THE SET OF QUANTUM CORRELATIONS IS NOT CLOSED

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

Received 18 October 2017; accepted 2 November 2018

Abstract
We construct a linear system nonlocal game which can be played perfectly using a limit of finite-dimensional quantum strategies, but which cannot be played perfectly on any finite-dimensional Hilbert space, or even with any tensor-product strategy. In particular, this shows that the set of (tensor-product) quantum correlations is not closed. The constructed nonlocal game provides another counterexample to the `middle' Tsirelson problem, with a shorter proof than our previous paper (though at the loss of the universal embedding theorem). We also show that it is undecidable to determine if a linear system game can be played perfectly with a finite-dimensional strategy, or a limit of finite-dimensional quantum strategies.

2010 Mathematics Subject Classification: 81P45 (primary); 20F05 (secondary)

1. Introduction

A two-player nonlocal game $G$ consists of finite question sets $\mathcal{I}_A$ and $\mathcal{I}_B$, finite output sets $\mathcal{O}_A$ and $\mathcal{O}_B$, and a function $V : \mathcal{O}_A \times \mathcal{O}_B \times \mathcal{I}_A \times \mathcal{I}_B \to \{0, 1\}$. During the game, the two players, commonly called Alice and Bob, are given inputs $x \in \mathcal{I}_A$ and $y \in \mathcal{I}_B$, respectively, and return outputs $a \in \mathcal{O}_A$ and $b \in \mathcal{O}_B$, respectively. The players win if $V(a, b \mid x, y) = 1$, and lose if $V(a, b \mid x, y) = 0$. The players know the rules of the game, and can decide ahead of time on their strategy. However, once the game is in progress, they are unable to communicate, meaning they do not know each other’s inputs or subsequent choices. This can make it impossible for the players to win some games with certainty.
Imagine that the game is played repeatedly. To an outside observer, Alice and Bob’s actions during the game are described by the probability $p(a, b \mid x, y)$ that Alice and Bob output $a \in \mathcal{O}_A$ and $b \in \mathcal{O}_B$ on inputs $x \in \mathcal{I}_A$ and $y \in \mathcal{I}_B$. The collection $\{p(a, b \mid x, y)\} \subset \mathbb{R}^{\mathcal{O}_A \times \mathcal{O}_B \times \mathcal{I}_A \times \mathcal{I}_B}$ is called a correlation matrix (or a behaviour). Which correlation matrices can be achieved depends on the physical model. For instance, a correlation matrix $\{p(a, b \mid x, y)\}$ is said to be classical if it can be achieved using classical shared randomness. Formally, this means that there must be some integer $k \geq 1$, a probability distribution $\{\lambda_i\}$ on $\{1, \ldots, k\}$, probability distributions $\{p_a^i\}$ on $\mathcal{O}_A$ for each $1 \leq i \leq k$ and $x \in \mathcal{I}_A$, and probability distributions $\{q_b^i\}$ on $\mathcal{O}_B$ for each $1 \leq i \leq k$ and $y \in \mathcal{I}_B$, such that

$$p(a, b \mid x, y) = \sum_{i=1}^k \lambda_i p_a^i q_b^i$$

for all $(a, b, x, y) \in \mathcal{O}_A \times \mathcal{O}_B \times \mathcal{I}_A \times \mathcal{I}_B$.

The set of classical correlation matrices is denoted by $\mathcal{C}_c(\mathcal{O}_A, \mathcal{O}_B, \mathcal{I}_A, \mathcal{I}_B)$, although we typically write $\mathcal{C}_c$ when the output and input sets are clear.

In quantum information, we are interested in what correlations can be achieved with a shared quantum state. Accordingly, a correlation matrix is said to be quantum if there are finite-dimensional Hilbert spaces $H_A$ and $H_B$, a quantum state $|\psi\rangle \in H_A \otimes H_B$, projective measurements $\{M^x_a\}_{a \in \mathcal{O}_A}$ on $H_A$ for every $x \in \mathcal{I}_A$, and projective measurements $\{N^y_b\}_{b \in \mathcal{O}_B}$ on $H_B$ for every $y \in \mathcal{I}_B$, such that

$$p(a, b \mid x, y) = \langle \psi | M^x_a \otimes N^y_b |\psi\rangle$$

for all $(a, b, x, y) \in \mathcal{O}_A \times \mathcal{O}_B \times \mathcal{I}_A \times \mathcal{I}_B$.

A projective measurement on a Hilbert space $H$ is a collection $\{P_x\}_{x \in X}$ of self-adjoint operators on $H$, such that $P_x^2 = P_x$ for all $x \in X$, and $\sum_{x \in X} P_x = 1$. The set $X$ is interpreted as the set of measurement outcomes.

The set of quantum correlation matrices is denoted by $\mathcal{C}_q \cong \mathcal{C}_q(\mathcal{O}_A, \mathcal{O}_B, \mathcal{I}_A, \mathcal{I}_B)$. There are two natural variations on this definition. We can drop the requirement that $H_A$ and $H_B$ be finite-dimensional, in which case we get another set of correlations often denoted by $\mathcal{C}_{qs}$. We can also look at correlations which can be realized as limits of finite-dimensional quantum correlations; the corresponding correlation set is the closure of $\mathcal{C}_q$, and is typically denoted by $\mathcal{C}_{qa}$. It is well known that $\mathcal{C}_{qs} \subseteq \mathcal{C}_{qa}$, and consequently $\mathcal{C}_{qa}$ is also the closure of $\mathcal{C}_{qs}$ [29].

Since $\mathcal{C}_{qs} \subseteq \mathcal{C}_{qa}$, we get a hierarchy of correlation sets

$$\mathcal{C}_c \subseteq \mathcal{C}_q \subseteq \mathcal{C}_{qs} \subseteq \mathcal{C}_{qa}.$$ 

All the sets involved are convex, and $\mathcal{C}_c$ and $\mathcal{C}_{qa}$ are both closed. Bell’s celebrated theorem [2] states that $\mathcal{C}_c \not= \mathcal{C}_q$, and furthermore that the two sets are strictly nested.
can be separated by a hyperplane. It has been a longstanding open problem to determine the relationship between the quantum correlation sets, and in particular to determine whether \( C_q \) and \( C_{qs} \) are closed (see [3, 9, 32, 33]). Part of the interest in this latter question comes from the resource theory of nonlocal games: \( C_q \neq C_{qa} \) if and only if there is a nonlocal game which can be played optimally (with respect to some probability distribution on inputs) using a limit of finite-dimensional quantum strategies, but cannot be played optimally using any fixed dimension. Numerical evidence has suggested that even very simple nonlocal games might have this property [18, 25]. For variants of nonlocal games (for instance, with quantum questions, or infinite output sets), there are several examples of games with this property [17, 21, 28].

The purpose of this paper is to show that there are indeed nonlocal games (with finite classical input and output sets) that cannot be played optimally using any fixed dimension. A perfect strategy for a nonlocal game \( \mathcal{G} \) is a correlation matrix \( \{p(a, b | x, y)\} \) such that Alice and Bob win with probability one on every pair of inputs \( x \) and \( y \). Formally, this means that for all \( (a, b, x, y) \in \mathcal{O}_A \times \mathcal{O}_B \times \mathcal{I}_A \times \mathcal{I}_B \), if \( V(a, b | x, y) = 0 \), then \( p(a, b | x, y) = 0 \).

**Theorem 1.** There is a nonlocal game with a perfect strategy in \( C_{qa} \), but no perfect strategy in \( C_{qs} \).

In particular, neither \( C_q \) or \( C_{qs} \) are closed. The proof is constructive, with the game in question having input sets of size 184 and 235, and output sets of size 8 and 2.

The set \( C_q \) is related to the cone of completely positive-semidefinite (cpsd) matrices defined in [16]. An \( n \times n \) matrix \( M \) is said to be cpsd if there are nonnegative operators \( P_1, \ldots, P_n \) on some finite-dimensional Hilbert space with \( M_{ij} = \text{tr}(P_i P_j) \) for all \( 1 \leq i, j \leq n \). By a theorem of Sikora and Varvitsiotis [30], the set \( C_q \) is an affine slice of the cone of cpsd matrices, so the cone of cpsd matrices is not closed as a consequence of Theorem 1.

The fact that \( C_{qs} \neq C_{qa} \) also has an interesting reformulation. Let \( G_i \) be the \( n \)-fold free product \( \mathbb{Z}_m * \cdots * \mathbb{Z}_m \), where \( n = |\mathcal{I}_i| \) and \( m = |\mathcal{O}_i| \), for \( i = A, B \). Let \( M_{ij}^a \) denote the \( i \)th spectral projector of the \( j \)th factor of \( G_A \) in the full group \( C^* \)-algebra \( C^*(G_A) \) of \( G_A \), and define \( M_{ij}^b \) similarly for \( C^*(G_B) \). For each \( i = A, B \), find a faithful representation \( \phi_i \) of \( C^*(G_i) \) on some Hilbert space \( H_i \). The minimal (or spatial) tensor product \( C^*(G_A) \otimes_s C^*(G_B) \) is the norm closure of the image \( \phi_A(C^*(G_A)) \otimes \phi_B(C^*(G_B)) \) in the \( C^* \)-algebra \( \mathcal{B}(H_A \otimes H_B) \). A correlation matrix \( \{p(a, b | x, y)\} \) belongs to \( C_{qa} \) if and only if there is a state \( \omega \) on the \( C^* \)-algebra \( C^*(G_A) \otimes_s C^*(G_B) \) with

\[
p(a, b | x, y) = \omega(M_{a}^x \otimes N_{b}^y)
\]
for all \((a, b, x, y) \in \mathcal{O}_A \times \mathcal{O}_B \times \mathcal{I}_A \times \mathcal{I}_B\) [9, 29]. On the other hand, the correlation matrix belongs to \(C_{qs}\) if and only if there are representations \(\phi_i\) of \(G_i\) on \(H_i\), \(i = A, B\), and a vector state \(|\psi\rangle \in H_A \otimes H_B\), with

\[
p(a, b \mid x, y) = \langle \psi | \phi_A(M^x_a) \otimes \phi_B(N^y_b) | \psi \rangle
\]

for all \((a, b, x, y) \in \mathcal{O}_A \times \mathcal{O}_B \times \mathcal{I}_A \times \mathcal{I}_B\). Since \(C_{qs} \neq C_{qa}\), there can be states on the minimal tensor product \(\mathcal{C}^*(G_A) \otimes_s \mathcal{C}^*(G_B)\) which do not come from vector states on some tensor product \(\phi_A \otimes \phi_B\) of representations \(\phi_A\) and \(\phi_B\).

There is another candidate set of quantum correlations, the commuting-operator correlations \(C_{qc}\), which contains \(C_{qa}\). Determining whether \(C_{qc}\) is equal to \(C_t\) for any \(t \in \{q, qs, qa\}\) is known as Tsirelson’s problem [7, 32]. A theorem of Ozawa [24] (see also [9, 13]) states that \(C_{qa} = C_{qc}\) if and only if there is an affirmative answer to the Connes embedding problem. In a previous paper [31], we showed that \(C_{qs} \neq C_{qc}\). By showing that \(C_{qs} \neq C_{qa}\), we provide another proof of this fact. The proof that \(C_{qs} \neq C_{qc}\) in [31] uses a universal embedding theorem, which states that every finitely presented group embeds in the solution group of a linear system game. In this paper, we follow a similar line, proving a restricted embedding theorem for a subclass of finitely presented groups which we call linear-plus-conjugacy groups. For the proof of this restricted embedding theorem, we use a completely different method from [31], with the result that the proof is much shorter. However, it remains an open problem to prove the universal embedding theorem via the new approach.

If we choose a probability distribution \(\pi\) on questions \(\mathcal{I}_A \times \mathcal{I}_B\), then the probability that the players win when their behaviour is described by correlation matrix \(\{p(a, b \mid x, y)\}\) is

\[
\sum_{x,y,a,b} \pi(x, y)V(a, b \mid x, y)p(a, b \mid x, y).
\]

The quantum value of the game (with probability distribution \(\pi\)) is the optimal winning probability for the game over correlations in \(C_{qa}\). A fundamental question about nonlocal games is whether it is possible to compute the quantum value, either exactly or approximately. As long as \(\pi(x, y) \neq 0\) for every \((x, y) \in \mathcal{I}_A \times \mathcal{I}_B\), then the quantum value of a game is equal to 1 if and only if the game has a perfect strategy in \(C_{qa}\). Thus as a special case of the exact computation problem, we can ask whether it is possible to determine whether a nonlocal game has a perfect quantum strategy. (There seem to be several reasonable ways to formalize the general problem of computing the exact value of a nonlocal game. One way is as the problem of determining, from a given nonlocal game and a rational number \(a\), whether the quantum value of the game is \(\geq a\). Another reasonable formalization is the problem of computing,
from a given nonlocal game and positive integer $n$, the $n$th decimal digit of
the quantum value of the game, where the decimal expansion used must be the
unique expansion which does not terminate in a repeating sequence of 9’s. Both
of these formalizations contain the problem of determining whether a game has
a perfect quantum strategy as a special case.)

Recall that a decision problem of the form ‘Is $P(x)$ true?’ is decidable if
there is a Turing machine which halts and outputs whether or not $P(x)$ is true
on every input $x$. If no such Turing machine exists, then the problem is said
to be undecidable. A set $\mathcal{L}$ is recursive if the decision problem ‘Is $x \in \mathcal{L}$?’ is
decidable. An easy consequence of the universal embedding theorem is that it is
undecidable to determine if a linear system game has a perfect strategy in $C_{qc}$.
In this paper we prove a stronger result by applying our restricted embedding
theorem to Kharlampovich’s example [14] of a finitely presented solvable group
with an undecidable word problem.

**Theorem 2.** There is a (recursive) family of linear system games such that:

(a) it is undecidable to determine if a game in the family has a perfect strategy
   in $C_{qa}$; and

(b) every game in the family has a perfect strategy in $C_{qc}$ if and only if it has a
   perfect strategy in $C_{qa}$.

Theorem 2 can be interpreted as saying that there is no algorithm to compute
the quantum value of a nonlocal game exactly.

A function $f : \mathbb{N} \to Y$ to some set $Y$ is computable if there is a Turing
machine which halts with output $f(n)$ on every input $n$, and computable in
$T(n)$-time if there is such a Turing machine which halts within $T(n)$ steps on
input $n$. The computation time of a Turing machine is often defined using the
size of inputs $n$, rather than the values of $n$. However, to avoid introducing an
extra parameter for the length, or discuss encodings of inputs, we always just
refer to computation time on specific inputs. Kharlampovich’s construction has
been extended by Kharlampovich, Myasnikov, and Sapir to show that the word
problem for finitely presented residually finite groups can be as hard as any
computable function [15]. The word problem for finitely presented residually
finite groups is always decidable, so this is the best possible lower bound. Using
this extension, we can show:

**Theorem 3.** Let $f : \mathbb{N} \to \mathbb{N}$ be a computable function. Then there is a family
of linear system games $\mathcal{G}_n$, $n \in \mathbb{N}$, such that:

(a) the games $\mathcal{G}_n$ have input sets of size $\exp(O(n))$, and the function $n \mapsto \mathcal{G}_n$ is
   computable in $\exp(O(n))$-time;
(b) for any algorithm accepting the language
\[ \{ n \in \mathbb{N} : G_n \text{ has a perfect strategy in } C_q \} , \]
the maximum running time over inputs \( n \leq N \) is at least \( f(N) \) when \( N \) is sufficiently large;
(c) \( G_n \) has a perfect strategy in \( C_{qc} \) if and only if it has a perfect strategy in \( C_q \).

Theorem 3 has the following corollary.

**Corollary 4.** It is undecidable to determine if a linear system game has a perfect strategy in \( C_q \).

Theorems 2 and 3 say nothing about whether it is possible to approximate the quantum value of a nonlocal game. Formally speaking, the approximation problem is to compute a rational number \( r \) from a given game and rational number \( \epsilon > 0 \), such that \( r \) is within \( \epsilon \) of the quantum value of the given game. This problem is complete for the complexity class \( \text{MIP}^* \) of multi-prover interactive proofs with entangled provers [12]. Determining whether there is an algorithm to approximate the quantum value (or equivalently, whether \( \text{MIP}^* \) is computable) is a major open problem in quantum complexity theory. Navascués, Pironio, and Acin have shown that such an algorithm would exist for two-player games if the Connes embedding problem has an affirmative answer [23].

## 2. Group theory preliminaries

### 2.1. Group presentations.

Given a set \( S \), let \( \mathcal{F}(S) \) denote the free group generated by \( S \). If \( H \) is a group, then homomorphisms \( \mathcal{F}(S) \to H \) can be identified with functions \( S \to H \), and we use these two types of objects interchangeably. If \( R \) is a subset of \( \mathcal{F}(S) \), then the quotient of \( \mathcal{F}(S) \) by the normal subgroup generated by \( R \) is denoted by \( \langle S : R \rangle \). If \( G = \langle S : R \rangle \) and \( R' \subset \mathcal{F}(S \cup S') \), then we write \( \langle G, S' : R' \rangle \) to mean \( \langle S \cup S' : R \cup R' \rangle \). If \( G = \langle S : R \rangle \) and \( H = \langle S' : R' \rangle \), then any homomorphism \( \psi : G \to H \) lifts to a homomorphism \( \Psi : \mathcal{F}(S) \to \mathcal{F}(S') \) such that for any \( w \in \mathcal{F}(S) \), the image of \( \Psi(w) \) in \( H \) is equal to \( \psi(w) \). Lifts of this type are not unique unless \( R' \subseteq \{ e \} \).

A group \( G \) is said to be finitely presentable if \( G = \langle S : R \rangle \) for some finite sets \( S \) and \( R \). A finitely presented group is a tuple \( (G, S, R) \), where \( G = \langle S : R \rangle \). In other words, a finitely presented group is a finitely presentable group along with a choice of finite presentation.

For the purposes of this paper, a representation of \( G \) will always mean a unitary representation, that is a homomorphism from \( G \) to the unitary group \( \mathcal{U}(H) \) of
some (possibly infinite-dimensional) Hilbert space $H$. If $G = \langle S : R \rangle$, then a representation is the same thing as a homomorphism $\phi : \mathcal{F}(S) \to \mathcal{U}(H)$ such that $\phi(r) = 1$ for all $r \in R$.

2.2. Approximate representations. Let $\| \cdot \|$ be the normalized Hilbert–Schmidt norm, so if $T$ is an endomorphism of a finite-dimensional Hilbert space $H$, then $\|T\| = \sqrt{\text{tr}(T^*T)}/\sqrt{\dim H}$.

**Definition 5.** Let $G = \langle S : R \rangle$ be a finitely presented group. A finite-dimensional $\epsilon$-approximate representation (or $\epsilon$-representation for short) of $G$ is a homomorphism $\phi : \mathcal{F}(S) \to \mathcal{U}(H)$ from $\mathcal{F}(S)$ to the unitary group $\mathcal{U}(H)$ of some finite-dimensional Hilbert space $H$, such that

$$\|\phi(r) - 1\| \leq \epsilon$$

for all $r \in R$.

The normalized Hilbert–Schmidt norm is invariant under conjugation by unitaries, so the set of $\epsilon$-representations is independent of the cyclic order of the relations $r \in R$. That means that, for instance, we can write the relation $x = y$ without worrying about whether we mean $xy^{-1} = e$ or $y^{-1}x = e$. We use the term approximate representation for any homomorphism $\phi : \mathcal{F}(S) \to \mathcal{U}(H)$, where $H$ is a finite-dimensional Hilbert space. Since the norm is not actually involved, this terminology is somewhat redundant, but it is helpful for situations where we do not want to specify $\epsilon$ or $H$. Note that, in contrast to ordinary representations, we assume that approximate representations are always finite-dimensional (however, both types of representations are always assumed to be unitary). As a result, a 0-representation of $G$ is the same thing as a finite-dimensional representation of $G$.

**Example 6.** The group $\mathbb{Z}_2^k$ has presentation

$$\langle x_1, \ldots, x_k : x_i^2 = e \text{ for all } 1 \leq i \leq k, [x_i, x_j] = e \text{ for all } 1 \leq i \neq j \leq k \rangle,$$

where $[x, y] := xyx^{-1}y^{-1}$. An $\epsilon$-representation of $\mathbb{Z}_2^k$ with this presentation is a homomorphism

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

for some $d$, such that

$$\|\phi(x_i)^2 - 1\| \leq \epsilon \quad \text{for all } 1 \leq i \leq k \quad (1)$$

and

$$\|\phi(x_i)\phi(x_j)\phi(x_i)^{*}\phi(x_j)^{*} - 1\| \leq \epsilon \quad \text{for all } 1 \leq i \neq j \leq k. \quad (2)$$
Since the homomorphism \( \phi \) is determined by its values on \( \{x_1, \ldots, x_k\} \), we can think of an \( \epsilon \)-representation concretely as a tuple of unitaries \( (\phi(x_1), \ldots, \phi(x_k)) \) in \( \mathcal{U}(\mathbb{C}^d) \) satisfying Equations (1) and (2).

There are several different notions of approximate representations in the literature. The notion we are using comes from the study of stable relations of \( C^* \)-algebras (see, for instance, [19, Section 4.1]). For the purposes of this paper, we could also use the closely related notion of approximate homomorphisms as in [4, Section II]. However, Definition 5 is very convenient for working with examples, as we frequently do in this paper. The main disadvantage of this definition is that it depends on the choice of presentation. We can work around this using the following easy lemma.

**Lemma 7.** Let \( \psi : G \to H \) be a homomorphism, where \( G = \langle S : R \rangle \) and \( H = \langle S' : R' \rangle \) are finitely presented groups. If \( \Psi : F(S) \to F(S') \) is a lift of \( \psi \), then there is a constant \( C > 0 \) such that if \( \phi \) is an \( \epsilon \)-representation of \( H \), then \( \phi \circ \Psi \) is a \( C \epsilon \)-representation of \( G \).

**Proof.** Since \( \Psi \) is the lift of a homomorphism, for any \( r \in R \) we have that \( \Psi(r) = a_1 r_1^{b_1} a_1^{-1} \cdots a_k r_k^{b_k} a_k^{-1} \) for some \( k \geq 0 \), \( a_1, \ldots, a_k \in F(S') \), \( r_1, \ldots, r_k \in R' \), and \( b_1, \ldots, b_k \in \{\pm 1\} \). Suppose that \( \phi \) is an \( \epsilon \)-representation of \( H \). The normalized Hilbert–Schmidt norm is invariant under right and left multiplication by unitaries (that is \( \|AU\| = \|UA\| = \|A\| \) for any unitary \( U \)), so \( \|\phi(a_i r_i^{b_i} a_i^{-1}) - 1\| = \|\phi(r_i) - 1\| \leq \epsilon \) for all \( i = 1, \ldots, k \). Hence

\[
\|\phi \circ \Psi(r) - 1\| = \|\phi(a_1 r_1 a_1^{-1} \cdots a_k r_k a_k^{-1}) - 1\|
\leq \|\phi(a_1 r_1 a_1^{-1}) \phi(a_2 \cdots a_k^{-1})\|
+ \|\phi(a_2 r_2 a_2^{-1}) \phi(a_3 \cdots a_k^{-1})\|
+ \cdots + \|\phi(a_k r_k a_k^{-1}) - 1\| \leq k \epsilon,
\]

where the last inequality again uses the invariance of the norm under right multiplication by unitaries. Since \( R \) is finite, we can take \( C \) to be the maximum \( k \) across all \( r \in R \). \( \Box \)

We record two other simple lemmas for later use.

**Lemma 8.** Let \( G = \langle S : R \rangle \), and let \( M \) be the length of the longest relation in \( R \). If \( \phi \) is an \( \epsilon \)-representation of \( G \), and \( \psi \) is an approximate representation of \( G \) with

\[
\|\psi(x) - \phi(x)\| \leq \delta
\]

for all \( x \in S \), then \( \psi \) is an \((M \delta + \epsilon)\)-representation.
Given approximate representations $\phi : \mathcal{F}(S) \rightarrow \mathcal{U}(H)$ and $\psi : \mathcal{F}(S) \rightarrow \mathcal{U}(H')$ of $G = \langle S : R \rangle$, we can form new approximate representations $\phi \oplus \psi : \mathcal{F}(S) \rightarrow \mathcal{U}(H \oplus H')$ and $\phi \otimes \psi : \mathcal{F}(S) \rightarrow \mathcal{U}(H \otimes H')$.

**Lemma 9.** Suppose $\phi$ and $\psi$ are $\epsilon$- and $\epsilon'$-representations of $G$, respectively. Then $\phi \oplus \psi$ is a $\text{max}(\epsilon, \epsilon')$-representation, and $\phi \otimes \psi$ is an $(\epsilon + \epsilon')$-representation.

A group $G$ is said to be residually finite-dimensional if every nontrivial element of $G$ is nontrivial in some finite-dimensional representation. More generally, the set of elements which are trivial in finite-dimensional representations forms a normal subgroup of $G$. We let $G^{\text{fin}}$ denote the quotient of $G$ by this normal subgroup (alternatively, $G^{\text{fin}}$ is the image of $G$ in its profinite completion). Any homomorphism $\phi : G \rightarrow H$ descends to a homomorphism $G^{\text{fin}} \rightarrow H^{\text{fin}}$.

**Definition 10.** A homomorphism $\phi : G \rightarrow H$ is a $\text{fin}$-embedding if the induced map $G^{\text{fin}} \rightarrow H^{\text{fin}}$ is injective, and a $\text{fin}^*$-embedding if $\phi$ is both injective and a $\text{fin}$-embedding.

Equivalently, $\phi$ is a $\text{fin}$-embedding if $\phi(g)$ is nontrivial in finite-dimensional representations whenever $g \in G$ is nontrivial in finite-dimensional representations.

We can similarly look at elements which are nontrivial in approximate representations:

**Definition 11.** Let $G$ be a finitely presentable group. An element $g \in G$ is nontrivial in (finite-dimensional) approximate representations if there is a finite presentation $G = \langle S : R \rangle$, a representative $w \in \mathcal{F}(S)$ for $g$, and some constant $\delta > 0$ such that, for all $\epsilon > 0$, there is an $\epsilon$-representation $\phi$ of $G$ with $\|\phi(w) - \mathbb{1}\| > \delta$.

Alternatively, if $g \in G = \langle S : R \rangle$, let

$$\ell^{fa} \left( g \right) := \lim_{\epsilon \to 0^+} \sup_{\phi} \|\phi(w) - \mathbb{1}\|,$$

where $w$ is a representative for $g$, and the supremum is across $\epsilon$-representations $\phi$ of $G$. It is easy to see that the right-hand side is independent of the choice of representative $w$. By Lemma 7, if $\psi : G \rightarrow H$ is a homomorphism, then $\ell^{fa}(g) \geq \ell^{fa}(\psi(g))$. Consequently, $\ell^{fa}(g)$ is independent of the chosen presentation $\langle S : R \rangle$, and $g$ is nontrivial in approximate representations if and
only if \( \ell^a(g) > 0 \). This makes it apparent that the choice of presentation \( (S : R) \) and representative \( w \) in Definition 11 is arbitrary.

Standard amplification arguments show that the constant \( \delta \) in Definition 11 is also somewhat arbitrary; in particular, any number in \((0, \sqrt{2})\) will work. Since we use these amplification arguments in a later section, we record them in the following lemma. If \( X \) is a linear operator on a finite-dimensional Hilbert space, let \( \widetilde{\text{tr}}(X) := \text{tr}(X)/\dim(H) \).

**Lemma 12.** Let \( G = \langle S : R \rangle \), and suppose that the image of \( w \in \mathcal{F}(S) \) in \( G \) is nontrivial in approximate representations of \( G \). Then for every \( \epsilon, \tau > 0 \) there is an \( \epsilon \)-representation \( \phi \) with \( 0 \leq \widetilde{\text{tr}}(\phi(w)) \leq \tau \).

For the proof, we follow [4, Section II.2.2]:

**Proof.** First, let \( 2 \geq \delta > 0 \) be such that for all \( \epsilon > 0 \), there is an \( \epsilon \)-representation \( \phi \) with \( \|\phi(w) - 1\| \geq \delta \) (since Definition 11 does not depend on the choice of presentation and representative, such a \( \delta \) exists).

Next, suppose \( \phi : \mathcal{F}(S) \to \mathcal{U}(\mathbb{C}^d) \) is an \( \epsilon \)-representation. Let \( \overline{\phi} \) be the approximate representation defined by \( \overline{\phi}(a) = \overline{\phi(a)} \), the entry-wise complex conjugate of \( \phi(a) \) with respect to the standard basis on \( \mathbb{C}^d \). It is not hard to check that \( \overline{\phi} \) is also an \( \epsilon \)-representation. Let \( \gamma \) be the direct sum of \( \phi \oplus \overline{\phi} \) with \( 2d \) copies of the trivial representation of \( \mathcal{F}(S) \). Then \( \gamma \) is an \( \epsilon \)-representation of \( G \) by Lemma 9, \( \text{tr}(\gamma(w)) = \text{tr}(\phi(w)) + \text{tr}(\overline{\phi(w)}) + 2d \geq 0 \), and \( \|\gamma(w) - 1\|^2 = \|\phi(w) - 1\|^2/2 \geq \delta^2/2 \). Since

\[
\|U - 1\|^2 = 2 - 2 \text{Re} \widetilde{\text{tr}}(U) \tag{3}
\]

for any unitary \( U \), we see that

\[
0 \leq \widetilde{\text{tr}}(\gamma(w)) = \text{Re} \widetilde{\text{tr}}(\gamma(w)) \leq 1 - \frac{\delta^2}{4}.
\]

Finally, suppose we are given \( \tau > 0 \). Since \( 2 \geq \delta > 0 \), we can find \( k \) such that \((1 - \delta^2/4)^k < \tau\). By the previous paragraph, for every \( \epsilon > 0 \), there is an \( \epsilon/k \)-representation \( \phi \) such that \( 0 \leq \widetilde{\text{tr}}(\phi(w)) \leq 1 - \delta^2/4 \). Thus, we can use the tensor-power trick: Since \( \widetilde{\text{tr}}(X^\otimes k) = \widetilde{\text{tr}}(X)^k \), Lemma 9 implies that \( \phi^\otimes k \) is an \( \epsilon \)-representation of \( G \) with

\[
0 \leq \widetilde{\text{tr}}(\phi^\otimes k(w)) \leq \left(1 - \frac{\delta^2}{4}\right)^k \leq \tau,
\]
as required. \( \square \)
The set of quantum correlations is not closed

By Equation (3), if \( \widetilde{r}(\phi(w)) \leq \tau \), then \( \| \phi(w) - 1 \| \geq \sqrt{2} - 2\tau \). Thus Lemma 12 immediately implies that we can take any \( \delta \in (0, \sqrt{2}) \) in Definition 11. Equivalently, we can say that \( \ell^{fa}(g) \) never takes values in \((0, \sqrt{2})\).

Let \( \mathbb{R}_{>0} \) be the set of positive reals. Recall from [27] that a group \( G \) is hyperlinear if for every for \( \epsilon > 0, \delta \in (0, \sqrt{2}) \), and finite subset \( F \subseteq G \), there is a function \( f : G \to \mathcal{U}(H) \), where \( H \) is a finite-dimensional Hilbert space, such that:

- \( f(e) = 1 \);
- \( \| f(gh) - f(g)f(h) \| \leq \epsilon \) for all \( g, h \in F \); and
- \( \| f(g) - 1 \| \geq \delta \) for all \( g \in F \setminus \{e\} \).

The proof of Lemma 12 can be used to show:

**Lemma 13.** A finitely presentable group \( G \) is hyperlinear if and only if every element of \( G \setminus \{e\} \) is nontrivial in approximate representations.

**Proof.** Fix a finite presentation \( G = \langle S : R \rangle \). We start with the claim that \( G \) is hyperlinear if and only if for every \( \epsilon > 0, \delta \in (0, \sqrt{2}) \), and finite subset \( F_0 \subset \mathcal{F}(S) \) such that every word in \( F_0 \) has nontrivial image in \( G \), there is an \( \epsilon \)-representation \( \phi \) with \( \| \phi(w) - 1 \| \geq \delta \). To prove one direction of the claim, suppose that \( G \) is hyperlinear, and that we are given \( \epsilon > 0, \delta \in (0, \sqrt{2}) \), and a finite subset \( F_0 \subset \mathcal{F}(S) \) such that every word in \( F_0 \) has nontrivial image in \( G \). Let \( M = \max\{|r| : r \in R \cup F_0 \cup \{2\} \} \), where \( |w| \) denotes the length of \( w \in \mathcal{F}(S) \), and let \( F \) be the image in \( G \) of all words in \( \mathcal{F}(S) \) of length \(< M \). Given a function \( f : G \to \mathcal{U}(H) \), we can define \( \phi : \mathcal{F}(S) \to \mathcal{U}(H) \) by setting \( \phi(s) = f(s) \) for all \( s \in S \). If \( \| f(gh) - f(g)f(h) \| \leq \epsilon \) for all \( g, h \in F \), then

\[
\| \phi(s_1 \cdots s_k) - f(s_1 \cdots s_k) \| = \| f(s_1) \cdots f(s_k) - f(s_1 \cdots s_k) \| \leq (k - 1)\epsilon
\]

for all \( 2 \leq k \leq M \). In particular, if \( f(e) = 1 \) then \( \phi \) will be an \((M - 1)\epsilon\)-representation of \( G \), and

\[
\| \phi(w) - 1 \| \geq \| f(w) - 1 \| - \| \phi(w) - f(w) \| \geq \| f(w) - 1 \| - (M - 1)\epsilon
\]

for all \( w \in F_0 \). It follows from the definition of hyperlinearity that for every \( \epsilon > 0 \) and \( \delta \in (0, \sqrt{2}) \), there is an \( \epsilon \)-representation \( \phi \) with \( \| \phi(w) - 1 \| \geq \delta \) for all \( w \in F_0 \). The other direction of the claim is similar, and we leave it to the reader.
Now suppose that $G$ is hyperlinear, and $g \in G \setminus \{e\}$. Applying the claim with $F_0 = \{w\}$, where $w \in \mathcal{F}(S)$ is some representative of $g$, we see that $g$ is nontrivial in approximate representations of $G$. For the converse, suppose that every element of $G \setminus \{e\}$ is nontrivial in approximate representations, and that $F_0 \subset \mathcal{F}(S)$ is a finite set of words such that every element of $F_0$ is nontrivial in $G$. By Lemma 12, for every $\epsilon > 0$, $\delta \in (0, \sqrt{2})$, and $w \in F$ there is an $\epsilon$-representation $\phi$ with $\|\phi(w) - 1\| \geq \delta$. By taking direct sums of these $\epsilon$-representations, we can conclude that for every $\epsilon > 0$ and $\delta \in (0, \sqrt{2})$, there is an $\epsilon$-representation $\phi$ such that $\|\phi(w) - 1\| \geq \delta/\sqrt{|F_0|}$ for every $w \in F_0$. To improve this lower bound and remove the dependence on $|F_0|$, note that the amplification procedure in the proof of Lemma 12 does not depend on $w$, and thus can be used to improve the lower bound on $\|\phi(w) - 1\|$ for all $w \in F_0$ simultaneously. It follows that for every $\epsilon > 0$ and every $\delta \in (0, \sqrt{2})$, there is an $\epsilon$-representation $\phi$ with $\|\phi(w) - 1\| \geq \delta$ for all $w \in F_0$.

Clearly $\ell^{fa}(g) \geq 0$ for all $g \in G$, and it is easy to see that

$$\ell^{fa}(gh) \leq \ell^{fa}(g) + \ell^{fa}(h)$$

and $\ell^{fa}(hgh^{-1}) = \ell^{fa}(g)$ for all $g, h \in G$. Thus the set of elements of $G$ which are trivial in approximate representations (that is for which $\ell^{fa}(g) = 0$) forms a normal subgroup of $G$. Let $G^{fa}$ be the quotient of $G$ by this normal subgroup. Because $\ell^{fa}$ is decreasing via homomorphisms, any homomorphism $\phi : G \to H$ between finitely presentable groups descends to a homomorphism $G^{fa} \to H^{fa}$.

**Definition 14.** A homomorphism $\phi : G \to H$ is an $fa$-embedding if the induced map $G^{fa} \to H^{fa}$ is injective, and an $fa^*$-embedding if $\phi$ is injective, a $fin$-embedding, and an $fa$-embedding.

Equivalently, $\phi$ is an $fa$-embedding if $\phi(g)$ is nontrivial in approximate representations whenever $g \in G$ is nontrivial in approximate representations.

If $\phi$ and $\psi$ are approximate representations, then we say that $\phi$ is a direct summand of $\psi$ if $\psi = \phi \oplus \phi'$ for some other approximate representation $\phi'$. We use the following simple trick to construct $fa^*$-embeddings.

**Lemma 15.** Let $G = \langle S : R \rangle$ and $H = \langle S' : R' \rangle$ be two finitely presented groups, and let $\Psi : \mathcal{F}(S) \to \mathcal{F}(S')$ be a lift of a homomorphism $\psi : G \to H$.

(a) Suppose that for every representation (resp. finite-dimensional representation) $\phi$ of $G$, there is a representation (resp. finite-dimensional representation) $\gamma$ of $H$ such that $\phi$ is a direct summand of $\gamma \circ \psi$. Then $\psi$ is injective (resp. a $fin$-embedding).
(b) Suppose that there is an integer \( N > 0 \) and a real number \( C > 0 \) such that for every \( d \)-dimensional \( \epsilon \)-representation \( \phi \) of \( G \), where \( \epsilon > 0 \), there is an \( N^d \)-dimensional \( C \epsilon \)-representation \( \gamma \) of \( H \) such that \( \phi \) is a direct summand of \( \gamma \circ \Psi \). Then \( \psi \) is an \( f a \)-embedding.

Proof. Part (a) is clear, so we prove (b). Suppose \( \phi \) is an \( \epsilon \)-representation of \( G \), where \( \epsilon > 0 \). If \( \gamma \circ \Psi = \phi \oplus \phi' \), where \( \phi \) is \( d \)-dimensional and \( \phi' \) is \( (N-1)d \)-dimensional, then

\[
\|\gamma(\Psi(w)) - 1\| = \|\phi(w) \oplus \phi'(w) - 1\| \geq \frac{1}{\sqrt{N}} \|\phi(w) - 1\|
\]

for all \( w \in \mathcal{F}(S) \). So \( \ell^{fa}(\gamma(g)) \geq \ell^{fa}(g)/\sqrt{N} \), and \( \psi \) is an \( f a \)-embedding. \( \square \)

In our applications it will be possible to check parts (a) and (b) of Lemma 15 simultaneously, in which case \( \psi \) will be an \( f a^* \)-embedding.

2.3. Groups over \( \mathbb{Z}_2 \). For convenience, we use the following definition from [31]: A group over \( \mathbb{Z}_2 \) is a pair \((G, J)\), where \( J \) is a central element of \( G \) of order two. Note that \( J \) is allowed to be the identity element. Typically we drop the pair notation, and just use the symbol \( J \) (or \( J_G \) where necessary) to refer to the special element of a group \( G \) over \( \mathbb{Z}_2 \), in the same way that we use \( e \) to refer to the identity element. If \( G \) and \( H \) are groups over \( \mathbb{Z}_2 \), then a morphism \( G \to H \) over \( \mathbb{Z}_2 \) is a group homomorphism \( G \to H \) which sends \( J_G \mapsto J_H \).

If a group \( G \) over \( \mathbb{Z}_2 \) is finitely presentable, then it has a finite presentation \( \langle S : R \rangle \) where \( J \in S \), and \( R \) includes the relations \( J^2 = e \) and \( [J, s] = e \) for every \( s \in S \setminus \{J\} \). We use presentations of this form often enough that it is helpful to have some notation for them. Suppose that \( S_0 \) is a set of indeterminates, and \( R_0 \subseteq \mathcal{F}(S_0 \cup \{J\}) \). Then we set

\[
\langle S_0 : R_0 \rangle_{\mathbb{Z}_2} := \langle S_0 \cup \{J\} : R_0 \cup \{[J, s] = e : s \in S_0\} \cup \{J^2 = e\} \rangle,
\]

and call \( \langle S_0 : R_0 \rangle_{\mathbb{Z}_2} \) a presentation over \( \mathbb{Z}_2 \). As with ordinary presentations, if \( G = \langle S : R \rangle \) or \( \langle S : R \rangle_{\mathbb{Z}_2} \), then \( \langle G, S' : R' \rangle_{\mathbb{Z}_2} := \langle S \cup S' : R \cup R' \rangle_{\mathbb{Z}_2} \).

Example 16. Consider the group \( \mathbb{Z}_2^3 \), presented as in Example 6. Then

\[
\langle \mathbb{Z}_2^{3}, y : x_1x_2x_3y = J \rangle_{\mathbb{Z}_2}
\]

is the finitely presented group

\[
\langle x_1, x_2, x_3, y, J : x_i^2 = [x_i, J] = e \text{ for all } 1 \leq i \leq 3, [x_i, x_j] = e \text{ for all } 1 \leq i \neq j \leq 3, [y, J] = J^2 = e, \text{ and } x_1x_2x_3y = J \rangle.
\]
3. Linear system games and solution groups

Let $Ax = b$ be an $m \times n$ linear system over $\mathbb{Z}_2$. To the system $Ax = b$, we can associate a nonlocal game, called a linear system game, as follows. For each $1 \leq i \leq m$, let $V_i = \{ j : A_{ij} \neq 0 \}$ be the set of indices of variables appearing in the $i$th equation. Let $S_i \subset \mathbb{Z}_2^{V_i}$ be the set of assignments to variables $x_j$, $j \in V_j$ satisfying the $i$th equation, that is $a \in \mathbb{Z}_2^{V_i}$ belongs to $S_i$ if and only if $\sum_{j \in V_i} a_j = b_i$. Then Alice receives an equation as input, represented by an integer $1 \leq i \leq m$, and must output an element $a \in S_i$. Bob receives a variable, represented by an integer $1 \leq j \leq n$, and must output an assignment $b$ for $x_j$. The players win if either $j \not\in V_i$, or $j \in V_i$ and $a_j = b$, that is, Alice’s and Bob’s outputs are consistent.

A quantum strategy (presented in terms of measurements) for a linear system game consists of:

(1) a pair of Hilbert spaces $H_A$ and $H_B$;

(2) a projective measurement $\{N^i_b\}_{b \in \mathbb{Z}_2}$ on $H_B$ for every integer $1 \leq j \leq n$;

(3) a projective measurement $\{M^i_a\}_{a \in S_i}$ on $H_A$ for every integer $1 \leq i \leq m$; and

(4) a quantum state $|\psi\rangle \in H_A \otimes H_B$.

The strategy is finite-dimensional if $H_A$ and $H_B$ are finite-dimensional. The associated quantum correlation matrix $\{p(a, b | i, j)\}$ is defined by

$$p(a, b | i, j) = \langle \psi | M^i_a \otimes N^j_b | \psi \rangle, \quad 1 \leq i \leq m, \ 1 \leq j \leq n, \ a \in S_i, \ b \in \mathbb{Z}_2.$$ 

As in the introduction, we also use the term strategy to refer to the correlation matrix $\{p(a, b | i, j)\}$. If $j \in V_i$, then the probability that Alice and Bob win on inputs $i$ and $j$ is

$$p_{ij} := \sum_{a, b : a_j = b} p(a, b | i, j).$$

A strategy is perfect if and only if $p_{ij} = 1$ for all $1 \leq i \leq m$ and $j \in V_i$.

For linear system games, it is often convenient to work with strategies presented in terms of $\pm 1$-valued observables—self-adjoint operators which square to the identity—rather than measurement operators. A quantum strategy (presented in terms of observables) consists of

(a) a pair of Hilbert spaces $H_A$ and $H_B$;

(b) a collection of self-adjoint operators $X_j$, $1 \leq j \leq n$, on $H_B$ such that $X_j^2 = 1$ for every $1 \leq j \leq n$;

W. Slofstra
14
The set of quantum correlations is not closed

(c) a collection of self-adjoint operators $Y_{ij}$, $1 \leq i \leq m$, $j \in V_i$ on $H_A$ such that:

(i) $Y_{ij}^2 = 1$ for every $1 \leq i \leq m$ and $j \in V_i$,

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

(iii) $Y_{ij} Y_{il} = Y_{il} Y_{ij}$ for every $1 \leq i \leq m$ and $j, l \in V_i$; and

(d) a quantum state $|\psi\rangle \in H_A \otimes H_B$.

Given a quantum strategy presented in terms of measurements, we can get a quantum strategy presented in terms of observables by setting $X_j = N_j^0 - N_j^1$ for every $1 \leq j \leq n$, and

$$Y_{ij} = \sum_{a \in S_i} (-1)^{a_j} M_{a}^i$$

for $1 \leq i \leq m$ and $j \in V_i$. Conversely, given a quantum strategy in terms of observables, we can recover the measurement presentation using the spectral decomposition of the observables. So the two notions of strategy are equivalent. Note that if $j \in V_i$, then

$$\langle \psi | Y_{ij} \otimes X_j | \psi \rangle = \sum_{a \in S_i} (-1)^{a_j} M_{a}^i \otimes \sum_{b \in \mathbb{Z}_2} (-1)^{b} N_{b}^j | \psi \rangle$$

$$= \sum_{a \in S_i, b \in \mathbb{Z}_2} (-1)^{a_j+b} p(a, b \mid i, j)$$

$$= 2 \left[ \sum_{a, b : a_j = b} p(a, b \mid i, j) \right] - 1 = 2p_{ij} - 1,$$  (4)

where $p_{ij}$ is, again, the probability that Alice and Bob win on inputs $i$ and $j$. The quantity $2p_{ij} - 1$ is called the winning bias on inputs $i$ and $j$.

To every linear system, we can also associate a finitely presented group over $\mathbb{Z}_2$, as follows.

**Definition 17.** Let $Ax = b$ be an $m \times n$ linear system. The solution group of this system is the group

$$\Gamma(A, b) := \langle x_1, \ldots, x_n : x_j^2 = e \text{ for all } 1 \leq j \leq n, \prod_{j=1}^n x_j^{A_{ij}} = J^{b_i} \text{ for all } 1 \leq i \leq m, \text{ and } x_j x_k = x_k x_j \text{ if } j, k \in V_i \text{ for some } 1 \leq i \leq m \rangle_{\mathbb{Z}_2}.$$
We say that a group over \( \mathbb{Z}_2 \) is a solution group if it has a presentation over \( \mathbb{Z}_2 \) of this form.

Solution groups and linear system games are related as follows.

**Theorem 18** ([6], see also [5]). Let \( G \) be the linear system game associated to a system \( Ax = b \). Then the following are equivalent:

(a) \( G \) has a perfect strategy in \( C_{qs} \).

(b) \( G \) has a perfect strategy in \( C_q \).

(c) \( J_\Gamma \) is nontrivial in some finite-dimensional representation of \( \Gamma = \Gamma(A, b) \).

Although we have not defined the set of commuting-operator correlations \( C_{qc} \), we can work with \( C_{qc} \) through the following result.

**Theorem 19** [5]. The linear system game associated to a system \( Ax = b \) has a perfect strategy in \( C_{qc} \) if and only if \( J_\Gamma \) is nontrivial in \( \Gamma = \Gamma(A, b) \).

The main point of this section is to prove an analogue of one direction of Theorem 18 for approximate representations.

**Proposition 20.** Let \( \Gamma = \Gamma(A, b) \) be a solution group. If \( J_\Gamma \) is nontrivial in finite-dimensional approximate representations of \( \Gamma \) then the linear system game associated to \( Ax = b \) has a perfect strategy in \( C_{qa} \).

The proof of Proposition 20 is a straightforward application of a number of easy stability lemmas. We start by pinning down what we want to prove.

**Lemma 21.** The linear system game associated to \( Ax = b \) has a perfect strategy in \( C_{qa} \) if and only if, for all \( \epsilon > 0 \), there is a finite-dimensional quantum strategy (presented in terms of observables) \( \{Y_{ij}, X_j, |\psi\rangle\} \) such that

\[
\langle \psi | Y_{ij} \otimes X_j |\psi\rangle \geq 1 - \epsilon \quad \text{for all } 1 \leq i \leq n, \ j \in V_i.
\]

**Proof.** Since \( C_{qa} \) is the closure of \( C_q \), the linear system game associated to \( Ax = b \) has a perfect strategy in \( C_{qa} \) if and only if, for every \( \epsilon > 0 \), there is a finite-dimensional quantum strategy such that the winning probability \( p_{ij} \geq 1 - \epsilon/2 \) for every \( 1 \leq i \leq m \) and \( j \in V_i \). But \( p_{ij} \geq 1 - \epsilon/2 \) if and only if the winning bias \( 2p_{ij} - 1 \geq 1 - \epsilon \), so the lemma follows from Equation (4).
Next, we come to the stability lemmas, which will allow us to turn approximate representations of the solution group \( \Gamma \) into quantum strategies. The following lemmas are all likely well known to experts (see, for instance, [8, 10]); we include the proofs for completeness.

**Lemma 22.** For any diagonal matrix \( X \), there is a diagonal matrix \( D \) with \( D^2 = 1 \) and
\[
\|D - X\| \leq \left(1 + \frac{1}{\sqrt{2}}\right)\|X^2 - 1\|.
\]

**Proof.** Suppose \( X \) is a \( d \times d \) matrix, and let \( D_{ii} = \text{sgn} \, \text{Re} \, X_{ii} \) for all \( 1 \leq i \leq d \), where \( \text{sgn} \, x = 1 \) if \( x \geq 0 \) and \(-1\) if \( x < 0 \). To show that the desired inequality holds, consider a complex number \( \alpha = a + bi \). Then
\[
|\alpha^2 - 1|^2 = |a^2 - b^2 - 1 + 2abi|^2 = [(a^2 - 1) - b^2]^2 + 4a^2b^2 = (a^2 - 1)^2 + 2b^2 + 2a^2b^2 + b^4.
\]
In particular, this implies that \( |\alpha^2 - 1|^2 \) is greater than or equal to \( (a^2 - 1)^2 \) and \( 2b^2 \). Consequently,
\[
\|\text{Re} \, X^2 - 1\| = \sqrt{\frac{1}{d} \sum_j \left[\text{Re} \, X_{jj}^2 - 1\right]^2} \leq \sqrt{\frac{1}{d} \sum_j |X_{jj}^2 - 1|^2} = \|X^2 - 1\|,
\]
and
\[
\|\text{Re} \, X - X\| = \|\text{Im} \, X\| = \sqrt{\frac{1}{d} \sum_j |\text{Im} \, X_{jj}|^2} \leq \sqrt{\frac{1}{2d} \sum_j |X_{jj}^2 - 1|^2} = \frac{1}{\sqrt{2}} \|X^2 - 1\|.
\]
By considering the cases \( a \geq 0 \) and \( a < 0 \) separately, we see that
\[
|a^2 - 1| = |1 + a||1 - a| = (1 + |a|)|\text{sgn} \, a - a| \geq |\text{sgn} \, a - a|
\]
for all \( a \in \mathbb{R} \). Thus, as above, \( \|D - \text{Re} \, X\| \leq \|\text{Re} \, X^2 - 1\| \), and the lemma follows. \( \square \)

**Lemma 23.** Suppose \( X_1, \ldots, X_n \) are commuting unitary matrices, with \( X_i^2 = 1 \) for all \( 1 \leq i \leq n \), and \( Y \) is a unitary matrix such that \( Y^2 = 1 \) and \( Y \) commutes
with $X_i$ for all $1 \leq i \leq n - 1$. Then there is a unitary matrix $Z$ such that $Z^2 = 1$, $Z$ commutes with $X_i$ for all $1 \leq i \leq n$, and

$$\|Z - Y\| \leq \left(1 + \frac{1}{2\sqrt{2}}\right) \|X_nY - YX_n\|.$$ 

Proof. Let $Z_0 = \frac{1}{2}(Y + X_nYX_n)$. Clearly $Z_0$ commutes with $X_i$ for all $1 \leq i \leq n - 1$. Since $X_n^2 = 1$, we also have that $X_nZ_0 = \frac{1}{2}(X_nY + YX_n) = Z_0X_n$. Since $Y^2 = 1 = (X_nYX_n)^2$ as well, we have that

$$\|Z_0^2 - 1\| = \frac{1}{4} \|YX_nYX_n + X_nYX_nY - 21\|$$

$$\leq \frac{1}{4} \|YX_nYX_n - 1\| + \frac{1}{4} \|X_nYX_nY - 1\| = \frac{1}{2} \|X_nY - YX_n\|.$$ 

Since $X_n$ and $Y$ are self-adjoint, $Z_0$ is self-adjoint, so we can simultaneously diagonalize $X_1, \ldots, X_n$ and $Z_0$. Hence by Lemma 22, there is a matrix $Z$ such that $Z^2 = 1$, $Z$ commutes with $X_i$ for all $1 \leq i \leq n$, and

$$\|Z - Z_0\| \leq \left(1 + \frac{1}{\sqrt{2}}\right) \|Z_0^2 - 1\| \leq \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right) \|X_nY - YX_n\|.$$ 

Finally,

$$\|Y - Z_0\| = \frac{1}{2} \|Y - X_nYX_n\| = \frac{1}{2} \|X_nY - YX_n\|,$$

so the lemma follows. \qed

Lemma 24. Consider $\mathbb{Z}_2^k$ as a finitely presented group with presentation

$$(x_1, \ldots, x_k : x_i^2 = e, [x_i, x_j] = e \text{ for all } i \neq j).$$

Then there is a constant $C > 0$, depending on $k$, such that if $\phi$ is an $\epsilon$-representation of $\mathbb{Z}_2^k$ on a Hilbert space $H$, then there is a representation $\psi$ of $\mathbb{Z}_2^k$ on $H$ with

$$\|\psi(x_i) - \phi(x_i)\| \leq C\epsilon$$

for all $1 \leq i \leq k$.

Proof. Suppose $\psi$ is an $\epsilon$-representation of $\mathbb{Z}_2^k$ such that the following properties hold for some $1 \leq l \leq k - 1$:

(a) $\psi(x_i)^2 = 1$ for all $1 \leq i \leq k$; and

(b) $\psi(x_i)$ commutes with $\psi(x_j)$ for all $1 \leq i \leq l - 1$ and $1 \leq j \leq k$. 

Downloaded from https://www.cambridge.org/core. IP address: 54.70.40.11, on 17 Aug 2019 at 11:07:03, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/fmp.2018.3
In particular, property (b) requires that $\psi(x_1), \ldots, \psi(x_l)$ pairwise commute. Then by Lemma 23, for each $l < j \leq k$ there is a unitary matrix $X_j$ such that $X_j^2 = 1$, $X_j$ commutes with $\psi(x_i)$ for all $1 \leq i \leq l$, and

$$\|X_j - \psi(x_j)\| \leq C_0 \|\psi(x_i)\psi(x_j) - \psi(x_j)\psi(x_i)\| \leq C_0 \epsilon,$$

where $C_0 = 1 + 1/(2\sqrt{2})$. Define an approximate representation $\psi'$ of $\mathbb{Z}_2^k$ by $\psi'(x_i) = \psi(x_i)$ if $i \leq l$ and $\psi'(x_i) = X_i$ if $i > l$. Then $\psi'(x_i)^2 = 1$ for all $1 \leq i \leq k$, and $\psi'(x_i)$ commutes with $\psi'(x_j)$ for all $1 \leq i \leq l$ and $1 \leq j \leq k$. In other words, $\psi'$ satisfies properties (a) and (b) with $l$ replaced by $l + 1$. Finally, $\|\psi'(x_i) - \psi(x_i)\| \leq C_0 \epsilon$ for all $1 \leq i \leq k$, so $\psi'$ is a $(4C_0 + 1)\epsilon$-representation by Lemma 8.

Now suppose that $\phi$ is any $\epsilon$-representation of $\mathbb{Z}_2^k$. By Lemma 22, there is an approximate representation $\psi_1$ of $\mathbb{Z}_2^k$ with $\psi_1(x_i)^2 = 1$ and $\|\psi_1(x_i) - \phi(x_i)\| \leq C_1 \epsilon$ for all $1 \leq i \leq k$, where $C_1 = (1 + 1/(\sqrt{2}))$. By Lemma 8, $\psi_1$ is a $(4C_1 + 1)\epsilon$-representation. Clearly, $\psi_1$ satisfies conditions (a) and (b) with $l = 1$. Using the argument in the previous paragraph, we can then iteratively define approximate representations $\psi_2, \ldots, \psi_{k-1}$, where $\psi_j$ satisfies conditions (a) and (b) with $l = j$ for all $1 \leq j \leq k - 1$. Let $\epsilon_l = (4C_0 + 1)^{l-1}(4C_1 + 1)\epsilon$, so $\psi_1$ is an $\epsilon_1$-representation. It is not hard to check that $\psi_l$ is an $\epsilon_l$-representation, and furthermore that

$$\|\psi_l(x_i) - \psi_1(x_i)\| \leq \frac{1}{4} \left( (4C_0 + 1)^{l-1} - 1 \right) \epsilon_1 = \frac{1}{4} \left( (4C_0 + 1)^{l-1} - 1 \right) (4C_1 + 1) \epsilon$$

for all $1 \leq i \leq k$. Since $\psi_{k-1}$ is an exact representation, we can take

$$C = \frac{1}{4} \left( (4C_0 + 1)^{k-2} - 1 \right) (4C_1 + 1) + C_1.$$

**Lemma 25.** Suppose $G = \langle S_0 : R_0 \rangle_{\mathbb{Z}_2^k}$, where $R_0$ includes the relations $s^2 = e$ for all $s \in S_0$. If $J_G$ is nontrivial in finite-dimensional approximate representations of $G$, then for every $\epsilon > 0$ there is an $\epsilon$-representation $\phi$ of $G$ such that $\phi(J) = -1$, and $\phi(s)^2 = 1$ for all $s \in S_0$.

**Proof.** Suppose $A$ is an $m \times n$ matrix, and let $S = S_0 \cup \{J\}$. If $J$ is nontrivial in approximate representations, then there is a $\delta > 0$ such that for all $\epsilon > 0$, there is an $\epsilon$-representation $\phi$ with $\|\phi(J) - 1\| > \delta$.

By Lemmas 8, 22, and 23, there are constants $C, C' > 0$ such that if $\phi$ is an $\epsilon$-representation, then there is a $C'\epsilon$-representation $\psi$ such that:

1. $\psi(x)^2 = 1$ for all $x \in S$;
2. $\psi(s)$ and $\psi(J)$ commute for all $s \in S_0$; and
3. $\|\psi(J) - \phi(J)\| \leq C \epsilon$.
We can take $C = (1 + 1/\sqrt{2})$, while $C'$ will depend on the length of the longest defining relation of $G$. If $\|\phi(J) - \mathbb{1}\| > \delta$, and $\epsilon < \delta/(2C)$, then

$$\delta < \|\phi(J) - \mathbb{1}\| \leq \|\phi(J) - \psi(J)\| + \|\psi(J) - \mathbb{1}\| \leq \frac{\delta}{2} + \|\psi(J) - \mathbb{1}\|,$$

so $\|\psi(J) - \mathbb{1}\| \geq \delta/2$. Thus we conclude that for all $\epsilon > 0$, there is an $\epsilon$-representation $\psi$ satisfying conditions (1) and (2), and with $\|\psi(J) - \mathbb{1}\| > \delta/2$.

Suppose $\psi$ is an $\epsilon$-representation satisfying conditions (1) and (2), and with $\|\psi(J) - \mathbb{1}\| > \delta/2$. Choose a basis with $\psi(J) = \mathbb{1}_{d_0} \oplus (-\mathbb{1}_{d_1})$. Since $\psi(s)$ commutes with $\psi(J)$ for all $s \in S_0$, we must have $\psi = \psi_0 \oplus \psi_1$, where $\psi_a$ is an approximate representation of dimension $d_a$, and $\psi_a(J) = (-\mathbb{1})^a$, $a = 0, 1$. Since $\psi(s)^2 = \mathbb{1}$, we also have $\psi_a(s)^2 = \mathbb{1}$ for all $s \in S_0$, $a = 0, 1$. To finish the proof, we just need to show that $\psi_1$ is a $C''\epsilon$-representation for some constant $C''$ independent of $\psi$. If $w \in \mathcal{F}(S)$, then

$$\|\psi(w) - \mathbb{1}\|^2 = \frac{d_0}{d_0 + d_1} \|\psi_0(w) - \mathbb{1}\|^2 + \frac{d_1}{d_0 + d_1} \|\psi_1(w) - \mathbb{1}\|^2.$$

If $w = J$, then $\|\psi_0(w) - \mathbb{1}\| = 0$ and $\|\psi_1(w) - \mathbb{1}\| = \|\mathbb{1} - 2\mathbb{1}\| = 4$, so we conclude that

$$\frac{\delta^2}{4} < \|\psi(J) - \mathbb{1}\|^2 = \frac{4d_1}{d_0 + d_1},$$

so $d_1/(d_0 + d_1) > \delta^2/16$. On the other hand, if $w = r$ is one of the defining relations of $G$, then

$$\epsilon^2 \geq \|\psi(r) - \mathbb{1}\|^2 \geq \frac{d_1}{d_0 + d_1} \|\psi_1(r) - \mathbb{1}\|^2 > \frac{\delta^2}{16} \|\psi_1(r) - \mathbb{1}\|^2.$$

Thus $\psi_1$ is a $4\epsilon/\delta$-representation with $\psi_1(J) = -\mathbb{1}$ and $\psi_1(s)^2 = \mathbb{1}$ for all $s \in S_0$. Since $\delta$ is a constant, the lemma follows. 

**Proof of Proposition 20.** For this proof, we use the notation $O(\epsilon)$ to hide constants which are independent of $\epsilon$, $\phi$, and so on. The constants can still depend on the linear system $Ax = b$, however. Suppose $J$ is nontrivial in finite-dimensional approximate representations of $\Gamma$. Given $\epsilon > 0$, let $\phi$ be an $\epsilon$-representation of $\Gamma$ with $\phi(J) = -\mathbb{1}$ and $\phi(x_j)^2 = \mathbb{1}$ for all $1 \leq j \leq n$, as in Lemma 25. Suppose $\phi$ has dimension $d$, and let $|v\rangle$ be the maximally entangled state on $\mathbb{C}^d \otimes \mathbb{C}^d$. For each $1 \leq j \leq n$, set $X_j = \phi(x_j)$. For each $1 \leq i \leq m$, let $j_i$ be the maximal element of $V_i$, and set $W_i := V_i \setminus \{j_i\}$. The restriction of $\phi$ to the subgroup $\langle x_j : j \in W_i \rangle$ is an $\epsilon$-representation of $\mathbb{Z}_2^{|W_i|}$, and by Lemma 24, there is a representation $\psi_i$ of $\mathbb{Z}_2^{|W_i|}$ with $\|\psi_i(x_j) - \phi(x_j)\| \leq O(\epsilon)$. Set $Y_{ij} := \psi_i(x_j)^T$.
(the transpose of $\psi_i(x_j)$ in a Schmidt basis for $|v\rangle$) for all $j \in W_i$, and set $Y_{ij} := (-1)^{bh} \prod_{j \in W_i} Y_{ij}$.

Suppose $j \in W_i$ for some $1 \leq i \leq m$. Since $Y_{ij}$ and $X_j$ are self-adjoint, we have that

$$2 - \frac{2}{d} \text{tr}(Y_{ij}^T X_j) = \|Y_{ij}^T - X_j\|^2 = \|\psi(x_j) - \phi(x_j)\|^2 \leq O(\epsilon^2),$$

so $(1/d) \text{tr}(Y_{ij}^T X_j) \geq 1 - O(\epsilon^2)$. For the remaining variable in $V_i$, we have that

$$\|Y_{ij}^T - X_j\| = \left\|(-1)^{bh} \prod_{j \in W_i} \psi_i(x_j) - \phi(x_j)\right\|$$

$$\leq \left\|(-1)^{bh} \prod_{j \in W_i} \phi(x_j) - \phi(x_j)\right\| + |W_i|\epsilon$$

$$= \left\|(-1)^{bh} \prod_{j \in V_i} \phi(x_j) - 1\right\| + |W_i|\epsilon \leq O(\epsilon),$$

where the last equality uses the fact that $\phi(x_j)^2 = 1$. Because the $Y_{ij}$'s commute for all $j \in W_i$, $Y_{ij}$ is also self-adjoint, so once again we conclude that

$$2 - \frac{2}{d} \text{tr}(Y_{ij}^T X_j) = \|Y_{ij}^T - X_j\|^2 \leq O(\epsilon^2)$$

or in other words that $(1/d) \text{tr}(Y_{ij}^T X_j) \geq 1 - O(\epsilon^2)$.

Now clearly $\{Y_{ij}\}, \{X_j\}, |v\rangle$ is a strategy for the linear system game associated to $Ax = b$. If $A$ and $B$ are any two $d \times d$ matrices, it follows from the definition of maximally entangled states that

$$\langle v| A \otimes B |v\rangle = \frac{1}{d} \text{tr}(A^T B).$$

We conclude that $\langle v| Y_{ij} \otimes X_j |v\rangle = (1/d) \text{tr}(Y_{ij}^T X_j) \geq 1 - O(\epsilon^2)$ for all $j \in V_i$, $1 \leq i \leq m$. The proposition follows from Lemma 21. \qed

4. Linear-plus-conjugacy groups

The goal of the next two sections is to show that there is a solution group $\Gamma$ such that $J_\Gamma$ is trivial in finite-dimensional representations, but nontrivial in approximate representations. In this section, we start by showing that it suffices to construct more general types of groups with these properties. Specifically,
in Section 4.1 we show that a certain class of groups, which we call linear-plus-conjugacy groups, \( f a^* \)-embeds over \( \mathbb{Z}_2 \) in solution groups. In Section 4.2 we show that if we drop the requirement that the embeddings be over \( \mathbb{Z}_2 \), then a more general class of groups, which we call extended homogeneous-linear-plus-conjugacy groups, can also be \( f a^* \)-embedded in solution groups.

### 4.1. Embeddings over \( \mathbb{Z}_2 \)

Given an \( m \times n \) linear system \( Ax = b \), we once again let \( V_i = V_i(A) := \{1 \leq j \leq n : A_{ij} \neq 0\} \).

**Definition 26.** Suppose \( Ax = b \) is an \( m \times n \) linear system over \( \mathbb{Z}_2 \), and \( C \subseteq [n] \times [n] \times [n] \), where \( [n] = \{1, \ldots, n\} \). Let

\[
\Gamma(A, b, C) := \langle \Gamma(A, b) : x_i x_j x_i = x_k \text{ for all } (i, j, k) \in C \rangle_{\mathbb{Z}_2}.
\]

Lacking a better term, we say that a group over \( \mathbb{Z}_2 \) is a **linear-plus-conjugacy group** if it has a presentation over \( \mathbb{Z}_2 \) of this form.

The conjugacy part of the name comes from the fact that since \( x_i \) is an involution, the relation \( x_i x_j x_i = x_k \) is equivalent to the relation \( x_i x_j x_i^{-1} = x_k \), so \( \Gamma(A, b, C) \) can be thought of as a solution group with additional conjugacy relations. In the context of linear-plus-conjugacy and related groups, we use the term **conjugacy relations** as a convenient shorthand for relations of the form \( x y x = z \). We also use the term **linear relation** \( x_1 \cdots x_n = e \) to refer to the set of relations

\[
\{x_1 \cdots x_n = e\} \cup \{[x_i, x_j] = e : 1 \leq i \neq j \leq n\}.
\]

Finally, observe that there are two ways to make generators \( x_i \) and \( x_j \) commute in a linear-plus-conjugacy group: we can add a conjugacy relation \( x_i x_j x_i = x_j \), or add an additional generator \( x_{n+1} \) and a linear relation \( x_i x_j x_{n+1} = e \). We pick and choose from these two methods based on what is convenient.

The main point of this section is to prove:

**Proposition 27.** Let \( G \) be a linear-plus-conjugacy group. Then there is an \( f a^* \)-embedding \( G \to \Gamma \) over \( \mathbb{Z}_2 \), where \( \Gamma \) is a solution group.

We prove Proposition 27 by first showing that linear-plus-conjugacy groups can be embedded in linear-plus-conjugacy groups of a certain form.

**Definition 28.** A linear-plus-conjugacy group is **nice** if it has a presentation of the form \( \Gamma(A, b, C) \), where \( A \) is an \( m \times n \) matrix over \( \mathbb{Z}_2 \), \( b \in \mathbb{Z}_2^m \), and \( C \subseteq [n] \times [n] \times [n] \) is such that if \( (i, j, k) \in C \), then \( j, k \in V_l \) for some \( 1 \leq l \leq m \).
This means that if \( x_i x_j x_i = x_k \) is a defining relation of a nice linear-plus-conjugacy group, then \( x_j x_k = x_k x_j \) will also be a defining relation.

**Lemma 29.** Let \( G \) be a linear-plus-conjugacy group. Then there is an \( f \mathcal{a}^* \)-embedding \( G \to K \) over \( \mathbb{Z}_2 \), where \( K \) is a nice linear-plus-conjugacy group.

**Proof.** Suppose \( G = \Gamma(A, b, C) \), where \( A \) is an \( m \times n \) matrix. Let

\[
K := \langle \Gamma(A, b), w_j, y_j, z_j \mid 1 \leq j \leq n \text{ and } f : \\
f^2 = e, \ y_j^2 = z_j^2 = w_j^2 = e \text{ for all } 1 \leq j \leq n, \\
x_j = y_j z_j = f w_j \text{ and } f y_j f = z_j \text{ for all } 1 \leq j \leq n, \\
y_j z_k = z_k y_j \text{ for all } (i, j, k) \in C, \text{ and} \\
w_i y_j w_i = z_k \text{ for all } (i, j, k) \in C \mathbb{Z}_2.
\]

Since the generators are involutions, note that the relations imply that \( f w_j = w_j f \), \( y_j z_j = z_j y_j \), and \( f z_j f = y_j \) for all \( 1 \leq j \leq n \). If \( (i, j, k) \in C \), then

\[
w_i z_j w_i = w_i f y_j f w_i = f w_i y_j w_i f = f z_k f = y_k,
\]

so

\[
x_i x_j x_i = f w_i y_j z_j f w_i = (f w_i y_j w_i f)(f w_i z_j w_i f) \\
= (f z_k f)(f y_k f) = y_k z_k = x_k
\]

in \( K \). Thus there is a homomorphism \( \psi : G \to K \) sending \( x_i \mapsto x_i \).

Suppose \( \phi \) is an \( \epsilon \)-representation of \( G \), where \( \epsilon > 0 \). Define an approximate representation \( \gamma \) of \( K \) by

\[
\gamma(x_i) = \begin{pmatrix} \phi(x_i) & 0 \\ 0 & \phi(x_i) \end{pmatrix}, \quad \gamma(J) = \begin{pmatrix} \phi(J) & 0 \\ 0 & \phi(J) \end{pmatrix}, \\
\gamma(y_i) = \begin{pmatrix} \phi(x_i) & 0 \\ 0 & 1 \end{pmatrix}, \quad \gamma(z_i) = \begin{pmatrix} 1 & 0 \\ 0 & \phi(x_i) \end{pmatrix}, \\
\gamma(w_i) = \begin{pmatrix} 0 & \phi(x_i) \\ \phi(x_i) & 0 \end{pmatrix}, \quad \text{and} \quad \gamma(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
\]

It is straightforward to check that \( \gamma \) is an \( \epsilon \)-representation of \( K \). If \( \Psi \) is the lift of \( \psi \) sending \( x_i \mapsto x_i \), then \( \gamma \circ \Psi = \phi \oplus \phi \). When \( \phi \) is an exact representation of dimension \( d \) (possibly infinite), the same construction gives an exact representation \( \gamma \) of dimension \( 2d \). By Lemma 15, \( \psi \) is an \( f \mathcal{a}^* \)-embedding.

Finally, we observe that \( K \) is a nice linear-plus-conjugacy group. Indeed, since the relation \( x_i = y_i z_i \) forces \( y_i \) and \( z_i \) to commute, this relation is equivalent to
the relations
\[ x_i y_i z_i = e = [x_i, y_i] = [x_i, z_i] = [y_i, z_i],\]
which means that we can make \( x_i = y_i z_i \), and similarly \( x_i = f w_i \), part of the ‘linear’ relations. By adding ancilla variables \( g_{jk} \), the commuting relations \( y_j z_k = z_k y_j \) can also be replaced with equivalent linear relations \( g_{jk} y_j z_k = e \). The conjugacy relations \( f y_j f = z_j \) and \( w_i y_w_i = z_k \) will then satisfy the requirements of Definition 28.

**Proof of Proposition 27.** By Lemma 29, we can assume that \( G \) is a nice linear-plus-conjugacy group. Let \( G = \Gamma(A, b, C) \) be a presentation satisfying the conditions of Definition 28. Augment the linear system \( Ax = b \) by adding additional variables \( y_{Ij} \) for each \( I \in C \) and \( 1 \leq j \leq 7 \), and additional relations
\[
\begin{align*}
  x_i + y_{I1} + y_{I2} &= 0, & x_j + y_{I2} + y_{I3} &= 0, & y_{I3} + y_{I4} + y_{I5} &= 0, \\
  x_i + y_{I5} + y_{I6} &= 0, & x_k + y_{I6} + y_{I7} &= 0, & y_{I1} + y_{I4} + y_{I7} &= 0
\end{align*}
\]
for every \( I = (i, j, k) \in C \). Let \( \Gamma \) be solution group of this augmented linear system, so
\[
\Gamma = \langle \Gamma(A, b), y_{Ij} \mid I \in C, 1 \leq j \leq 7 : R \rangle_{\mathbb{Z}_2},
\]
where \( R \) consists of the new relations (now written in multiplicative form)
\[
x_i y_{I1} y_{I2} = x_j y_{I2} y_{I3} = y_{I3} y_{I4} y_{I5} = x_i y_{I5} y_{I6} = x_k y_{I6} y_{I7} = y_{I1} y_{I4} y_{I7} = e \quad (1)
\]
for every \( I = (i, j, k) \in C \), as well as the corresponding commutation relations. In \( \Gamma \), we have that
\[
x_i x_j x_i = (y_{I1} y_{I2}) (y_{I2} y_{I3}) (y_{I5} y_{I6}) = y_{I1} (y_{I3} y_{I5}) y_{I6} = y_{I1} y_{I4} y_{I6} = y_{I7} y_{I6} = x_k
\]
for every \( I = (i, j, k) \in C \). So once again we get a homomorphism \( \psi : G \to \Gamma \) sending \( x_i \mapsto x_i \).

Suppose \( \phi \) is an \( \epsilon \)-representation of \( G \). Define an approximate representation \( \gamma \) of \( \Gamma \) by
\[
\begin{align*}
  \gamma(x_i) &= \left( \begin{array}{cc} \phi(x_i) & 0 \\
 0 & \phi(x_i) \end{array} \right), & \gamma(y_{I1}) &= \left( \begin{array}{cc} 0 & \phi(x_i) \\
 \phi(x_i) & 0 \end{array} \right), \\
  \gamma(y_{I2}) &= \left( \begin{array}{cc} 0 & 1 \\
 1 & 0 \end{array} \right), & \gamma(y_{I3}) &= \left( \begin{array}{cc} 0 & \phi(x_j) \\
 \phi(x_j) & 0 \end{array} \right), \\
  \gamma(y_{I4}) &= \left( \begin{array}{cc} 0 & \phi(x_j x_i) \\
 \phi(x_j x_i) & 0 \end{array} \right), & \gamma(y_{I5}) &= \left( \begin{array}{cc} \phi(x_j x_i) & 0 \\
 0 & \phi(x_i) \end{array} \right), \\
  \gamma(y_{I6}) &= \left( \begin{array}{cc} \phi(x_j x_k) & 0 \\
 0 & 1 \end{array} \right), & \text{and} & \gamma(y_{I7}) &= \left( \begin{array}{cc} \phi(x_j) & 0 \\
 0 & \phi(x_k) \end{array} \right)
\end{align*}
\]
for all $I = (i, j, k) \in C$. It is straightforward to show that $\gamma$ is a $C\epsilon$-representation of $\Gamma'$, where $C$ is a positive constant $\leq 15$. For instance, consider the relation $y_{15}^2 = e$. To show that $\gamma(y_{15})^2 \approx 1$, we need to show that $\phi(x_i x_j x_k)^2 \approx 1$. Write $X \approx \epsilon Y$ to mean that $\|X - Y\| \leq \epsilon$. Since $\phi(x^2) \approx 1$ and $\phi(x^3) \approx 1$, we have $\phi(x_i x_j x_k)^2 \approx 3\epsilon 1$. We can conclude from this that $\gamma(y_{15})^2 \approx 3\epsilon 1$ (we can do slightly better by averaging over the blocks of $\gamma(y_{15})$, but we ignore this to simplify the analysis). We can similarly show that $\gamma(y_{1j})^2 \approx 3\epsilon 1$ for all $1 \leq j \leq 7$, and that the linear relations in Equation (1) hold to within $3\epsilon$.

This leaves the commuting relations. Consider the relation $y_{13} y_{14} y_{15} = e$. We want to show that $\gamma(y_{13})$, $\gamma(y_{14})$, and $\gamma(y_{15})$ approximately commute. But since $\gamma(y_{13}) \gamma(y_{14}) \gamma(y_{15}) \approx 3\epsilon 1$ and $\gamma(y_{1j})^2 \approx 3\epsilon 1$, we conclude that

$$\gamma(y_{14}) \gamma(y_{15}) \approx 3\epsilon \gamma(y_{13})^* \approx 3\epsilon \gamma(y_{13}) \approx 3\epsilon \gamma(y_{15})^* \approx 6\epsilon \gamma(y_{15}) \gamma(y_{14}),$$

or in other words, $\gamma(y_{14}) \gamma(y_{15}) \approx 15\epsilon \gamma(y_{15}) \gamma(y_{14})$. The other commuting relations follow similarly.

Let $\Psi$ be the lift of $\psi$ sending $x_i \mapsto x_i$. Then $\gamma \circ \Psi = \phi \oplus \phi$. Once again, the same construction applies when $\psi$ is an exact representation, so $\psi$ is an $f a^*$-embedding by Lemma 15.

Note that if $j = k$ in a relation $x_i x_j x_i = x_j$, then the system in Equation (1) is precisely the Mermin–Peres magic square [22, 26]. The magic square has previously been used by Ji to show that linear system games can require a (finite but) arbitrarily high amount of entanglement to play perfectly [11].

The proof of Proposition 27 has several interesting features:

**Remark 30.** Let $G = \Gamma(A, b, C)$ be an $m \times n$ linear-plus-conjugacy group, and let $\Gamma' = \Gamma'(A', b')$ be the solution group constructed in the proof of Proposition 27. Then, accounting for Lemma 29, the system $A' x = b'$ has $11n + 8c + 1$ variables and $8n + m + 7c$ equations, where $c = |C|$ is the number of conjugacy relations. A presentation for $\Gamma'$ can be constructed in polynomial time in $m, n$, and $c$.

The proofs of Lemma 29 and Proposition 27 show that there is a constant $C > 0$, and a lift $\Psi$ of the homomorphism $G \to \Gamma'$ to the defining free groups, such that for any $d$-dimensional $\epsilon$-representation $\phi$ of $G$, there is a $4d$-dimensional $C\epsilon$-representation $\psi$ of $\Gamma'$ with $\psi \circ \Psi = \phi \oplus^4$. Taking into account the fact that we have to change the presentation of the group $K$ in the proof of Lemma 29, we can take the constant $C \leq 75$. The lift $\Psi$ can be chosen to send the generators of $G$ to generators of $\Gamma'$ (although not every generator of $\Gamma'$ will lie in the image of $\Psi$).
4.2. **Embeddings not over** $\mathbb{Z}_2$. It is important for our argument that the $fa^*$-embedding in Proposition 27 is over $\mathbb{Z}_2$. However, we can go a little further in what type of groups can be embedded if we drop this requirement.

**Definition 31.** Suppose $A$ is an $m \times n$ matrix over $\mathbb{Z}_2$, and $C \subseteq [n] \times [n] \times [n]$. Let

$$\Gamma_0(A, C) := \left\langle x_1, \ldots, x_n : x_j^2 = e \text{ for all } 1 \leq j \leq n, \prod_{j=1}^{n} x_j^{A_{ij}} = e \text{ for all } 1 \leq i \leq m, \ x_j x_k = x_k x_j \text{ if } j, k \in V_i(A) \text{ for some } 1 \leq i \leq m, \text{ and } x_i x_j x_i = x_k \text{ for all } (i, j, k) \in C \right\rangle.$$  

We say that a group $G$ is a **homogeneous-linear-plus-conjugacy group** if it has a presentation of this form.

Since $\Gamma_0(A, C)$ is not presented over $\mathbb{Z}_2$, a homogeneous-linear-plus-conjugacy group is *not* a linear-plus-conjugacy group. However, the two types of groups are closely related, as $\Gamma_0(A, C) \times \mathbb{Z}_2 = \Gamma(A, 0, C)$.

**Definition 32.** Suppose $A$ is an $m \times n$ matrix over $\mathbb{Z}_2$, $C_0 \subseteq [n] \times [n] \times [n]$, $C_1 \subseteq [\ell] \times [n] \times [n]$, and $L$ is an $\ell \times \ell$ lower-triangular matrix with nonnegative integer entries. Let

$$EG_0(A, C_0, C_1, L) := \left\langle \Gamma_0(A, C_0), \ y_1, \ldots, y_\ell : y_i x_j y_i^{-1} = x_k \text{ for all } (i, j, k) \in C_1, \text{ and } y_i y_j y_i^{-1} = y_j^{L_{ij}} \text{ for all } i > j \text{ with } L_{ij} > 0 \right\rangle.$$  

We refer to the generators $x_i$ in this presentation as **involutary generators**, and to the generators $y_j$ as **noninvolutary generators**. We say that a group $G$ is an **extended homogeneous-linear-plus-conjugacy group** if it has a presentation of this form.

**Proposition 33.** Let $G = EG_0(A, C_0, C_1, L)$ as in Definition 32, where $A$ is an $m \times n$ matrix. Then there is an $m \times n'$ matrix $A'$ and a set $C' \subseteq [n'] \times [n'] \times [n']$, where $n \leq n'$, such that there is an $fa^*$-embedding $\psi : G \to \Gamma_0(A', C')$ with $\psi(x_i) = x_i$ for all $1 \leq i \leq n$.

**Proof.** Suppose $G$ has $\ell$ noninvolutary generators, and let

$$G' = \langle G, z, w : z^2 = w^2 = e, y_1 = zw, z y_i = y_i z \text{ for } i = 2, \ldots, \ell \rangle.$$
We claim that the natural morphism $\psi : G \to G'$ is an $f a^*$-embedding. Indeed, let $\Psi : \mathcal{F}(S) \to \mathcal{F}(S \cup \{z, w\})$ be the natural inclusion, where $S = \{x_1, \ldots, x_n, y_1, \ldots, y_\ell\}$. Given an $\epsilon$-representation $\phi$ of $G$, define an approximate representation $\gamma$ of $G'$ by

$$
\gamma(x_i) = \begin{pmatrix} \phi(x_i) & 0 \\ 0 & 1 \end{pmatrix}, \quad \gamma(z) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},
$$

$$
\gamma(w) = \begin{pmatrix} 0 & \phi(y_1)^* \\ \phi(y_1) & 0 \end{pmatrix}, \quad \gamma(y_1) = \begin{pmatrix} \phi(y_1) & 0 \\ 0 & \phi(y_1)^* \end{pmatrix}, \quad \text{and}
$$

$$
\gamma(y_i) = \begin{pmatrix} \phi(y_i) & 0 \\ 0 & \phi(y_i) \end{pmatrix} \quad \text{for } i = 2, \ldots, \ell.
$$

Because $L$ is lower-triangular, $G'$ has no defining relations of the form $y_1 y_i y_1^{-1} = y_i^{L_{i1}}$. Suppose $L_{i1} > 0$, so that $\phi(y_i) \phi(y_1) \phi(y_i)^* \approx_\epsilon \phi(y_1)^{L_{i1}}$, where once again $X \approx_\epsilon Y$ means that $\|X - Y\| \leq \epsilon$. Then $\phi(y_i) \phi(y_1)^* \phi(y_i)^* \approx_\epsilon \phi(y_1)^{-L_{i1}}$, so $\psi(y_i) \psi(y_1) \psi(y_i)^* \approx_\epsilon \psi(y_1)^{L_{i1}}$. It is easy to see that the remaining defining relations of $G'$ hold to within $\epsilon$, so $\psi$ is an $\epsilon$-representation of $G'$. Since $\phi$ is a direct summand of $\gamma \circ \Psi$, we can apply Lemma 15 with $N = 2$ and $C = 1$ to see that $\psi$ is an $f a$-embedding. The same construction for exact representations shows that $\psi$ is an $f a^*$-embedding.

Next, observe that $G'$ is an extended homogeneous-linear-plus-conjugacy group with $\ell - 1$ noninvolutory generators. Indeed, suppose $(1, j, k) \in C_1$. Then the defining relation $y_1 x_j y_1^{-1} = x_k$ is equivalent to the relation $zw x_j wz = x_k$. By adding an ancilla variable $Z_{jk}$ with $Z_{jk}^2 = e$, we can replace this relation with the two conjugacy relations $wx_j w = Z_{jk}$ and $zZ_{jk} z = x_k$. Similarly, suppose $L_{i1} > 0$. Then the relation $y_i y_1 y_i = y_i^{L_{i1}}$ is equivalent to the relation $y_i w y_i^{-1} = w(zw)^{L_{i1}}$. Once again, we can replace this relation with a sequence of conjugacy relations by adding ancilla variables. For instance, if $L_{i1} = 3$, then we would add ancilla variables $W_{i0}$ and $W_{i1}$ with $W_{i0}^2 = W_{i1}^2 = e$, and conjugacy relations $zwz = W_{i0}$, $wW_{i0} w = W_{i1}$, and $y_i w y_i^{-1} = W_{i1}$. After making these replacements, the only relation containing $y_1$ is $y_1 = zw$, so we can remove $y_1$ from the set of generators. The commuting relations added in $G'$ are equivalent to $y_i z y_i^{-1} = z$ for all $2 \leq i \leq \ell$, so $G'$ is an extended homogeneous-linear-plus-conjugacy group. The additional variables (including the ancilla) are involutory generators, so $G'$ has $\ell - 1$ noninvolutory generators.

Iterating this construction, we get a sequence of $f a^*$-embeddings terminating in a homogeneous-linear-plus-conjugacy group, as desired.

The reason the above argument does not apply for groups over $\mathbb{Z}_2$ is that, if we set $\gamma(J) = \phi(J) \oplus 1$, then $\gamma(J)$ would not commute with $\gamma(z)$ and $\gamma(w)$, while
if we set $\gamma(J) = \phi(J) \oplus \phi(J)$, then any linear relations containing $J$ would not be satisfied.

**Remark 34.** The above proof shows that, in Proposition 33, we can take

$$n' = n + 2\ell + \binom{\ell}{2} + |C_1| + \text{sum}(L)$$

and

$$|C'| = |C_0| + 2|C_1| + 2\binom{\ell}{2} + \text{sum}(L) + \#(L),$$

where $\ell$ is the number of noninvolutory generators, $\text{sum}(L)$ is the sum of the entries of $L$, and $\#(L)$ is the number of nonzero entries of $L$. The matrix $A'$ and set $C'$ can be constructed in polynomial time in $m, n, \ell, |C_0|, |C_1|, \text{and sum}(L)$.

Since a homogeneous-linear-plus-conjugacy group can be turned into a linear-plus-conjugacy group by taking a product with $\mathbb{Z}_2$, Propositions 27 and 33 imply that every extended homogeneous-linear-plus-conjugacy group can be $f a^*$-embedded in a solution group. When using these embedding theorems in the next two sections, we start with an extended homogeneous-linear-plus-conjugacy group of interest, and embed it in a homogeneous-linear-plus-conjugacy group using Proposition 33. We then turn this homogeneous-linear-plus-conjugacy group into a linear-plus-conjugacy group by taking the product with $\mathbb{Z}_2$ and adding relations involving $J$. Finally, we embed this linear-plus-conjugacy group into a solution group over $\mathbb{Z}_2$ using Proposition 27.

### 5. Proof of Theorem 1

The point of this section is to prove the following proposition, and hence finish the proof of Theorem 1.

**Proposition 35.** There is a solution group $\Gamma$ for which $J$ is trivial in finite-dimensional representations, but nontrivial in finite-dimensional approximate representations.

For the proof of Proposition 35, it is convenient to work with sofic groups. We do not need to know the definition of soficity, just that the class of sofic groups has the following properties:

1. Amenable groups are sofic.

2. Sofic groups are hyperlinear.
(3) If $H$ is an amenable subgroup of a sofic group $G$, and $\alpha : H \to G$ is injective homomorphism, then the HNN extension

$$\langle G, t : t h t^{-1} = \alpha(h) \rangle$$

(where $t$ is a new indeterminate)

of $G$ by $\alpha$ is sofic.

An expository treatment of sofic groups can be found in [4]. In particular, the last ‘closure property’ can be found in [4, Section II.4].

We need one more general-purpose lemma before proceeding to the proof.

**Lemma 36.** Suppose $G = \langle S : R \rangle$ is a finitely presented group, where $R$ contains the relation $a^2 = e$ for some $a \in S$. Let

$$\hat{G} := \langle G, t : t^2 = e, tat = Ja \rangle_{\mathbb{Z}_2},$$

where $J, t \notin S$. If $a$ is nontrivial in approximate representations of $G$, then $J$ is nontrivial in approximate representations of $\hat{G}$.

Note that $\hat{G}$ is the ‘$\mathbb{Z}_2$-HNN extension’ of $G \times \mathbb{Z}_2$, where $J$ is the generator of the $\mathbb{Z}_2$ factor, by the order-two automorphism sending $a \mapsto Ja$ and $J \mapsto J$.

**Proof.** Recall that if $X$ is a linear operator on a finite-dimensional Hilbert space $H$, then $\tilde{\text{tr}}(X) := \text{tr}(X)/\dim H$. Suppose $\phi$ is an $\epsilon$-representation of $G$ with $\phi(a)^2 = 1$ and $\text{tr}(\phi(a)) \geq 0$ (we explain how to fulfill these hypotheses shortly). Because the eigenvalues of $\phi(a)$ belong to $\{\pm 1\}$, we can choose a basis so that $\phi(a) = \mathbb{1}_{d_0} \oplus (-\mathbb{1})_{d_0} \oplus \mathbb{1}_{d_1}$, where $d_1 = \text{tr}(\phi(a))$. Define an approximate representation $\psi$ of $\hat{G}$ by

$$\psi(x) = \phi(x) \quad \text{for all } x \in S, \quad \psi(J) = -\mathbb{1}, \quad \text{and} \quad \psi(t) = \begin{pmatrix} 0 & \mathbb{1}_{d_0} & 0 \\ \mathbb{1}_{d_0} & 0 & 0 \\ 0 & 0 & \mathbb{1}_{d_1} \end{pmatrix}. $$

Clearly $\|\psi(r) - \mathbb{1}\| = \|\phi(r) - \mathbb{1}\| \leq \epsilon$ for all relations $r \in R$, $\psi([J, s]) = \mathbb{1}$ for all $s \in S \cup \{t\}$, and $\psi(t)^2 = \psi(J)^2 = \mathbb{1}$. For the remaining relation,

$$\|\psi(tat) - \psi(Ja)\| = \|0_{2d_0} \oplus 2\mathbb{1}_{d_1}\| = 2\sqrt{\frac{d_1}{2d_0 + d_1}} = 2\sqrt{\tilde{\text{tr}}(\phi(a))}. $$

So $\psi$ will be a max$(\epsilon, 2\sqrt{\tilde{\text{tr}}(\phi(a))})$-representation with $\|\psi(J) - \mathbb{1}\| = 2$.

Suppose $a$ is nontrivial in approximate representations of $G$, and fix $\tau > 0$. By Lemma 12, for every $\epsilon > 0$ there is an $\epsilon$-representations $\gamma$ of $G$ with
0 ≤ \tilde{\tau}(\gamma(a)) ≤ \tau. We want to show that it is possible to find ε-representations γ of this form with γ(a)^2 = 1. In the proof of Lemma 12, the approximate representations γ can be constructed from any family of ε-representations φ_ε, ε > 0, such that \|φ_ε(a) - 1\| ≥ δ for all ε > 0, where δ > 0 is some fixed constant. By Lemmas 8 and 22, it is possible to find such a family with φ_ε(a)^2 = 1 for all ε > 0. The approximate representations γ are then constructed by taking tensor powers of direct sums of the approximate representations φ_ε, \overline{φ_ε}, and copies of the trivial representation, and will also satisfy γ(a)^2 = 1.

Taking τ = ε^2/4, the argument of the first paragraph gives ε-representations ψ of \widehat{G} with \|ψ(J) - 1\| = 2, so J is nontrivial in approximate representations of \widehat{G} as desired.

We are now ready to prove Proposition 35. Note that any hyperlinear but nonresidually finite group has an element which is trivial in finite-dimensional representations, but nontrivial in approximate representations. To prove Proposition 35, we show that

$$K = \langle x, y, a, b : a^2 = b^2 = e, ab = ba, yay^{-1} = a, yby^{-1} = ab, xyx^{-1} = y^2 \rangle$$

is an extended homogeneous-linear-plus-conjugacy group which is hyperlinear but nonresidually finite. Indeed, to see that K has a presentation as in Definition 32, we can introduce a third variable c with c^2 = e and c = ab. Then K is equivalent to the extended homogeneous-linear-plus-conjugacy group with three involutary generators a, b, c, one linear relation abc = e (along with the corresponding commuting relations), two noninvolutary generators x and y, and three conjugacy relations yay^{-1} = a, yby^{-1} = c, and xyx^{-1} = y^2. For the remainder of this section, K will refer to this group.

**Lemma 37.** K is sofic, and the element a ∈ K is nontrivial.

**Proof.** K_1 := \langle y, a, b : a^2 = b^2 = e, ab = ba, yay^{-1} = a, yby^{-1} = ab \rangle is isomorphic to a semidirect product \mathbb{Z} \rtimes (\mathbb{Z}_2 \times \mathbb{Z}_2), and in particular is solvable (hence amenable). The group K is the HNN extension of K_1 by the injective endomorphism of \langle y \rangle \cong \mathbb{Z} sending y \mapsto y^2. Hence K is sofic by properties (1) and (3) of sofic groups above. In addition, the natural morphism K_1 → K is injective. Since a is clearly nontrivial in K_1, we conclude that a is nontrivial in K.

The following lemma comes from discussions with Tobias Fritz.

**Lemma 38.** The element a ∈ K is trivial in all finite-dimensional representations of K.
Proof. By a theorem of Mal’cev [20], it suffices to show that $a$ is trivial in finite representations, rather than finite-dimensional representations. So let $\phi : G \rightarrow H$ be a homomorphism from $G$ to a finite group $H$. Now the order $k$ of $\phi(x)$ is finite, so $\phi(y) = \phi(x)^k \phi(y) \phi(x)^{-k} = \phi(y)^{2^k}$. It follows that the order $m = |\phi(y)|$ of $\phi(y)$ divides $2^k - 1$, and in particular is odd. Since $\phi(y)\phi(b)\phi(y)^{-1} = \phi(ab)$ and $\phi(y)\phi(ab)\phi(y)^{-1} = \phi(b)$, we conclude that $\phi(b) = \phi(y)^m \phi(b)\phi(y)^{-m} = \phi(ab)$. Consequently $\phi(a) = 1$ as desired.

Proof of Proposition 35. By Proposition 33, there is an $fa$-embedding of $K$ to a homogeneous-linear-plus-conjugacy group $G = \Gamma_0(A, C)$, in which $a \in K$ is mapped to a generator $x_i$ of $G$. Let

$$\widehat{G} = \langle G, t : t^2 = e, tx_i t = Jx_i \rangle_{\mathbb{Z}_2}. $$

The relation $tx_i t = Jx_i$ can be replaced with the relations $tx_i t = Z$ and $Zx_i = J$, where $Z$ is an ancilla variable with $Z^2 = e$. With this presentation, $\widehat{G}$ is a linear-plus-conjugacy group. By Proposition 27, there is an $fa$-embedding over $\mathbb{Z}_2$ of $\widehat{G}$ to a solution group $\Gamma$.

By Lemma 37, $a$ is nontrivial in approximate representations of $K$, and hence $x_i$ is nontrivial in approximate representations of $G$. By Lemma 36, $J\widehat{G}$ is nontrivial in approximate representations of $\widehat{G}$, and we conclude that $J\Gamma$ is nontrivial in approximate representations of $\Gamma$.

Finally, there is a morphism from $K$ to $\widehat{G}$ which sends $a$ to $x_i$, so $x_i$ will be trivial in all finite-dimensional representations of $\widehat{G}$ by Lemma 38. But since $J\widehat{G} = [t, x_i]$, this means that $J\widehat{G}$ (and hence $J\Gamma$) is trivial in all finite-dimensional representations of $\widehat{G}$.

Proof of Theorem 1. Let $\Gamma$ be the solution group from Proposition 35, and let $G$ be the associated game. Since $J$ is trivial in finite-dimensional representations, Theorem 18 implies that $G$ does not have a perfect strategy in $C_{qs}$. But since $J$ is nontrivial in approximate representations, Proposition 20 implies that $G$ has a perfect strategy in $C_{qa}$.

REMARK 39. By Remarks 30 and 34, the linear system constructed in the proof of Theorem 1 will have 235 variables and 184 equations.

6. Proofs of Theorems 2 and 3

To prove Theorem 2, we want to find a hyperlinear group with an undecidable word problem, which $fa$-embeds in a solution group. For Theorem 3, we want to
find a family of residually finite groups with arbitrarily hard (albeit computable) word problems, which \(\text{fin}\)-embed in solution groups. Fortunately, such groups are provided by Kharlampovich [14] and Kharlampovich, Myasnikov, and Sapir [15]. Since the presentations are rather complicated, we do not repeat them here. Instead, we summarize some points of the construction from [15] in the following theorem.

It is helpful to use the following notation: given \(S_0 \subseteq S_1\), let \(\mathcal{N}(S_0, S_1)\) denote the normal subgroup generated by \(S_0\) in the free group \(F(S_1)\). Note that if \(S_1 \subseteq S\), then \(\mathcal{N}(S_0, S_1)\) is a (not necessarily normal) subgroup of \(F(S)\) in a natural way. Also, if \(x, y\) are group elements, recall that \([x, y] = xyx^{-1}y^{-1}\), and \(x^y = yxy^{-1}\).

(This is the reverse of the convention in [15], where \([x, y] = x^{-1}y^{-1}xy\) and \(x^y = y^{-1}xy\).) Finally, recall that a set \(X \subseteq \mathbb{N}\) is said to be recursively enumerable if there is a Turing machine which takes integers \(n\) as input, and halts with output \(\text{true}\) if and only if \(n \in X\) (if \(n \notin X\) then the machine can either halt with output \(\text{false}\), or run forever).

**Theorem 40 ([15], see also [14]).** Let \(X \subseteq \mathbb{N}\) be recursively enumerable. Then there is a finitely presented solvable group \(K_X = \langle S : R \rangle\) with the following properties:

1. The set \(S\) is divided into three subsets \(L_i, i = 0, 1, 2\).
2. The relations in \(R\) come in three types:
   - (a) \(R\) contains the relations \(x^2 = e\) for all \(x \in L_0 \cup L_1\).
   - (b) \(R\) also contains commuting relations of the form \(xy = yx\), for certain pairs \(x, y \in S\).
   - (c) For every other relation \(r \in R\), there are some subsets \(S_1 \subseteq S\) and \(S_0 \subseteq (L_0 \cup L_1) \cap S_1\) such that \(r \in \mathcal{N}(S_0, S_1)\), and the image of \(\mathcal{N}(S_0, S_1)\) in \(K_X\) is abelian.
3. The image of \(\mathcal{N}(L_0, S)\) in \(K_X\) is abelian.
4. There are elements \(z_0, z_1 \in L_0, A_1, A_2 \in L_1,\) and \(a, a' \in L_2\), such that \(n \in X\) if and only if
   \[
   [A_2, [A_1, w(2^n)]] = [A_2, [A_1, z_0]]
   \]
   in \(K_X\), where \(w(m)\) is defined by
   
   \[
   w(m) := \begin{cases} 
   z_1 & m = 0, \\
   w(m-1)w(m-1)^{a-1}w(m-1)^aw(m-1)^{a'} & m \geq 1.
   \end{cases}
   \]

5. If \(X\) is recursive, then \(K_X\) is residually finite.
Note that there is some overlap between relations of type (2b) and (2c). Indeed, if \( [x, y] = e \) is a relation, then the image of \( \mathcal{N}(\{x\}, \{x, y\}) \) in \( K_X \) is equal to \( \langle x \rangle \), and in particular is abelian. Since \( [x, y] \) belongs to \( \mathcal{N}(\{x\}, \{x, y\}) \), any relation \( [x, y] = e \) of type (2b) with \( x \in L_0 \cup L_1 \) can also be regarded as a relation of type (2c).

**Proof.** For the convenience of the reader, we explain how to recover Theorem 40 from [15]. First, if \( X \) is recursively enumerable, then by [15, Theorem 2.7] there is a 2-glass deterministic Minsky machine \( MM_2 \) enumerating \( X \), in the sense that \( n \in X \) if and only if the machine takes input configuration \( (1; 2^n, 0) \) to the accept configuration \( (0; 0, 0) \) (this and other unexplained terminology and notation is as in [15]). Furthermore, if \( X \) is recursive, then we can take \( MM_2 \) to be a sym-universally halting 2-glass Minsky machine deciding \( X \). We let \( K_X \) be the finitely presented group \( G(MM_2) \) defined in relations (G1)–(G8) of [15, Section 4.1], with parameter \( p = 2 \). Then part (1) of the theorem is exactly the notation used in [15], part (3) follows immediately from [15, Lemma 4.5], and part (5) is [15, Theorem 4.18].

It remains to show that \( K_X \) satisfies parts (2) and (4) of the theorem. The relations in part (2a) follow from the relations in (G1). For the rest of part (2), we need to show that every other relation in (G1)–(G8) falls under one of (2b) or (2c). It is immediate that the relations in (G1) (excepting the relations from (2a)) and the relations in (G2) fall under (2b). The relations in (G3) and (G4) are written down in parts (1)–(3) of [15, Lemma 4.1]. The relations in part (1) of this lemma are just commuting relations, and hence fall under (2b). Using the notation of this lemma, if we set \( S_0 = X \) and \( S_1 = X \cup F \cup F' \), then the relations in parts (2) and (3) of Lemma 4.1 belong to \( \mathcal{N}(S_0, S_1) \). The fact that \( \mathcal{N}(S_0, S_1) \) is abelian is the conclusion of Lemma 4.1. The specific set \( X \) and sets \( F, F' \) used in (G3) and (G4) are contained in \( L_1 \) and \( L_2 \), respectively, so these relations fall under (2b) and (2c). Finally, the relations in (G5)–(G8) belong to \( \mathcal{N}(L_0, S) \), and hence fall under (2c) by part (3) of the theorem.

For part (4), note that the element \( w(m) \) belongs to \( \mathcal{N}(L_0, S) \), which is abelian. Since \( z_1^2 = e \) by part (2a), we can conclude by induction that \( w(m)^2 = e \) for all \( m \geq 0 \). Using the notation of [15, Theorem 4.3(b)], let \( z_0 = x(q_0A_0) \), \( z_1 = x(q_1A_1) \), \( a = a_1 \), \( a' = a'_1 \), and \( A_1, A_2 \) be the symbols of the same name. Using again that \( \mathcal{N}(L_0, S) \) is abelian, it follows that \( w(m) = \mathcal{N}(L_0, S) \) is the element denoted in [15] by \( x(q_1A_0) * a_1^{(m)} \), that \( [A_2, [A_1, w(m)]] \) is the element \( x(q_1A_0) * a_1^{(m)} * A_1 * A_2 \), and that \( [A_2, [A_1, z_0]] \) is the element \( x(q_0A_0) * A_1 * A_2 \). Thus [15, Theorem 4.3(b) and Section 3.1] imply that \( [A_2, [A_1, w(m)]] = [A_2, [A_1, z_0]] \) if and only if \( MM_2 \) takes input configuration \( (1; m, 0) \) to accept configuration \( (0; 0, 0) \). Hence part (4) of the theorem follows from the definition of \( MM_2 \).
We show that every group with a presentation as in parts (1) and (2) of Theorem 40 is an extended homogeneous-linear-plus-conjugacy group (as in Definition 32). We first illustrate this with an example.

**Example 41.** Let

$$K = \langle x_1, x_2, y_1, y_2 : x_1^2 = x_2^2 = [x_1, x_2] = [x_2, y_1] = [y_1, y_2] = x_1x_2y_1^{-1}x_1^{-1}y_1^{-1} = e \rangle.$$  

This presentation of $K$ satisfies parts (1) and (2) of Theorem 40 with $L_0 \cup L_1 = \{x_1, x_2\}$ and $L_2 = \{y_1, y_2\}$. Indeed, all the defining relations fall under (2a) and (2b) except the last, which belongs to $\mathcal{N}(S_0, S_1)$ with $S_0 = \{x_1, x_2\}$, $S_1 = \{x_1, x_2, y_1\}$. Since $x_1^{y_1}, x_1^{y_1^{-1}}$, and $x_1x_2$ are all involutions in $K$, the last relation implies that $x_1^{y_1}, x_1^{y_1^{-1}}$, and $x_1x_2$ commute in $K$. But $x_2$ commutes with both $x_1$ and $y_1$, and hence commutes with $x_1^{y_1}$ and $x_1^{y_1^{-1}}$, so $x_1$ also commutes with $x_1^{y_1}$ and $x_1^{y_1^{-1}}$. It follows that the image of $\mathcal{N}(S_0, S_1)$ is abelian in $K$, and hence the last defining relation falls under (2c).

To construct a presentation of $K$ as an extended homogeneous-linear-plus-conjugacy group, note that the generators split into involutary generators $x_1, x_2$ and noninvolutary generators $y_1, y_2$. The order in which we list the involutary generators does not matter, but for the noninvolutary generators the order is significant. Most of the defining relations for $K$ are commuting relations, and these can be replaced with the conjugacy relations $x_1x_2x_1 = x_2$, $y_1x_2y_1^{-1} = x_2$, and $y_2y_1^{-1} = y_2$. Note that we could choose $x_2x_1x_2 = x_1$ in place of $x_1x_2x_1 = x_2$, but that we are forced to pick $y_2y_1^{-1} = y_1$ over $y_1y_2y_1^{-1} = y_2$ by the ordering on $y_1$ and $y_2$. For the last defining relation, add two new involutary generators $x_3$ and $x_4$, along with relations $y_1x_1y_1^{-1} = x_3$ and $y_1x_4y_1^{-1} = x_1$. The relation $x_1x_2x_1y_1x_1^{y_1^{-1}} = e$ can then be replaced with the linear relation $x_1x_2x_3x_4 = e$. Note that, following the convention introduced in Section 4.1, this last linear relation is really the set of relations $\{x_1x_2x_3x_4 = e\} \cup \{[x_i, x_j] = e \text{ for } i \neq j\}$. However, the additional relations $[x_i, x_j] = e$ do not change the group, since $x_1^{y_1}, x_1^{y_1^{-1}}, x_1, x_2$ commute in $K$.

We conclude that if $A = (1 \ 1 \ 1 \ 1), \ C_0 = \{(1, 2, 2)\}, \ C_1 = \{(1, 2, 2), (1, 1, 3), (1, 4, 1)\}$, and $L = (0 \ 0 \ 0)$, then the natural map

$$K \to E\Gamma_0(A, C_0, C_1, L) : x_i \mapsto x_i, i = 1, 2 \quad \text{and} \quad y_i \mapsto y_i, i = 1, 2$$

is an isomorphism. Note that the relation $x_1x_2x_1 = x_2$ coming from including $(1, 2, 2)$ in $C_0$ is actually redundant, since this relation is also implied by the linear relation, but we include it to illustrate how to handle this type of commuting relation in general.
We now turn to the general case.

**Lemma 42.** Suppose $K = \langle S : R \rangle$ is a finitely presented group satisfying properties (1) and (2) of Theorem 40. Then $K$ is an extended homogeneous-linear-plus-conjugacy group. Furthermore, if $S_0 \subseteq S_1 \subseteq S$ are two subsets such that $S_0 \subseteq L_0 \cup L_1$, and the image of $\mathcal{N}(S_0, S_1)$ in $K$ is abelian, then for every $w \in \mathcal{N}(S_0, S_1)$, there is a presentation of $K$ as an extended homogeneous-linear-plus-conjugacy group in which $w$ is equal in $K$ to one of the involutary generators $x_j$.

**Proof.** The proof is exactly as in Example 41. The generators $S$ of $K$ split into involutary generators $L_0 \cup L_1 = \{x_1, \ldots, x_{n_0}\}$ and noninvolutary generators $L_2 = \{y_1, \ldots, y_{\ell}\}$, where we choose an arbitrary order on each set (although only the order on $L_2$ is significant). We show that there are $m \geq 0$, $n \geq n_0$, an $m \times n$ matrix $A$ over $\mathbb{Z}_2$, subsets $C_0 \subseteq [n] \times [n] \times [n]$ and $C_1 \subseteq [\ell] \times [n] \times [n]$, and an $\ell \times \ell$ lower-triangular matrix $L$, such that the natural inclusion

$$K \to E \Gamma_0(A, C_0, C_1, L) : \ x_i \mapsto x_i \quad \text{for } 1 \leq i \leq n_0 \quad \text{and}$$

$$y_i \mapsto y_i \quad \text{for } 1 \leq i \leq \ell$$

is an isomorphism which sends the element $w$ to some $x_j$. Indeed, in property (2) of Theorem 40, the defining relations for $K$ split into three types, type (2a), (2b), and (2c). The relations of type (2a) simply say that the generators $L_0 \cup L_1$ are involutions, and the commuting relations in type (2b) are equivalent to conjugacy relations (where, for relations $y_i y_j = y_j y_i$, we choose either $y_i y_j y_i^{-1} = y_j$ or $y_j y_i y_j^{-1} = y_i$ depending on whether $i > j$ or $i < j$). Thus the only difficulty in constructing the presentation $E \Gamma_0(A, C_0, C_1, L)$ is handling the relations of type (2c). For these relations, we prove the following claim: if $S_0 \subseteq S_1 \subseteq S$ are subsets such that $S_0 \subseteq L_0 \cup L_1$, and the image of $\mathcal{N}(S_0, S_1)$ in $K$ is abelian, then there is a set of indeterminates $S_w = \{x_{n_0} + 1, \ldots, x_q\}, q \geq n_0$, and a set of relations $R_w$, such that:

(i) $R_w$ consists of linear relations $x_{i_1} \cdots x_{i_r} = e, 1 \leq i_1 < \cdots < i_r \leq q$ (which, as in Section 4.1, include the commuting relations $[x_{i_j}, x_{i_k}] = e$ for all $1 \leq j, k \leq r$), conjugacy relations $x_i x_j x_i = x_k, 1 \leq i, j, k \leq q$, and conjugacy relations $y_i x_j y_i^{-1} = x_k, 1 \leq i \leq \ell$ and $1 \leq j, k \leq q$;

(ii) the relations

$$\tilde{R}_w := R_w \cup \{s^2 = e : s \in L_0 \cup L_1 \cup S_w\}$$

imply that $w$ is equal to an element of $S_0 \cup S_w$; and
(iii) the added generators $S_w$ and relations $R_w$ do not change the group, that is the inclusion

$$K \to \langle K, S_w : R_w \cup \{ s^2 = e : s \in S_w \} \rangle$$

is an isomorphism.

To prove the claim, we use induction on the length of $w$ in $\mathcal{F}(S_1)$. The claim is trivially true if $w \in S_0 \cup S_0^{-1}$. Suppose $w = zw'z^{-1}$, where $w' \in \mathcal{N}(S_0, S_1)$ has length less than $w$, and $z \in S_1$. By induction, there is a set of ancilla variables $S_w' = \{ x_{n_0+1}, \ldots, x_{q'} \}$ and relations $R_w'$ satisfying properties (i)–(iii) for $w'$. In particular, the relations $\tilde{R}_w'$ imply that $w'$ is equal to some $X \in S_0 \cup S_w'$. Then we can set $q := q' + 1$, $S_w := S_w' \cup \{ x_q \}$, and $R_w := R_w' \cup \{ x_q = zXz \}$ or $R_w' \cup \{ x_q = zXz^{-1} \}$ depending on whether $z \in L_0 \cup L_1$ or $z \in L_2$. If $w = z^{-1}w'z$, then we do the same thing, but using $z x_q z^{-1} = X$ in place of $x_q = zXz^{-1}$. In both cases, the relations $\tilde{R}_w$ imply that $w$ is equal to $x_q$, so properties (i)–(iii) hold for $R_w$. Similarly, suppose that $w = w_1 \cdots w_k$, where each $w_i \in \mathcal{N}(S_0, S_1)$ has smaller length than $w$. By induction, there are sets $S_{w_i} = \{ x_{q_i+1}, \ldots, x_{q_i} \}$, where $n_0 = q_0 \leq q_1 \leq \cdots \leq q_k$, and relations $\tilde{R}_{w_i}$ implying that $w_i$ is equal to some $X_i \in S_0 \cup S_{w_i}$. We then set $q := q_k + 1$, $S_w := \bigcup S_{w_i} \cup \{ x_q \}$, and

$$R_w := \bigcup \tilde{R}_{w_i} \cup \{ x_q X_1 \cdots X_k = e = [x_q, X_i] = [X_i, X_j] \text{ for all } 1 \leq i, j \leq k \}$$

(in other words $R_w$ contains all relations from $R_{w_i}$ and the single linear relation $x_q X_1 \cdots X_n = e$). Since the image of $\mathcal{N}(S_0, S_1)$ in $K$ is abelian, adding the relations $R_w$ does not change $K$, so properties (i)–(iii) hold again, finishing the proof of the claim.

Now suppose that $K$ has a defining relation $r$ of type (2c), so $r \in \mathcal{N}(S_0, S_1)$ for some $S_0 \subseteq L_0 \cup L_1$ and $S_1 \subseteq S$ such that the image of $\mathcal{N}(S_0, S_1)$ is abelian in $K$. If $r = zr'z^{-1}$ for some $r' \in \mathcal{N}(S_0, S_1)$ and $z \in S_1 \cup S_1^{-1}$, then $r$ can be replaced with the simpler relation $r'$. Hence we can assume without loss of generality that $r = r_1 \cdots r_n$, where each $r_i \in \mathcal{N}(S_0, S_1)$. By the claim, we can add ancilla variables and relations as in property (i) such that each $r_i$ is equal to an involutary generator $X_i$ in $K$, and the relation $r$ can be replaced with the linear relation $X_1 \cdots X_n = e$. By repeatedly applying the claim to all relations of type (2c), we can construct a presentation of $K$ as an extended homogeneous-linear-plus-conjugacy group. Furthermore, the claim immediately implies that we can construct such a presentation sending a specified element $w$ to some involutary generator, as required. \hfill \Box

We now come to the main result of this section.

**Proposition 43.** Let $X \subseteq \mathbb{N}$ be a recursively enumerable set. Then there is a family of solution groups $\Gamma_n = \Gamma(A^{(n)}, b^{(n)}), n \geq 1$, such that:
The set of quantum correlations is not closed

(a) \( A^{(n)}x = b^{(n)} \) is an \( \exp(O(n)) \times \exp(O(n)) \) linear system;
(b) the function \( n \mapsto (A^{(n)}, b^{(n)}) \) is computable in \( \exp(O(n)) \)-time;
(c) \( J_{\Gamma_n} \) is nontrivial in \( \Gamma_n \) if and only if \( n \in X \);
(d) if \( J_{\Gamma_n} \) is nontrivial in \( \Gamma_n \), then \( J_{\Gamma_n} \) is nontrivial in approximate representations; and
(e) if \( X \) is recursive and \( J_{\Gamma} \) is nontrivial in \( \Gamma_n \), then \( J_{\Gamma_n} \) is nontrivial in finite-dimensional representations.

Before giving the proof, we need the following exact version of Lemma 36.

**Lemma 44.** Suppose \( G = \langle S : R \rangle \) is a finitely presented group, where \( R \) contains the relation \( a^2 = e \) for some \( a \in S \). Let

\[
\widehat{G} := \langle G, t : t^2 = e, tat = Ja \rangle_{\mathbb{Z}_2},
\]

where \( J, t \notin S \). If \( a \) is nontrivial in finite-dimensional representations of \( G \), then \( J \) is nontrivial in finite-dimensional representations of \( \widehat{G} \).

**Proof.** Suppose \( a \) is nontrivial in finite-dimensional representations of \( G \). A theorem of Baumslag states that the free product of two residually finite groups amalgamated over a finite subgroup is residually finite [1]. Let \( \widetilde{G} := G \times \mathbb{Z}_2 \), where the generator of the \( \mathbb{Z}_2 \) factor is denoted by \( J \), and let

\[
H = \langle t, a : t^2 = a^2 = e, tat = aJ \rangle_{\mathbb{Z}_2} \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2.
\]

Then \( \widehat{G} \) is isomorphic to the amalgamated free product of \( \widetilde{G} \) and \( H \) over \( \langle a, J \rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \), a finite group. While \( \widehat{G} \) is not necessarily residually finite, the group \( \widehat{G}^\text{fin} \) is residually finite by definition, and there is a natural map from \( \widehat{G} \) to the amalgamated free product of \( \widehat{G}^\text{fin} \) and \( H \) over \( \mathbb{Z}_2 \times \mathbb{Z}_2 \). The image of \( J_{\widehat{G}} \) is nontrivial in \( \widehat{G}^\text{fin} \), and hence in the amalgamated product of \( \widehat{G}^\text{fin} \) and \( H \). So \( J \) is nontrivial in finite-dimensional representations of \( \widehat{G} \) by Baumslag's result.

**Proof of Proposition 43.** Given a recursively enumerable subset \( X \subseteq \mathbb{N} \), let \( K_X = \langle S : R \rangle \) be the associated group from Theorem 40. Using the notation from property (4) of Theorem 40, let \( c(n) = [A_2, [A_1, w(2^n)]]A_2, [A_1, z_0]^{-1} \), so that \( c(n) = e \) in \( K_X \) if and only if \( n \in X \). Since \( c(n) \) belongs to \( \mathcal{N}(L_0, S) \), Lemma 42 and property (3) of Theorem 40 implies that \( K_X \) has a presentation \( E \Gamma_0(A^{(n)}, C_0^{(n)}, C_1^{(n)}, L^{(n)}) \) as an extended homogeneous-linear-plus-conjugacy group, in which \( c(n) \) is equal to some involuntary generator \( x_i \). For the purposes of this
Given a computable function \( f \), suppose \( X \subseteq \mathbb{N} \) is a recursively enumerable but nonrecursive set, and take the family \( \{ G_n : n \in \mathbb{N} \} \) of games associated to the solution groups \( \{ \Gamma_n : n \in \mathbb{N} \} \) constructed in Proposition 43. By Theorem 19 and part (c) of Proposition 43, \( G_n \) will have a perfect strategy in \( C_{qa} \) if and only if \( n \in X \). By Proposition 20 and part (d) of Proposition 43, \( G_n \) will have a perfect strategy in \( C_{qc} \) if and only if it has a perfect strategy in \( C_{qa} \). Because the function \( n \mapsto G_n \) is computable by part (b) of Proposition 43, it is undecidable to determine if the games in this family have perfect strategies in \( C_{qa} \).

Proof of Theorem 3. Given a computable function \( f(n) \), let \( X \subseteq \mathbb{N} \) be a recursive subset such that for any Turing machine accepting \( X \), the running time
over inputs \( n \leq N \) is at least \( f(N) \) when \( N \) is sufficiently large. (As mentioned in the introduction, often when talking about the running time, we look at the maximum running time over inputs of size \( n \leq N \), rather than value \( n \leq N \). However, thinking of the running time in terms of the values of the inputs does not change the fact that such sets \( X \) exist.) As in the proof of Theorem 2, we can take the family of games \( \{ G_n : n \in \mathbb{N} \} \) associated to the solution groups \( \{ \Gamma_n : n \in \mathbb{N} \} \) from Proposition 43. Then part (a) of Theorem 3 follows from parts (a) and (b) of Proposition 43, while parts (b) and (c) of Theorem 3 follow from parts (c) and (e) of Proposition 43, as well as Theorems 18 and 19.

**Proof of Corollary 4.** Suppose there is an algorithm to decide if a linear system game has a perfect strategy in \( C_q \). Let \( g(n) \) be the maximum running time of this algorithm on games coming from linear systems with at most \( n \) rows and columns. Note that \( g(n) \) is an increasing function. Let \( f(n) \) be any computable function such that

\[
f(n) > g(2n^2) + 2n^2
\]

for all \( n \geq 1 \). Let \( \mathcal{G}_n \) be the family of games associated to \( f(n) \) as in Theorem 3. Then there is a constant \( C \) such that \( \mathcal{G}_n \) has size \( \leq 2^{Cn} \) for all \( n \geq 1 \), and the function \( n \mapsto \mathcal{G}_n \) is computable in time \( 2^{Cn} \). Plugging \( \mathcal{G}_n \) into the algorithm to decide whether a linear system game has a perfect strategy in \( C_q \), we get an algorithm for the language

\[
X = \{ n \in \mathbb{N} : \mathcal{G}_n \text{ has a perfect strategy in } C_q \}
\]

with running time at most \( g(2^{CN}) + 2^{CN} \) on inputs \( n \leq N \). But by part (b) of Theorem 3, when \( N \) is sufficiently large the maximum running time on inputs \( n \leq N \) for any algorithm for \( X \) must be at least \( f(N) \). Since \( N^2 \) will eventually be larger than \( CN \), we get a contradiction. Thus there is no algorithm to decide if a linear system game has a perfect strategy in \( C_q \).

**Acknowledgements**

Acknowledgements I thank Jason Crann, Richard Cleve, Tobias Fritz, Li Liu, Martino Lupini, Narutaka Ozawa, Vern Paulsen, Mark Sapir, Jamie Sikora, and Thomas Vidick for helpful comments and conversations. This work was partially supported by NSERC grant number 2018-03968.

**References**


The set of quantum correlations is not closed


