j} P - \widetilde{X}_{i_j} \right\|_F \leq O(\epsilon^{1/4}) \operatorname{tr}(P)^{1/2} ,$$

so the claim follows. Thus if $\phi' : \mathcal{F}(S) \to \mathcal{U}(H_B)$ is defined by sending $x_j \mapsto X_j$ and $J \mapsto -\mathbb{1}$, then

$$\begin{aligned} \|\phi(r) - P\|_f &= \frac{1}{\operatorname{tr}(P)^{1/2}} \|\phi(r) - P\|_F \\ &\leq \frac{1}{\operatorname{tr}(P)^{1/2}} \left(\|\phi(r) - \phi'(r)P\|_F + \|(\phi'(r) - \mathbb{1})P\|_F \right) \\ &\leq O(\epsilon^{1/4}). \end{aligned}$$

for any of the defining relations r of $\Gamma(A, b)$. □

Remark 5.7. *When combined with Proposition 4.4, part (a), Theorem 5.1 implies that any ϵ-perfect strategy can be turned into an $O(\epsilon^{1/2})$-strategy with a maximally entangled state. This can also be proved directly using Lemma 5.5, and when combined with part (b) of Proposition 4.4, this gives another route to prove Theorem 5.1. The Connes joint distribution trick can also be used to prove similar statements about other classes of games, such as synchronous games* [PSS+16] *and weak projection games* [Man14]*. Determining the largest class of games to which this idea can be applied seems to be an interesting question for further work.*

6. Bounds on hyperlinear profile of a specific solution group

To apply Corollary 5.2, we need examples of solution groups with bounds on the hyperlinear profile of J. In this section, we use Proposition 3.2 to prove:

Proposition 6.1. *There is a solution group $\Gamma(A, b)$ and constants $C, C' > 0$ such that J is non-trivial in approximate representations of $\Gamma(A, b)$, and furthermore*

$$\frac{C}{\epsilon^{2/3}} \leq \operatorname{hlp}(J, 2, \epsilon) \leq \frac{C'}{\epsilon}$$

for all $\epsilon > 0$.

To prove Proposition 6.1, recall from [Slo17] that a group is an *extended homogeneous-linear-plus-conjugacy* group if it has a presentation $\langle S : R \rangle$, where

(a) the set S of generators is split into two disjoint subsets S_0 and S_1, where $S_1 = \{y_1, \ldots y_\ell\}$ is linearly-ordered,

(b) R contains the relations $x^2 = e$ for all $x \in S_0$,

(c) there is a collection $\mathcal{V}$ of ordered subsets of S_0 such that for all $V \in \mathcal{V}$, R contains the relations

$$\prod_{x \in V} x = e, \text{ and } xy = yx \text{ for all distinct } x, y \in V$$

(these relations are similar to the relations of a solution group), and

(d) every other relation of R is of the form

(i) $xyx = z$ for some (not necessarily distinct) $x, y, z \in S_0$,

(ii) $xyx^{-1} = z$ for some $x \in S_1$ and (not necessarily distinct) $y, z \in S_0$, or

(iii) $y_i y_j y_i^{-1} = y_j^k$ for some $1 \leq j < i \leq n$ and $k > 0$.

We call a presentation $G = \langle S : R \rangle$ of this form a *presentation of G as an extended homogeneous-linear-plus-conjugacy group*. We can use the main argument of [Slo17] to prove:

Proposition 6.2. *Let $G = \langle S : R \rangle$ be a presentation of G as an extended homogeneous-linear-plus-conjugacy group, and suppose $w \in S_0$ represents a central involution in G which is non-trivial in approximate representations. Then there is a solution group $\Gamma(A, b)$ and constants $C > C' > 0$, $N \in \mathbb{N}$, such that J is non-trivial in approximate representations of $\Gamma(A, b)$, and furthermore*

$$\operatorname{hlp}(w, 2, C\epsilon) \leq \operatorname{hlp}(J, 2, \epsilon) \leq N \operatorname{hlp}(w, 2, C'\epsilon)$$

for all $\epsilon > 0$.

We need one lemma before proving Proposition 6.2.

Lemma 6.3. *If A and B are unitary matrices, then*

$$\|\mathbb{1} - A\|_f \geq 1 - \frac{1}{2} \|\mathbb{1} + [A, B]\|_f .$$

Proof. Observe that $\|\mathbb{1} - [A, B]\|_f \leq 2 \|\mathbb{1} - A\|_f$. Hence

$$2 = \|2\mathbb{1}\|_f \leq \|\mathbb{1} - [A, B]\|_f + \|\mathbb{1} + [A, B]\|_f \leq 2 \|\mathbb{1} - A\|_f + \|\mathbb{1} + [A, B]\|_f .$$

□

Proof of Proposition 6.2. We combine several propositions from [Slo17]. By [Slo17, Proposition 4.8], G embeds in a homogeneous-linear-plus-conjugacy group G', and the embedding can be chosen so that the image of w is non-trivial in approximate representations of G'. We then make G' into a linear-plus-conjugacy group $\widehat{G}$ by adding two new generators t and J, along with relations stating that J is central, $J^2 = t^2 = e$, and $twt = J$. Since $\widehat{G}$ is an HNN extension of $G' \times \mathbb{Z}_2$, G' embeds in $\widehat{G}$, and [Slo17, Lemma 5.2] implies that J

is non-trivial in approximate representations of $\widehat{G}$. Finally, [Slo17, Proposition 4.2] implies that $\widehat{G}$ embeds in a solution group $\Gamma(A,b)$, where $J \in \widehat{G}$ maps to $J \in \Gamma(A,b)$, and J is non-trivial in approximate representations of $\Gamma(A,b)$. Suppose A is an $m \times n$ matrix. We conclude that there is a homomorphism $\Psi : \mathcal{F}(S) \to \mathcal{F}(x_1, \ldots, x_n, J)$ descending to an embedding $G \hookrightarrow \Gamma(A,b)$, such that if z is the image of t in $\mathcal{F}(x_1, \ldots, x_n, J)$, then $[\Psi(w), z] = J$ in $\Gamma(A,b)$.

To show the lower bound on $\operatorname{hlp}(J, 2, \epsilon)$, note that there is a constant $C_0 \geq 1$ such that if ϕ is an ϵ-representation of $\Gamma(A,b)$ with $\phi(J) = -\mathbb{1}$, then

$$\|[\phi(\Psi(w)), \phi(z)] + \mathbb{1}\|_f \leq C_0 \epsilon.$$

By Lemma 6.3, $\|\phi(\Psi(w)) - \mathbb{1}\|_f \geq 1 - O(\epsilon)$. We conclude that

$$\operatorname{hlp}(w, 1 - O(\epsilon), O(\epsilon)) \leq \operatorname{hlp}(\Psi(w), 1 - O(\epsilon), \epsilon) \leq \operatorname{hlp}(J, 2, \epsilon),$$

where the first inequality follows from Lemma 2.4. Finally, since w represents a central involution, we conclude from Proposition 2.7 that $\operatorname{hlp}(w, 2, O(\epsilon)) \leq \operatorname{hlp}(J, 2, \epsilon)$.

The upper bound on $\operatorname{hlp}(J, 2, \epsilon)$ requires careful attention to the construction of Ψ described above. Suppose ϕ is a d-dimensional ϵ-representation of G with $\phi(w) = -\mathbb{1}$. Then the proof of [Slo17, Proposition 4.8] (including the change of presentation) implies that there is a $2d$-dimensional ϵ-representation γ of G' with $\gamma(w) = -\mathbb{1}_d \oplus \mathbb{1}_d$. Setting

$$\gamma(J) = -\mathbb{1} \text{ and } \gamma(t) = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix},$$

we get a $2d$-dimensional ϵ-representation γ of $\widehat{G}$ with $\gamma(J) = -\mathbb{1}$. Finally [Slo17, Remark 4.5] implies that there is an $8d$-dimensional $O(\epsilon)$-representation γ' of $\Gamma(A,b)$ with $\gamma'(J) = -\mathbb{1}$. We conclude that $\operatorname{hlp}(J, 2, O(\epsilon)) \leq 8 \operatorname{hlp}(w, 2, \epsilon)$, so the proposition follows with $N = 8$. □

Proof of Proposition 6.1. Consider the group K defined in Section 3. Since $abc = e$ and a, b, and c are involutions, this relation implies that the elements a, b, and c pairwise commute. Hence K is an extended homogeneous-linear-plus-conjugacy group. Changing the presentation by adding the relations $ab = ba$, $ac = ca$, and $bc = cb$ does not change the asymptotics of the hyperlinear profile of c, so by Proposition 6.2 there is a solution group $\Gamma(A,b)$ with

$$\Omega(1/\epsilon^{2/3}) \leq \operatorname{hlp}(J, 2, \epsilon) \leq O(1/\epsilon).$$

□

Proof of Theorem 1.1. Upper and lower bounds both follow directly from Corollary 5.2 and Proposition 6.1. □

7. A presentation-independent version of hyperlinear profile

A weakness of Definition 2.3 is that it depends on the presentation of G. Following [Cor13], we can define another version of hyperlinear profile which is independent of the presentation, while still being in the spirit of Definition 2.3.

Definition 7.1. *Let E be a finite subset of a group G containing the identity. Let $\eta(E)$ be the function $\mathbb{R}_{>0} \times \mathbb{R}_{>0} \to \mathbb{N} \cup \{+\infty\}$ such that $\eta(E;\delta,\epsilon)$ is the smallest integer d for which there is a function $\phi : E \to \mathcal{U}(\mathbb{C}^d)$ with*

(a) $\phi(e) = \mathbb{1}$,

(b) $\|\phi(x)\phi(y) - \phi(xy)\|_f \leq \epsilon$ whenever x, y, and xy are in E, and

(c) $\|\phi(x) - \phi(y)\|_f \geq \delta$ whenever $x \neq y$ in E,

or $+\infty$ if no such function ϕ exists.

Given a finite subset E of G containing the identity, let $\sigma(E;\delta,\epsilon)$ be defined similarly to $\eta(E;\delta,\epsilon)$, but with $\mathcal{U}(\mathbb{C}^d)$ replaced by the group S_d of $d \times d$ permutation matrices. Then $\sigma(E)$ is a two-parameter version of the sofic profile defined in [Cor13]. Specifically, the *sofic profile* of E is

$$\mathbb{N} \to \mathbb{N} \cup \{+\infty\} : n \mapsto \sigma\left(E; \sqrt{2 - \frac{2}{n}}, \sqrt{\frac{2}{n}}\right). \tag{7.1}$$

More precisely, the sofic profile of E is the equivalence class of this function under the equivalence relation which identifies non-decreasing functions with the same asymptotic growth rate. However, for our purposes we can ignore this distinction. The square roots in Equation (7.1) come from the fact that sofic profile is defined in terms of the normalized Hamming metric on S_n, which for $u, v \in S_n$ is $\|u - v\|_f^2 / 2$.

Let $G = \langle S : R \rangle$ be a finitely-presented group. Then the collection of functions $\{\eta(E)\}$ and $\{\mathrm{hlp}(T)\}$ are roughly equivalent, in the sense that any function in one collection is bounded (asymptotically) by some function in the other collection.

Proposition 7.2. *Let $G = \langle S : R \rangle$ be a finitely-presented group.*

(a) Suppose $T \subset \mathcal{F}(S)$ is a finite subset not containing any element which is trivial in G. Then there is a constant $C > 0$ and a finite subset E of G containing the identity such that for all $\epsilon, \delta > 0$,

$$\mathrm{hlp}(T, \delta - C\epsilon, C\epsilon) \leq \eta(E, \delta, \epsilon).$$

(b) Suppose $E \subset G$ is a finite subset containing the identity. Then there is a constant $C > 0$ and a finite subset $T \subset \mathcal{F}(S)$ not containing any

element which is trivial in G, such that for all $\epsilon, \delta > 0$,

$$\eta(E, \delta, C\epsilon) \leq \mathrm{hlp}(T, \delta, \epsilon).$$

Proof. For (a), suppose $T \subset \mathcal{F}(S)$. Let $E \subset G$ be the set of all elements of the form $x_1^{a_1} x_2^{a_2} \cdots x_i^{a_i}$, where $x_1^{a_1} \cdots x_n^{a_n}$ is an element of $T \cup R \cup S$ for some (not necessarily distinct) elements $x_1, \ldots, x_n \in S$, integers $a_1, \ldots, a_n \in \mathbb{Z}$, and $0 \leq i \leq n$. Note that E contains the identity. Let C be the length of the longest word in $T \cup R$, and suppose $\phi : E \to \mathcal{U}(\mathbb{C}^d)$ has the property that $\|\phi(x)\phi(y) - \phi(xy)\|_f \leq \epsilon$ whenever x, y, and xy are in E. Define $\psi : \mathcal{F}(S) \to \mathcal{U}(\mathbb{C}^d)$ by $\psi(s) = \phi(s)$ for all $s \in S$. By the definition of E, if $w \in T \cup R$ then $\|\psi(w) - \phi(w)\|_f \leq C\epsilon$. Thus $\|\psi(r) - \mathbb{1}\|_f \leq C\epsilon$ for all $r \in R$, and $\|\psi(w) - \mathbb{1}\|_f \geq \|\phi(w) - \mathbb{1}\|_f - C\epsilon$ for all $w \in T$. We conclude that $\mathrm{hlp}(T, \delta - C\epsilon, C\epsilon) \leq \eta(E, \delta, \epsilon)$ for all $\delta, \epsilon > 0$.

For part (b), suppose $E \subset G$ is finite. Choose a representative $w(u) \in \mathcal{F}(S)$ for every $u \in E \setminus \{e\}$, and set $w(e_G) = e \in \mathcal{F}(S)$. Let $T \subset \mathcal{F}(S)$ to be the set of elements of the form $w(u)w(v)^{-1}$, where u and v are distinct elements of E, and choose C such that for any triple u, v, and uv in E, the word $w(u)w(v)$ can be turned into $w(uv)$ by at most C applications of the relations in R. Suppose $\psi : \mathcal{F}(S) \to \mathcal{U}(\mathbb{C}^d)$ is an ϵ-representation of G, and define $\phi : E \to \mathcal{U}(\mathbb{C}^d)$ by $\phi(u) = \psi(w(u))$. Note that $\phi(e) = \mathbb{1}$. Then

$$\|\phi(u)\phi(v) - \phi(uv)\|_f = \|\psi(w(u)w(v)) - \psi(w(uv))\|_f \leq C\epsilon$$

for all triples u, v, and uv in E. If u and v are distinct elements of E, then $\|\phi(u) - \phi(v)\|_f = \|\psi(w(u)w(v)^{-1}) - \mathbb{1}\|_f$, so we conclude that $\eta(E, \delta, C\epsilon) \leq \mathrm{hlp}(T, \delta, \epsilon)$ for all $\delta, \epsilon > 0$. □

We do not know whether there is a finer way to regard the families $\{\eta(E)\}$ and $\{\mathrm{hlp}(T)\}$ as asymptotically equivalent, aside from the rough comparison in Proposition 7.2. This raises the question of whether we should call $\{\eta(E)\}$ the hyperlinear profile of G, and find a different term for $\{\mathrm{hlp}(T)\}$. However, the point of hyperlinear profile (and this is also true of sofic profile) is to see how fast the functions $\eta(E)$ or $\mathrm{hlp}(T)$ can grow. Introducing another term would also raise the possibility of confusion with another related concept, the sofic dimension growth of Arzhantseva and Cherix (see [Cav16]), which is in a somewhat different spirit. Thus we suggest that both families $\{\eta(E)\}$ and $\{\mathrm{hlp}(T)\}$ should be regarded as the hyperlinear profile of G, with a specific definition selected to fit the context.

References

[Cav16] Matteo Cavaleri, *Algorithms and quantifications in amenable and sofic groups*, Ph.D. thesis, Università di Roma La Sapienza, 2016.

[CLS16] Richard Cleve, Li Liu, and William Slofstra, *Perfect commuting-operator strategies for linear system games*, Journal of Mathematical Physics (2016), to appear (arXiv:1606.02278).

[CM14] Richard Cleve and Rajat Mittal, *Characterization of Binary Constraint System Games*, Automata, Languages, and Programming, Lecture Notes in Computer Science, no. 8572, Springer Berlin Heidelberg, 2014, arXiv:1209.2729, pp. 320–331.

[CN16] Matthew Coudron and Anand Natarajan, *The parallel-repeated Magic Square game is rigid*, Tech. report, arXiv:1609.06306, 2016.

[Col16] Andrea W. Coladangelo, *Parallel self-testing of (tilted) EPR pairs via copies of (tilted) CHSH*, Tech. report, arXiv:1609.03687, 2016.

[Con76] A. Connes, *Classification of Injective Factors Cases II1, II, III, 1*, Annals of Mathematics **104** (1976), no. 1, 73–115.

[Cor13] Yves Cornulier, *Sofic profile and computability of Cremona groups*, The Michigan Mathematical Journal **62** (2013), no. 4, 823–841.

[CRSV17] Rui Chao, Ben W. Reichardt, Chris Sutherland, and Thomas Vidick, *Test for a large amount of entanglement, using few measurements*, Proceedings of the 2017 Conference on Innovations in Theoretical Computer Science (ITCS), 2017.

[Fri13] Tobias Fritz, *On infinite-dimensional state spaces*, Journal of Mathematical Physics **54** (2013), no. 5, 052107.

[HS17] Don Hadwin and Tatiana Shulman, *Stability of group relations under small Hilbert-Schmidt perturbations*, arXiv:1706.08405.

[JOP16] Marius Junge, Timur Oikhberg, and Carlos Palazuelos, *Reducing the number of inputs in nonlocal games*, Journal of Mathematical Physics **57** (2016), no. 10, 102203.

[JP11] M. Junge and C. Palazuelos, *Large Violation of Bell Inequalities with Low Entanglement*, Communications in Mathematical Physics **306** (2011), no. 3, 695–746.

[JPPG+10] M. Junge, C. Palazuelos, D. Pérez-García, I. Villanueva, and M. M. Wolf, *Unbounded Violations of Bipartite Bell Inequalities via Operator Space Theory*, Communications in Mathematical Physics **300** (2010), no. 3, 715–739.

[Man14] Laura Mančinska, *Maximally entangled state in pseudo-telepathy games*, Computing with New Resources, Springer, 2014, pp. 200–207.

[NV17] Anand Natarajan and Thomas Vidick, *A quantum linearity test for robustly verifying entanglement*, Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (New York, NY, USA), STOC 2017, ACM, 2017, pp. 1003–1015.

[OV16] Dimiter Ostrev and Thomas Vidick, *Entanglement of approximate quantum strategies in XOR games*, arXiv preprint arXiv:1609.01652 (2016).

[PGWP+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** (2008), no. 2, 455–486.

[PSS+16] Vern I Paulsen, Simone Severini, Daniel Stahlke, Ivan G Todorov, and Andreas Winter, *Estimating quantum chromatic numbers*, Journal of Functional Analysis **270** (2016), no. 6, 2188–2222.

[Slo11] William Slofstra, *Lower bounds on the entanglement needed to play xor nonlocal games*, Journal of Mathematical Physics **52** (2011), no. 10, 102202.

[Slo17] ———, *The set of quantum correlations is not closed*, arXiv:1703.08618.