Constant-sized correlations are sufficient for robust self-testing of maximally entangled states with unbounded dimension

Honghao Fu[1] and Carl Miller[1,2]

[1]Department of Computer Science, Institute for Advanced Computer Studies, and Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, MD 20742, USA
[2]National Institute of Standards and Technology, 100 Bureau Dr., Gaithersbug, MD 20899, USA

May 3, 2019

Abstract

We show that for any prime odd integer d with primitive r, there exists a correlation of size $\Theta(r^2)$ that can robustly self-test a maximally entangled state of dimension $4d-4$. Our result is inspired by the robust self-testing result by Bamps and Pironio (*Physical Review A* 91.5 (2015)) and techniques introduced by Slofstra (*Forum of Mathematics, Pi.* Vol. 7, 2019). Since there are infinitely many prime numbers with primitive root at most 5 (*The Mathematical Intelligencer* 10.4 (1988)), our result implies that constant-sized correlations are sufficient for robust self-testing of maximally entangled states with unbounded local dimension.

1 Introduction

Self-testing is a unique phenomenon of quantum mechanics. It has many applications in quantum delegated computation [**?**, **?**] and device independent quantum cryptography [**?**, **?**, **?**, **?**, **?**].

The case of self-testing 2-dimensional EPR pair is fully understood. One can robustly self-test one copy of it by the CHSH inequality [**?**]. and self-test many copies of the 2-dimensional EPR pair in parallel [**?**, **?**]. Self-testing general d-dimensional EPR pairs is a harder task. Proving robustness is harder.

At a high level, general d-dimensional maximally entangled states are self-tested by modifying the correlation and enlarging the size of the correlation. A natural question to ask is whether maximally entangled state with large local dimension can be self-tested with fixed-sized correlation. An equivalent question to ask is whether it is possible to self-test maximally entangled state of some local dimension more efficiently, with constant correlation size. In this report, we give an affirmative answer to this question by proving the following theorem.

Theorem 1.1 (Informal)**.** *There exists an infinite-sized set D of odd prime numbers such that, for any $d \in D$, the maximally entangled state of local dimension $d-1$ can be self-tested with constant-sized question and answer alphabets.*

The set D is easily characterizable as it contains all the odd prime numbers with smallest primitive root 2, 3 or 5. It has been shown that there are infinitely many prime numbers with smallest primitive root in the set $\{2, 3, 5\}$ [?], so the set D has infinitely many elements. To prove Theorem 1.1, we give explicit self-testing proof of the maximally entangled state with local dimension $d-1$ where the primitive root of d is 2,3, or 5, by explicitly giving the correlation that achieves self-testing. Our correlation is denoted by $C(d^{(r)})$ for prime d. We use the superscript (r) to denote the primitive root of d. Note that although the size of $C(d^{(r)})$ does not depend on d, the optimal correlation does.

In order to accomplish our goal, we introduce new techniques for self-testing. First of all, we use a different variant of the weighted CHSH inequality, which was first introduced in [?], to enforce the eigenvalue of some unknown operator which is the product of two binary observables used in the weighted CHSH test. The variant of the CHSH inequality that we use is not used in the self-testing literature before. Secondly, we give a new way to decompose unitaries of arbitrary order into binary observables which maintains certain commutation relations. Such decomposition is different from what Slofstra used in his work [?]. Our work is heavily based on [?], and we need to make that clearer. -Carl Intuitively, such decomposition can be seen as the inverse of the Jordan's lemma decomposition. The third contribution is that we prove self-testing without using anti-commutation relations between Pauli operators, which is the core idea in most of the previous self-testing results. Are we sure about that ("all")? -Carl Instead, we find a new pair of operators that can generate the ring of matrices over complex numbers.

Structure of the paper. We start with notations and background information in Section 2. Since the correlation we designed can win a special linear system game and satisfy an extended weighted CHSH test, we introduce the linear system game in Section 3 and the extended weighted CHSH test in Section 4. Our main result is based on the combination of the two tests and presented in Section 6.

2 Preliminaries and notations

2.1 Notations

The maximally entangled state of local dimension $d-1$ for some odd prime d that we are interested in is denoted by

$$|\Sigma^{(d-1)}\rangle = \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} |j\rangle|d-j\rangle. \tag{1}$$

The superscript $(d-1)$ stresses the dimension of the local Hilbert space and we follow this convention through this paper. Usually, we denote a general Hilbert space by $\mathcal{H}$. That last sentence is too vague. Does $\mathcal{H}$ refer to a specific Hilbert space? -Carl How about now? -H.F. Note that we only work with finite Hilbert spaces in this work. When there are multiple Hilbert spaces, we label them with subscripts, for example, $\mathcal{H}_A$ and $\mathcal{H}_B$. The d-th root of unity is denoted by $\omega_d := e^{2\pi i/d}$. In the qubit case, the maximally entangled state is denoted by

$$|\Sigma^{(2)}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \tag{2}$$

and the Pauli operators are given by

$$\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0| \qquad \sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|. \tag{3}$$

For quantum states and operators on different Hilbert spaces, we use subscripts to label them. For example, O_A is an operator on Alice's side and $|0\rangle_B$ is a state on Bob's side. The only exception is for projectors used in a quantum strategy defined below, where the subscript is the input and the superscript is the output. For example, M_x^a is Alice's projector for input x and output a. Operators defined by the projectors follow the same convention, for example, $M_x := M_x^0 - M_x^1$.

We use the following notation for the closeness between quantum states.

$$|u\rangle \simeq_\epsilon |v\rangle \iff \| |u\rangle - |v\rangle \| \leq \epsilon. \tag{4}$$

Similarly, for numbers $a, b \in \mathbb{C}$, the closeness is denoted by $a \simeq_\epsilon b \iff \|a - b\| \leq \epsilon$.

We use some basic number theory in our work.

Definition 2.1. *A primitive root of a prime number d is an integer r such that $r \pmod{d}$ has multiplicative order $d-1$ in the multiplicative group of integers modulo d, $\mathbb{Z}_d^\times$. I'm skeptical about whether that's the standard definition. (For example, the definition on wikipedia doesn't include that "smallest integer" part.) -Carl How about now? -H.F.*

In other words, r is the generator of $\mathbb{Z}_d^\times$. In our work, we always pick r to be the smallest primitive root of d.

2.2 Nonlocal games

In a nonlocal game, there are two players, Alice and Bob and each of them is requested to give an answer for a randomly chosen question. We denote Alice's question set by $\mathcal{X}$ and her answer set by $\mathcal{A}$. Similarly, Bob's question set is denoted by $\mathcal{Y}$ and his answer set is denoted by $\mathcal{B}$. The nonlocal game also comes with two functions: $\pi : \mathcal{X} \times \mathcal{Y} \to [0,1]$, which is the probability distribution over the questions, and $V : \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y} \to \{0,1\}$, which is the scoring function. The score 0 and 1 correspond to losing and winning. Such games are nonlocal because Alice and Bob cannot communicate after getting their questions but they may share some strategy beforehand.

In our work, we consider a special case of nonlocal games, which we call a nonlocal test. There are three differences between a nonlocal game and a nonlocal test:

- only a subset of all the question pairs are given to Alice and Bob;
- the answer sets for different questions are different;
- the scoring function maps input-output pairs to real numbers ($\mathbb{R}$).

A quantum strategy of a nonlocal game presented in terms of projective measurements consists of projective measurements $\{\{M_x^a\}_a\}_x$ on Alice's side, $\{\{N_y^b\}_b\}_y$ on Bob's side, and a shared state $|\psi\rangle$, where $(M_x^a)^2 = M_x^a = (M_x^a)^\dagger$ and $(N_y^b)^2 = N_y^b = (N_y^b)^\dagger$. Then Alice and Bob's quantum strategy produces the conditional probability distribution

$$P(ab|xy) = \langle\psi|M_x^a \otimes N_y^b|\psi\rangle \text{ for all } (a,b,x,y) \in \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y}, \tag{5}$$

For simplicity we omit the tensor product between operators acting on different Hilbert spaces in the rest of the work.

Definition 2.2. *Given sets $\mathcal{X}, \mathcal{Y}, \mathcal{A}, \mathcal{B}$, a correlation is a collection of conditional probability distributions $\{P(ab|xy) : a \in \mathcal{A}, b \in \mathcal{B}\}_{x \in \mathcal{X}, y \in \mathcal{Y}}$.*

A particular type of nonlocal games that we are interested in is called linear system games, which is defined below.

Definition 2.3 (Linear system game). *Let $H\underline{x} = \underline{c}$ be an $m \times n$ system of linear equations over $\mathbb{Z}_2 = \{0, 1\}$, where H is an m-by-n matrix with entries in $\mathbb{Z}_2$ and $\underline{c}$ is a length-n vector with entries in $\mathbb{Z}_2$. The associated linear system game involves two players Alice and Bob, where Alice is given an equation number $i \in \mathcal{X} = \{1 \ldots m\}$ and replies with $\underline{a} \in \mathcal{A} = \mathbb{Z}_2^{\times n}$, and Bob is given a variable number $j \in \mathcal{Y} = \{1 \ldots n\}$ and replies with an assignment $b \in \mathcal{B} = \mathbb{Z}_2$. The scoring function is defined by*

$$V(\underline{a}, b, i, j) = \begin{cases} 1, & \text{if } \sum_{k=1}^{n} H(i,k)\underline{a}(k) \equiv c(i) \pmod 2 \text{ and } \underline{a}(j) = b \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

A widely-used example of linear system games is the Magic Square game introduced in Ref. [?]. For the formulation of the Magic Square game as a linear system game and the general linear system games over $\mathbb{Z}_d$, we refer to Section 3.1 of Ref. [?].

In the rest of the paper, we will work with quantum strategies for the linear system game presented in terms of binary observables.

Definition 2.4 (Quantum strategy of a linear system game). *A quantum strategy presented in terms of binary observables for the linear system game $(H\underline{x} = \underline{c})$ consists of*

1. *a pair of finite-dimensional Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$;*
2. *a collection of binary observables N_j, $1 \leq j \leq n$, on $\mathcal{H}_B$ such that $N_j^2 = \mathbb{1}$ for every $1 \leq j \leq n$;*
3. *a collection of binary observables M_{ij}, $1 \leq i \leq m$, $1 \leq j \leq n$ on $\mathcal{H}_A$ such that*
   (a) *$M_{ij}^2 = \mathbb{1}$ for every i, j,*
   (b) *$\Pi_j M_{ij}^{H(i,j)} = (-\mathbb{1})^{\underline{c}(i)}$ for every i, and*
   (c) *$M_{il}M_{ik} = M_{ik}M_{il}$ for every i and $H(i,l) = H(i,k) = 1$;*

   and
4. *a quantum state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$.*

Note that any quantum strategy presented in terms of binary observables can be converted to a quantum strategy presented in terms of projective measurement by looking into the spectral decompositions of the observables. More specifically, the fact that all the observables $\{M_{ij}\}_{j:H(i,j)=1}$ commute with each other implies that they share a common eigenspace. The collection of eigenvalues associated with one eigenstate is the assignment to the whole equation when this eigenstate is measured in a projective measurement strategy. If you're going to make that claim, you should make it more precise. Add a sentence explaining briefly how the construction is done. -Carl

A quantum strategy using binary observables for the linear system game associated with $H\underline{x} = \underline{c}$ can be constructed from a finite-dimensional representation of a finitely presented group over $\mathbb{Z}_2$, which is called the solution group. Here we skip the background about group representations and presentations. For concepts related to group presentations, we refer to Sec. 2 of Ref. [?]. For concepts related to group representations, we refer to Sec. 2.5 of Ref. [?]. The phrase "the finite-dimensional representation" is confusing – it sound like you are implying that there is only one representation. -Carl

Definition 2.5 (Solution group of a linear system game). *Let $H\underline{x} = \underline{c}$ be an $m \times n$ linear system. The solution group of this system is the group*

$$\begin{aligned}\Gamma(H, \underline{c}) := \langle x_1, \dots x_n, J : & J^2 = x_i^2 = e, Jx_i = x_iJ \text{ for all } 1 \leq i \leq n, \\ & \Pi_{j=1}^n x_j^{H(i,j)} = J^{\underline{c}(i)} \text{ for all } 1 \leq i \leq m, \text{ and} \\ & x_l x_k = x_k x_l \text{ if } H(i,k) = H(i,l) = 1 \text{ for some } i\rangle.\end{aligned}$$

We remark that a linear system game can be extracted from a solution group and *vice versa*. In later parts of this work, we give the solution group instead of the set of linear equations when we defines a linear system game.

Let Ψ be a finite-dimensional representation of $\Gamma(H, \underline{c})$ such that $\Psi(J) = -\mathbb{1}$ on the carrier space $\mathbb{C}^d$, then a perfect strategy for the linear system game ($H\underline{x} = \underline{c}$) is given by

1. $M_{ij} = \Psi(x_j)$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$,
2. $N_j = \Psi(x_j)$ for all $1 \leq j \leq n$, Do you need a transpose symbol in here somewhere? -Carl I don't think we need to take transpose because these are binary observables. -HF A binary observable can still be different from its transpose. -Carl
3. the maximally entangled state $|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{k=1}^d |k\rangle|k\rangle$.

However, our strategy uses a different maximally entangled state, so we justify our strategy more carefully.

2.3 The weighted CHSH inequality [?].

The weighted CHSH inequality a variation of the CHSH inequality [?]. We formulate it as a nonlocal test. The input sets and output sets of Alice and Bob are $\mathcal{X} = \mathcal{Y} = \{1, 2\}$ and $\mathcal{A} = \mathcal{B} = \mathbb{Z}_2$. The scoring function for the α-weighted CHSH test, for $|\alpha| \geq 1$, is defined by

$$V(a, b, x, y)_{\alpha-CHSH} = \begin{cases} \alpha(-1)^{a+b}, & \text{if } x = 1 \\ (-1)^{a+b}, & \text{if } x = 2,\ y = 1 \\ (-1)^{1+a+b}, & \text{if } x = y = 2. \end{cases} \tag{7}$$

The input distribution is $\pi(x, y) = 1/4$ for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$. Let Alice and Bob's quantum strategy in terms of projectors be $(\{\{M_x^a\}_a\}_x, \{\{N_y^b\}_b\}_y, |\psi\rangle)$. We define binary observables $M_x := M_x^0 - M_x^1$ and $N_y := N_y^0 - N_y^1$ from Alice and Bob's projectors. The weighted CHSH inequality states that

$$\langle \mathcal{I}_\alpha \rangle = \alpha\langle M_1 N_1 \rangle + \alpha\langle M_1 N_2 \rangle + \langle M_2 N_1 \rangle - \langle M_2 N_2 \rangle \leq 2|\alpha|, \tag{8}$$

where $\langle M_x N_y \rangle := \langle \psi | M_x N_y | \psi \rangle$ is the expectation value of the observables. The weighted CHSH inequality is true if Alice and Bob share product state $|\phi\rangle = |\phi_A\rangle \otimes |\phi_B\rangle$. However, if they share an entangled state $|\psi\rangle$, the value of $\langle \mathcal{I}_\alpha \rangle$ can be as large as $\langle \mathcal{I}_\alpha \rangle_{\max} = 2\sqrt{1 + \alpha^2}$ [?]. You need to cite the results that you're summarizing here (or refer to an appendix). -Carl I cited the paper in the title of this subsection. Maybe I add a sentence saying that this subsection summarizes results in [?]? -H.F.

Definition 2.6 (Ideal strategy to achieve $\langle \mathcal{I}_\alpha \rangle_{\max}$ [?]). *Define $\mu = \arctan(1/\alpha)$. The ideal strategy for weighted CHSH with parameter α, which achieves the value $\langle \mathcal{I}_\alpha \rangle_{\max}$, consists of the joint state $|\Sigma^{(2)}\rangle$ and observables $\tilde{M}_1 = \sigma_z$, $\tilde{M}_2 = \sigma_x$, $\tilde{N}_1 = \cos(\mu)\sigma_z + \sin(\mu)\sigma_x$ and $\tilde{N}_2 = \cos(\mu)\sigma_z - \sin(\mu)\sigma_x$.*

An interesting observation of the weighted CHSH inequality is that if some strategy can achieve value $\langle \mathcal{I}_\alpha \rangle$ close to $\langle \mathcal{I}_\alpha \rangle_{\max}$, then the strategy is close to the ideal strategy up to some local isometry, which is a phenomenon referred as a robust self-test. We give the formal statement of this self-testing property of $\langle \mathcal{I}_\alpha \rangle_{\max}$ in Section 4. In Section 5 we design a nonlocal test call the extended weighted CHSH test based on the weighted CHSH inequality.

2.4 Approximation tools

In the rest of the paper, we use the following approximation results implicitly and explicitly.

Proposition 2.7. *Let $|v\rangle$ and $|v'\rangle$ be vectors such that the relation $|v\rangle \simeq_\epsilon |v'\rangle$ holds, then for any unit vector $|u\rangle$,*

$$\langle v|u\rangle \simeq_\epsilon \langle v'|u\rangle.$$

Check the grammar in this proposition. -Carl

Proof. We can write $|v'\rangle = |v\rangle + |v''\rangle$, where $\| |v''\rangle \| \leq \epsilon$, then

$$\| \langle v|u\rangle - \langle v'|u\rangle \| = \| \langle v''|u\rangle \| \leq \| |u\rangle \| \| |v''\rangle \| \leq \epsilon.$$

□

Question: Do we actually need the assumption that v, v' have length less than 1? -Carl You are right. We don't. -H.F.

Proposition 2.8. *Let $|u_i\rangle$ and $|v_i\rangle$ for $i = 1, 2$ be vectors with norm less than equal to 1 such that $|u_1\rangle \simeq_{\epsilon_1} |v_1\rangle$ and $|u_2\rangle \simeq_{\epsilon_2} |v_2\rangle$, then*

$$\langle u_1|u_2\rangle \simeq_{\epsilon_1+\epsilon_2} \langle v_1|v_2\rangle.$$

Proof. By direct calculation we have

$$\begin{aligned} \| \langle u_1|u_2\rangle - \langle v_1|v_2\rangle \| \leq & \| \langle u_1|u_2\rangle - \langle v_1|u_2\rangle \| + \| \langle v_1|u_2\rangle - \langle v_1|v_2\rangle \| \\ \leq & \epsilon_1 + \epsilon_2, \end{aligned}$$

where we used Proposition 2.7. □

Proposition 2.9. *Let H be a Hermitian matrix and $|u\rangle$ and $|v\rangle$ be two vectors with norm less or equal to 1 such that*

$$H|u\rangle \simeq_{\epsilon_1} a|u\rangle \qquad\qquad H|v\rangle \simeq_{\epsilon_2} b|v\rangle,$$

for some $a > b \in \mathbb{R}$, then

$$\| \langle u|v\rangle \| \leq \frac{\epsilon_1 + \epsilon_2}{a - b}.$$

Proof. We write $H|u\rangle = a|u\rangle + |u'\rangle$ such that $\| |u'\rangle \| \leq \epsilon_1$. Similarly, we write $H|v\rangle = b|v\rangle + |v'\rangle$ such that $\| |v'\rangle \| \leq \epsilon_2$. Then, we know

$$\begin{aligned} \langle v|H|u\rangle &= a\langle v|u\rangle + \langle v|u'\rangle, \\ \langle v|H|u\rangle &= b\langle v|u\rangle + \langle v'|u\rangle, \end{aligned}$$

Subtracting the second equation from the first equation gives us that

$$\|\langle v|u\rangle\| = \frac{\|\langle v'|u\rangle - \langle v|u'\rangle\|}{a-b} \leq \frac{\||v\rangle\|\||u'\rangle\| + \||u\rangle\|\||v'\rangle\|}{a-b} \leq \frac{\epsilon_1+\epsilon_2}{a-b}.$$

□

Proposition 2.10. *Let $|u\rangle$ and $|v\rangle$ be two vectors such that $\||u\rangle\| \simeq_{\epsilon_1} 1$, $\||v\rangle\| \simeq_{\epsilon_2} 1$, and $\langle u|v\rangle \geq 1-\epsilon_3$, then*

$$|u\rangle \simeq_{O(\sqrt{\epsilon_1+\epsilon_2+\epsilon_3})} |v\rangle.$$

Fix the grammar in the proposition above. -Carl

Proof. Direct calculation gives us that

$$\begin{aligned}\||u\rangle - |v\rangle\|^2 &= \||u\rangle\|^2 + \||v\rangle\|^2 - 2\langle u|v\rangle \\ &\leq (1+\epsilon_1)^2 + (1+\epsilon_2)^2 - 2(1-\epsilon_3) \\ &= 2(\epsilon_1+\epsilon_2+\epsilon_3) + \epsilon_1^2 + \epsilon_2^2 \\ &= O(\epsilon_1+\epsilon_2+\epsilon_3).\end{aligned}$$

□

Proposition 2.11. *Let $|u\rangle$ be a vector with norm less than or equal to 1 such that for a unitary U,*

$$U|u\rangle \simeq_\epsilon a|u\rangle \text{ for } a \in \mathbb{C},$$

then

$$U^n|u\rangle \simeq_{n\epsilon} a^n|u\rangle \text{ for all } n \geq 1.$$

Proof. We can prove it by induction. For $k=1$, the statement is satisfied. Assume it holds for $k=n-1$ and consider $k=n$, then we have

$$\begin{aligned}&\|U^n|u\rangle - a^n|u\rangle\| \\ \leq&\|U\|\|U^{n-1}|\psi\rangle - a^{n-1}|\psi\rangle\| + \|a^{n-1}\|\|U|\psi\rangle - a|\psi\rangle\| \\ \leq&(n-1)\epsilon + \epsilon = n\epsilon,\end{aligned}$$

where we used the fact that $\|U\| = 1$ and $\|a\| \leq 1$. □

3 The linear system game $LS(\Gamma_r)$

In this section, we introduce a linear system game which is a component of the full test. This game is designed to enforce relations that should be satisfied by Alice and Bob's observables. The inspiration comes from Slofstra's seminal work [?], in which he embeds the group

$$K = \langle x, y, a, b : a^2 = b^2 = e, ab = ba, yay^{-1} = a, yby^{-1} = ab, xyx^{-1} = y^2\rangle \tag{9}$$

into a linear system game such that the relations in the definition of K can be derived from the combination of the linear equations. Then he showed the difference between exact representations and approximate representations of K, which implies that the set of quantum correlation is not closed. We generalize the relation $xyx^{-1} = y^2$ and embed the following relation $uou^{-1} = o^r$ for some $r \in \mathbb{N}$ into a linear system game.

3.1 The embedding of $uou^{-1} = o^r$

The main result of this subsection is summarized in the following theorem.

Theorem 3.1. *The relation $uou^{-1} = o^r$ can be embedded in a linear system game which has $n(r) := 18r + 85$ variables and $m(r) := 17r + 77$ equations, where each equation has 3 variables.*

We prove this by constructing the solution group Γ_r of the linear system game, so we denote the linear system game by $LS(\Gamma_r)$. Do you prefer Γ_r or $\Gamma(r)$? Or you have any suggestion about a symbol for this group? -H.F. The notation that you're using seems backwards – I think we should express the solution group as a function of the linear system, not the other way around. -Carl We also prove some properties of the generators of Γ_r along the way.

We embed $uou^{-1} = o^r$ into Γ_r in three steps, which introduce two intermediate groups G_0 and G_1 defined below.

Definition 3.2. *The presentation of G_0 has order-2 generators:*

$$\{o_i\}_{i=1}^r \cup \{u_i\}_{i=1}^5$$

and relations:

$$\begin{aligned} u_3 &= u_2 o_1 u_2 & u_4 &= u_2 o_2 u_2 \\ u_5 &= u_1 u_3 u_1 & o_2 &= u_1 u_4 u_1 \\ u_5 &= o_1 o_r o_1, \end{aligned}$$

and when r is even

$$o_{1+2j} = o_1 o_{2j} o_1 \qquad o_{2+2j} = o_2 o_{1+2j} o_2 \qquad \text{for } j = 1 \ldots r/2 - 1;$$

when r is odd

$$o_3 = o_2 o_1 o_2 \qquad o_{2+2j} = o_1 o_{1+2j} o_1 \qquad o_{3+2j} = o_2 o_{2j+2} o_2 \qquad \text{for } j = 1 \ldots (r-3)/2.$$

Proposition 3.3. *The group G_0 embeds the relation*

$$(u_1 u_2)^j o_1 o_2 (u_2 u_1)^j = (o_1 o_2)^{r^j}, \tag{10}$$

for any $j \in \mathbb{N}$.

The proof follows the proof of Proposition 4.8 of Ref. [?]

Proof. We prove it by induction. Direct calculation gives us that

$$\begin{aligned} u_1 u_2 o_1 o_2 u_2 u_1 =& u_1 (u_2 o_1 u_2)(u_2 o_2 u_2) u_1 \\ =& (u_1 u_3 u_1)(u_1 u_4 u_1) \\ =& u_5 o_2 \\ =& o_1 o_r o_1 o_2. \end{aligned}$$

Then we substitute the definitions of o_k for $k = r \ldots 3$ into the equation above and proves the $j = 1$ case.

Now assume the statement is true for $j = n$, then for $j = n + 1$

$$\begin{aligned}(u_1u_2)^{n+1}o_1o_2(u_2u_1)^{n+1} =& u_1u_2(o_1o_2)^{r^n}u_2u_1 \\ =& (u_1u_2o_1o_2u_2u_1)^{r^n} \\ =& (o_1o_2)^{r^{n+1}},\end{aligned}$$

where from the first line to the second line, we insert $(u_2u_1)(u_1u_2)$ between every pair of o_1o_2. Hence, the statement is true. □

To better characterize all the conjugation relations above and simplify the form of G_1 defined below, we relabel u_i and o_i's as a_i's such that

$$a_j := \begin{cases} o_j \text{ for } j = 1, 2 \\ u_{j-2} \text{ for } j = 3 \dots 7 \\ o_{j-5} \text{ for } j = 8 \dots r + 5. \end{cases} \tag{11}$$

Then we define the set $C(r)$ of tuples (i, j, k) such that

$$(i, j, k) \in C(r) \iff a_ia_ja_i = a_k.$$

The size of $C(r)$ is $r + 3$. In this way, the group G_0 can be presented by $G_0 = \langle \{a_i\}_{i=1}^{r+5} : C(r) \rangle$. Then we define the group G_1.

Definition 3.4. *The presentation of G_1 has order-2 generators:*

$$\{a_i, b_i, c_i, d_i, v_i\}_{i=1}^{r+5}, f, g, \{h_{jk}\}_{(i,j,k)\in C(r)};$$

and relations:

$$\begin{aligned} &a_i = b_ic_i = fd_i = gv_i, && fb_if = c_i && \textit{for } i = 1 \dots r + 5 \\ &h_{jk}b_jc_k = e, && d_ib_jd_i = c_k && \textit{for all } (i, j, k) \in C(r). \end{aligned}$$

Note that the group G_1 is slightly different from the group K used in the proof of Lemma 4.4 of Ref. [**?**]. We introduce the generator g so that later we can associate f with g in linear relations of the Magic Square game. Such relations are important for the self-testing proof. Two important properties of G_1 are given in propositions below. The first property of G_1 was first proved in Lemma 4.4 of [**?**].

Proposition 3.5. *The group G_1 embeds relation $x_ix_jx_i = x_k$ for all $(i, j, k) \in C(r)$.*

Proof. We first show that $d_ic_jd_i = b_k$, which is because

$$d_ic_jd_i = d_i(fb_jf)d_i = f(d_ib_jd_i)f = fc_kf = b_k.$$

Then, we can prove that

$$a_ia_ja_i = fd_ib_jc_jfd_i = (fd_ib_jd_if)(fd_ic_jd_if) = (fc_kf)(fb_kf) = b_kc_k = a_k.$$

□

Proposition 3.6. *In the group G_1, the generators f and g commute with a_i for all $i = 1 \dots r + 5$.*

Proof. We can check that

$$fa_ifa_i = d_id_i = e \qquad\qquad ga_iga_i = v_iv_i = e,$$

for all $i = 1 \dots r+5$. □

Note that there are two types of conjugacy relations in G_1. One is of the form $fb_if = c_i$ for all $i = 1 \dots r+5$. The other one is of the form $d_ib_jd_i = c_k$ for all $(i,j,k) \in C(r)$. For each type of the conjugacy relations, we are going to embed it into a different set of linear relations in Γ_r.

Definition 3.7. *The presentation of Γ_r has order-2 generators:*

$$J, \{a_i, b_i, c_i, d_i, v_i, w_i, \{k_{i,j}\}_{j=1}^{5}\}_{i=1}^{r+5}, \{f_i, g_i, m_i\}_{i=0}^{2}, \{h_{jk}\}_{(i,j,k)\in C(r)}, \{\{k_{l,j}\}_{j=1}^{6}\}_{l\in C(r)},$$

and relations:

- *for all $i = 1 \dots r+5$:*

$$\begin{aligned}
&a_i = b_ic_i = f_0d_i = g_0v_i \\
&g_2a_iw_i = e \\
&f_0f_1f_2 = e \qquad && b_if_2k_{i,1} = e \qquad && k_{i,1}k_{i,2}k_{i,3} = e \\
&f_0k_{i,3}k_{i,4} = e \qquad && c_ik_{i,4}k_{i,5} = e \qquad && f_1k_{i,2}k_{i,5} = e,
\end{aligned}$$

  which embed $f_0b_if_0 = c_i$;

- *for all $l = (i,j,k) \in C(r)$:*

$$\begin{aligned}
&h_{jk}b_jc_k = e \\
&d_ik_{l,1}f_2 = e \qquad && b_jf_2k_{l,2} = e \qquad && k_{l,2}k_{l,3}k_{l,4} = e \\
&d_ik_{l,4}k_{l,5} = e \qquad && c_kk_{l,5}k_{l,6} = e \qquad && k_{l,1}k_{l,3}k_{l,6} = e,
\end{aligned}$$

  which embed $d_ib_jd_i = c_k$;

- *relations of a Magic Square game:*

$$\begin{aligned}
&f_0f_1f_2 = e \qquad && g_0g_1g_2 = e \qquad && m_0m_1m_2 = e \\
&f_0g_2m_0 = e \qquad && f_2g_0m_1 = e \qquad && f_1g_1m_2 = J,
\end{aligned}$$

  which embed $f_0g_0f_0g_0 = J$ and $f_2g_2f_2g_2 = J$;

- *J commutes with all the other generators.*

The special properties of the generators of Γ_r are summarized in the two propositions below.

Proposition 3.8. *The group Γ_r embeds all the conjugacy relations of G_1. It also embeds the relations: $f_0g_0f_0g_0 = J$ and $f_2g_2f_2g_2 = J$.*

The proof of the proposition above follows the proof of Proposition 4.2 of Ref. [**?**].

Proposition 3.9. *In the group Γ_r, the generator f_2 commutes with f_0 and a_i for all $i = 1 \dots r+5$; the generator g_2 commutes with g_0 and a_i for all $i = 1 \dots r+5$.*

Proof. We show the commutativity property of f_2 as an example. The proof for g_2 is very similar. The commutativity between f_0 and f_2 is immediate from the relations. From the relation, we can also deduce that f_2 commutes with b_i for all i. Since $a_i = b_ic_i = b_if_0b_if_0$, we know f_2 commutes with a_i. □

3.2 A representation of Γ_r

In this section we give a representation of Γ_r which is built upon the representation Ψ_0 of G_0 and Ψ_1 of G_1. We need the representation of Γ_r to construct the ideal strategy of $LS(\Gamma_r)$.

Before giving the representations, we define two Hilbert spaces W_{d-1} and W_2, which will form the carrier spaces of the representations,

$$W_{d-1} := \operatorname{span}(\{|x_j\rangle\}_{j=1}^{d-1}) \qquad\qquad W_2 = \operatorname{span}(|x_1\rangle, |x_2\rangle). \tag{12}$$

Note that another form of the basis of W_{d-1} is $\{|x_{r^j}\rangle\}_{j=0}^{d-2}$, where the subscript r^j is taken modulo d implicitly. We use the second form of the basis when it is convenient. We denote the identity operator for these two spaces by $\mathbb{1}_{W_{d-1}}$ and $\mathbb{1}_{W_2}$. On the Hilbert space W_2, we define

$$X_{W_2} = |x_1\rangle\langle x_2| + |x_2\rangle\langle x_1| \qquad Y_{W_2} = i|x_1\rangle\langle x_2| - i|x_2\rangle\langle x_1| \qquad Z_{W_2} = |x_1\rangle\langle x_1| - |x_2\rangle\langle x_2|,$$

Definition 3.10. *The representation, Ψ_0, of G_0 on W_{d-1} is defined by*

$$\begin{aligned}
\Psi_0(a_1) &= \sum_{j=1}^{(d-1)/2} \omega_d^j |x_j\rangle\langle x_{d-j}| + \omega_d^{-j} |x_{d-j}\rangle\langle x_j| \\
\Psi_0(a_2) &= \sum_{j=1}^{d-1} |x_j\rangle\langle x_{d-j}| \\
\Psi_0(a_3) &= |u_0\rangle\langle u_0| + \omega_{d-1}^{(d-1)/2} |u_{(d-1)/2}\rangle\langle u_{(d-1)/2}| \\
&\quad + \sum_{k=1}^{(d-3)/2} \left(\omega_{d-1}^{k} |u_k\rangle\langle u_{d-1-k}| + \omega_{d-1}^{-k} |x_{d-1-k}\rangle\langle x_k|\right) \\
\Psi_0(a_4) &= |u_0\rangle\langle u_0| + |u_{(d-1)/2}\rangle\langle u_{(d-1)/2}| \\
&\quad + \sum_{k=1}^{(d-3)/2} \left(|u_{d-1-k}\rangle\langle u_k| + |u_k\rangle\langle u_{d-1-k}|\right),
\end{aligned}$$

where $\{|u_j\rangle\}_{j=0}^{d-2}$ is a another basis of W_{d-1} given by

$$|u_k\rangle = \frac{1}{\sqrt{d-1}} \sum_{j=0}^{d-2} \omega_{d-1}^{jk} |x_{r^j}\rangle.$$

The representation of the rest of the generators can be constructed from $\Psi_0(a_1)$, $\Psi_0(a_2)$, $\Psi_0(a_3)$ and $\Psi_0(a_4)$ following the conjugation relations.

We can check that

$$\Psi_0(a_1 a_2) = \sum_{j=0}^{d-2} \omega_d^{r^j} |x_{r^j}\rangle\langle x_{r^j}| \qquad\qquad \Psi_0(a_3 a_4) = \sum_{j=0}^{d-2} |x_{r^{j-1}}\rangle\langle x_{r^j}|,$$

which satisfy the relation $\Psi_0(a_3a_4)\Psi_0(a_1a_2)\Psi_0(a_4a_3) = \Psi_0(a_1a_2)^r$ and $[\Psi_0(a_3a_4), \Psi_0(a_2)] = \mathbb{1}_{W_{d-1}}$.

Following the proof of Lemma 4.4 of Ref. **[?]**, we construct a representation of G_1.

Definition 3.11. *The representation Ψ_1 of G_1 on $W_2 \otimes W_{d-1}$ is defined by*

- *for f_0 and g_0*

$$\begin{aligned}\Psi_1(f_0) &= X_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\ \Psi_1(g_0) &= Z_{W_2} \otimes \mathbb{1}_{W_{d-1}},\end{aligned}$$

- *for $i = 1 \dots r+5$,*

$$\begin{aligned}\Psi_1(a_i) &= \mathbb{1}_{W_2} \otimes \Psi_0(a_i) \\ \Psi_1(b_i) &= |x_1\rangle\langle x_1| \otimes \Psi_1(a_i) + |x_2\rangle\langle x_2| \otimes \mathbb{1}_{W_{d-1}} \\ \Psi_1(c_i) &= |x_1\rangle\langle x_1| \otimes \mathbb{1}_{W_{d-1}} + |x_2\rangle\langle x_2| \otimes \Psi_1(a_i) \\ \Psi_1(d_i) &= X_{W_2} \otimes \Psi_0(a_i) \\ \Psi_1(v_i) &= Z_{W_2} \otimes \Psi_0(a_i).\end{aligned}$$

- *for $(i,j,k) \in C(r)$,*

$$\Psi_1(h_{jk}) = |x_1\rangle\langle x_1| \otimes \Psi_1(a_j) + |x_2\rangle\langle x_2| \otimes \Psi_1(a_k).$$

Finally, we follow the proof of Proposition 4.2 of Ref. [**?**] to construct the representation of Γ_r.

Definition 3.12. *The representation Ψ of Γ_r on $W_2 \otimes W_2 \otimes W_{d-1}$ is defined by:*

- *for any x which is also a generator of G_1,*

$$\Psi(x) = \mathbb{1}_{W_2} \otimes \Psi_1(x);$$

- *for the generators involved in the Magic square test*

$$\begin{aligned}\Psi(f_1) &= X_{W_2} \otimes X_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\ \Psi(g_1) &= Z_{W_2} \otimes Z_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\ \Psi(f_2) &= X_{W_2} \otimes \mathbb{1}_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\ \Psi(g_2) &= Z_{W_2} \otimes \mathbb{1}_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\ \Psi(m_0) &= Z_{W_2} \otimes X_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\ \Psi(m_1) &= X_{W_2} \otimes Z_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\ \Psi(m_2) &= Y_{W_2} \otimes Y_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\ \Psi(J) &= -\mathbb{1}_{W_2} \otimes \mathbb{1}_{W_2} \otimes \mathbb{1}_{W_{d-1}};\end{aligned}$$

- *for the generators $\{w_i\}_{i=1}^{r+5}$:*

$$\Psi(w_i) = Z_{W_2} \otimes \mathbb{1}_{W_{d-1}} \otimes \Psi_0(a_i);$$

- *for the generators to embed the relation $fb_if = c_i$:*

$$\begin{aligned}
\Psi(k_{i,1}) &= X_{W_2} \otimes \Psi_1(b_i) = X_{W_2} \otimes (|x_1\rangle\langle x_1| \otimes \Psi_0(a_i) + |x_2\rangle\langle x_2| \otimes \mathbb{1}_{W_{d-1}}) \\
\Psi(k_{i,2}) &= |x_1\rangle\langle x_2| \otimes \Psi_1(b_i f_0) + |x_2\rangle\langle x_1| \otimes \Psi_1(f_0 b_i) \\
&= |x_1x_1\rangle\langle x_2x_2| \otimes \Psi_0(a_i) + (|x_1x_2\rangle\langle x_2x_1| + |x_2x_1\rangle\langle x_1x_2|) \otimes \mathbb{1}_{W_{d-1}} \\
&+ |x_2x_2\rangle\langle x_1x_1| \otimes \Psi_0(a_i) \\
\Psi(k_{i,3}) &= |x_1\rangle\langle x_1| \otimes \Psi_1(b_i f_0 b_i) + |x_2\rangle\langle x_2| \otimes \Psi_1(f_0) \\
&= |x_1\rangle\langle x_1| \otimes X_{W_2} \otimes \Psi_0(a_i) + |x_2\rangle\langle x_2| \otimes X_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\
\Psi(k_{i,4}) &= |x_1\rangle\langle x_1| \otimes \Psi_1(b_i c_i) + |x_2\rangle\langle x_2| \otimes \mathbb{1}_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\
&= |x_1\rangle\langle x_1| \otimes \mathbb{1}_{W_2} \otimes \Psi_0(a_i) + |x_2\rangle\langle x_2| \otimes \mathbb{1}_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\
\Psi(k_{i,5}) &= |x_1\rangle\langle x_1| \otimes \Psi_1(b_i) + |x_2\rangle\langle x_2| \otimes \Psi_1(c_i) \\
&= |x_1x_1\rangle\langle x_1x_1| \otimes \Psi_0(a_i) + (|x_1x_2\rangle\langle x_1x_2| + |x_2x_1\rangle\langle x_2x_1|) \otimes \mathbb{1}_{W_{d-1}} \\
&+ |x_2x_2\rangle\langle x_2x_2| \otimes \Psi_0(a_i);
\end{aligned}$$

- *for the generators to embed the relation $d_ib_jd_i = c_k$ with $l = (i, j, k)$:*

$$\begin{aligned}
\Psi(k_{l,1}) &= X_{W_2} \otimes \Psi_1(d_i) = X_{W_2} \otimes X_{W_2} \otimes \Psi_0(a_i) \\
\Psi(k_{l,2}) &= X_{W_2} \otimes \Psi_1(b_j) = X_{W_2} \otimes (|x_1\rangle\langle x_1| \otimes \Psi_0(a_j) + |x_2\rangle\langle x_2| \otimes \mathbb{1}_{W_{d-1}}) \\
\Psi(k_{l,3}) &= |x_1\rangle\langle x_2| \otimes \Psi_1(b_j d_i) + |x_2\rangle\langle x_1| \otimes \Psi_1(d_i b_j) \\
&= |x_1x_1\rangle\langle x_2x_2| \otimes \Psi_0(a_j a_i) + (|x_1x_2\rangle\langle x_2x_1| + |x_2x_1\rangle\langle x_1x_2|) \otimes \Psi_0(a_i) \\
&+ |x_2x_2\rangle\langle x_1x_1| \otimes \Psi_0(a_i a_j) \\
\Psi(k_{l,4}) &= |x_1\rangle\langle x_1| \otimes \Psi_1(b_j d_i b_j) + |x_2\rangle\langle x_2| \otimes \Psi_1(d_i) \\
&= |x_1\rangle\langle x_1| \otimes (|x_2\rangle\langle x_1| \otimes \Psi_0(a_i a_j) + |x_1\rangle\langle x_2| \otimes \Psi_0(a_j a_i)) + |x_2\rangle\langle x_2| \otimes X_{W_2} \otimes \Psi_0(a_i) \\
\Psi(k_{l,5}) &= |x_1\rangle\langle x_1| \otimes \Psi_1(b_j c_k) + |x_2\rangle\langle x_2| \otimes \mathbb{1}_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\
&= |x_1x_1\rangle\langle x_1x_1| \otimes \Psi_0(a_j) + |x_1x_2\rangle\langle x_1x_2| \otimes \Psi_0(a_k) + |x_2\rangle\langle x_2| \otimes \mathbb{1}_{W_2} \otimes \mathbb{1}_{W_{d-1}} \\
\Psi(k_{l,6}) &= |x_1\rangle\langle x_1| \otimes \Psi_1(b_j) + |x_2\rangle\langle x_2| \otimes \Psi_1(c_k) \\
&= |x_1x_1\rangle\langle x_1x_1| \otimes \Psi_0(a_j) + (|x_1x_2\rangle\langle x_1x_2| + |x_2x_1\rangle\langle x_2x_1|) \otimes \mathbb{1}_{W_{d-1}} \\
&+ |x_2x_2\rangle\langle x_2x_2| \otimes \Psi_1(a_k).
\end{aligned}$$

The validation of Ψ_1 and Ψ are done in Slofstra's work [?].

4 Robust self-testing of $|\Sigma^{(2)}\rangle$ *via* $\mathcal{I}_\alpha$

In this section, we show how the weighted CHSH inequality can be used to self-test $|\Sigma^{(2)}\rangle$ robustly. Intuitively, it means that if some quantum strategy achieves a value $\langle \mathcal{I}_\alpha \rangle$ that is close to $\langle \mathcal{I}_\alpha \rangle_{\max}$, then the shared state must be close to $|\Sigma^{(2)}\rangle$ and their observables are close to the rotated Pauli operators defined in Definition 2.6, up to some local isometry. The formal statement is given in the theorem below. Since the proof uses the same techniques as introduced in Ref. [?], we defer the proof till Appendix A. Instead, we only list the key relations used in the proof because they will be reused in Section 6 where we prove our robust self-testing result.

Theorem 4.1. *Suppose the quantum strategy* $(|\psi\rangle, \{M_x := M_x^0 - M_x^1\}_{x=1,2}, \{N_y := N_y^0 - N_y^1\}_{y=1,2})$ *achieves that* $\langle \mathcal{I}_\alpha \rangle \geq 2\sqrt{1+\alpha^2} - \epsilon$ *for some* $\epsilon > 0$*, then we define* $\mu := \arctan(1/\alpha)$ *and there exists a local isometry* $\Phi = \Phi_A \otimes \Phi_B$ *and an auxiliary state* $|aux\rangle$ *such that*

$$\|\Phi(M_x \otimes N_y|\psi\rangle) - |aux\rangle \otimes (\tilde{M}_x \otimes \tilde{N}_y)|\Sigma^{(2)}\rangle\| = O(\frac{1}{\cos^2(\mu)\sin^{1/2}(\mu)} + \frac{1}{\sin(\mu)^{3/2}}))\sqrt{\epsilon}$$

for $x, y \in \{0,1,2\}$ *where the subscript* 0 *refers to the identity operator and* $\tilde{M}_x, \tilde{N}_y$ *are defined in Definition 2.6. You're using both "aux" and "junk" for the auxilliary state. Also you need to be clearer about what label you are using for Alice's and Bob's systems (see my previous comment on this). -Carl*

We follow the convention and define $c := \cos(\mu)$, $s := \sin(\mu)$, and

$$Z_A := M_1 \qquad\qquad X_A := M_2$$
$$Z_B := \frac{N_1 + N_2}{2c} \qquad\qquad X_B := \frac{N_1 - N_2}{2s}.$$

The key relations that we will reuse are

$$Z_A|\psi\rangle \simeq_{\frac{\sqrt{s\epsilon}}{c}} Z_B|\psi\rangle, \tag{13}$$
$$X_A|\psi\rangle \simeq_{\frac{\sqrt{\epsilon}}{\sqrt{s}}} X_B|\psi\rangle, \tag{14}$$
$$X_A(\mathbb{1} + Z_B)|\psi\rangle \simeq_{\frac{c+1}{c\sqrt{s}}\sqrt{\epsilon}} X_B(\mathbb{1} - Z_A)|\psi\rangle, \tag{15}$$
$$Z_A(\mathbb{1} + X_B)|\psi\rangle \simeq_{\frac{s+1}{c\sqrt{s}}\sqrt{\epsilon}} Z_B(\mathbb{1} - X_A)|\psi\rangle, \tag{16}$$
$$Z_A X_A|\psi\rangle \simeq_{\frac{1+c+s}{c\sqrt{s}}\sqrt{\epsilon}} -X_A Z_A|\psi\rangle, \tag{17}$$
$$X_A Z_A|\psi\rangle \simeq_{\frac{2s^2+cs+c+s}{cs\sqrt{s}}\sqrt{\epsilon}} -X_B Z_B|\psi\rangle, \tag{18}$$

where eq. (18) can be derived from eq. (13), eq. (14) and eq. (17).

5 The full test

The full test consists of three subtest: the linear system game $LS(\Gamma_r)$, the extended weighted CHSH test and a commutation test. We explain how the three tests are conducted and combined below.

Recall that the the linear system game with solution group Γ_r defined in Definition 3.7 has variables labelled by $j \in \{1, 2, \dots, n(r)\}$ and equations labelled by $i \in \{1, 2 \dots m(r)\}$. We denote the associated linear system by $H\underline{x} = \underline{c}$ and define $X_i := \{\underline{x} : \sum_k H(i,k)\underline{x}(k) = \underline{c}(i)\}$ and $I_i = \{k : H(i,k) = 1\}$. From the construction of Γ_r, we know that $|I_i| \leq 3$. The correspondence between $\underline{x}$ and the generators of Γ is that

$$\underline{x}(j) := a_j \qquad\qquad\qquad \text{for } j = 1 \dots r+5$$
$$\underline{x}(r+6+j) := f_j \qquad \underline{x}(r+9+j) := g_j \qquad \text{for } j = 0,1,2$$

The correspondence can be arbitrary for $j > r + 11$. We conduct the linear constraint test in a way such that Bob is always get a variable label but Alice gets either an equation label, in which case she needs to give the full assignment to this equation, or a pair of equation label and variable label, in which case she only needs to give the assignment to the particular variable in the given equation.

In this way, we force Alice and Bob to use an binary observable strategy for the linear constraint test.

In the extended weighted CHSH test, all the question pairs are $\{(*,*)\} \cup \{(*,y),(x,*)\}_{x,y=n(r)+1,n(r)+2} \cup \{(x,y)\}_{x=y\in\{n(r)+1,n(r)+2\}} \cup \{(x,y)\}_{x=y\in\{1,2\}} \cup \{(x,y)\}_{x=1,2,y=n(r)+1,n(r)+2} \cup \{(x,y)\}_{x=n(r)+1,n(r)+2,y=1,2}$. For each question, there is an associated answer set:

- when the question is $*$, the answer set is $\{\diamond, \perp\}$;
- when the question is 1 or 2, the answer set is $\{0,1\}$;
- when the question is $n(r)+1$ or $n(r)+2$, the answer set is $\{0,1,\perp\}$.

Note that when the referee wants to ask Alice $x = 1,2$, which also labels a variable of $LS(\Gamma_r)$, the referee gives Alice a pair $(i_x.x)$ where i_x labels an equation containing the variable labelled by x. For simplicity we omit the i_x part in the alphabet $\mathcal{X}$, since $LS(\Gamma_r)$ forces Alice to use a binary observable strategy. Later we will show that the i_x part can be made implicit. The test is conducted in the following way:

1. if $x = y = *$, Alice and Bob should answer with $a = b \in \{\diamond, \perp\}$;
2. if $x = y \in \{1,2\}$, Alice and Bob should answer with $a = b \in \{0,1\}$, as in the linear constraint test;
3. if $x = y \in \{n+1, n+2\}$, Alice and Bob should answer with $a = b \in \{0,1,\perp\}$;
4. if $x \in \{1,2\}$ and $y \in \{n+1,n+2\}$,
   (a) if they answer with $a,b \in \{0,1\}$, then they should maximize $\langle \mathcal{I}_{\cot(-\pi/d)} \rangle$ with Alice and Bob's roles flipped,
   (b) if Bob answers with $\perp$, then Alice's answer doesn't matter;
5. if $x \in \{n+1,n+2\}$ and $y \in \{1,2\}$,
   (a) if they answer with $a,b \in \{0,1\}$, then they should maximize $\langle \mathcal{I}_{\cot(-\pi/d)} \rangle$,
   (b) if Alice answers with $\perp$, then Bob's answer doesn't matter;
6. $x \in \{n+1,n+2\}, y = *$
   (a) if Bob answers with $\diamond$, Alice should answer with $\{0.1\}$ but not $\perp$,
   (b) if Bob answers with $\perp$ Alice should answer with $\perp$ too;
7. $x = *, y \in \{n+1,n+2\}$
   (a) if Alice answers with $\diamond$, Bob should answer with $\{0.1\}$ but not $\perp$,
   (b) if Alice answers $\perp$, Bob should answer $\perp$ too.

In the commutation test, we reuse questions from the two other tests so the answer set for each question is the same across the tests. There are 8 test cases. In each test case, Alice is given one of two questions. One of the possible question is to assign a value to some variable of $LS(\Gamma_r)$. Again, the equation label is implicit when this variable is given. The other possibility is to answer $n(r)+1$ or $n(r)+2$ from the extended weighted CHSH test. Bob is given both of the question so he needs to give answers to both questions. Their answers should agree on the question that is given to both.

The $\mathcal{Y}$ has 8 pairs, $\mathcal{Y} = \{(n(r)+1, j), (n(r)+2, j)\}_{j=r+6,r+8,r+9,r+11}$. Alice's input is determined by the pair so she receives one of the two questions of a given pair with equal probability.

The full correlation, denoted by $\mathfrak{C}(d, r)$, is generated by the ideal strategy that can pass all the three tests. We give the ideal strategy first.

5.1 The ideal strategy

Recall the two subspaces W_{d-1} and W_2 defined in eq. (12). In the ideal strategy Alice and Bob share the state

$$\begin{aligned}|\psi\rangle =& \frac{1}{2}(|x_1\rangle|x_1\rangle + |x_2\rangle|x_2\rangle)^{\otimes 2} \otimes \left(\frac{1}{\sqrt{d-1}}\sum_{j=1}^{d-1}|x_j\rangle|x_{d-j}\rangle\right)\\ =& \frac{1}{2}(|x_1\rangle|x_1\rangle + |x_2\rangle|x_2\rangle)^{\otimes 2} \otimes \left(1/\sqrt{d-1}\sum_{j=1}^{(d-1)/2} e^{ij\pi/d}(|j\rangle|j\rangle + |d-j\rangle|d-j\rangle)\right).\end{aligned}$$

Here we introduce the second basis $\{|j\rangle\}_{j=1}^{d-1}$ of W_{d-1} so that the observables used in the extended weighted CHSH test have a simpler form. The two basis are related by

$$|x_j\rangle = \frac{-1}{\sqrt{2}}(|j\rangle + i|d-j\rangle), \qquad |x_{d-j}\rangle = \frac{-e^{ij\pi/d}}{\sqrt{2}}(|j\rangle - i|d-j\rangle) \qquad \text{for } j = 1\ldots\frac{d-1}{2}.$$

We define two subspaces $V = \text{span}\{|1\rangle, |d-1\rangle\}$ and $V^{\perp} = W_{d-1} \setminus \text{span}\{|1\rangle, |d-1\rangle\}$ on W_{d-1} and define Π_V and $\Pi_{V^{\perp}}$ to be the corresponding projectors. Note that V is the subspace on which they should maximize $\langle \mathcal{I}_{\cot(-\pi/d)}\rangle$. We define the states for the measurements used in the extended weighted CHSH test as follows

$$\begin{aligned}|1_+\rangle &= \cos(-\frac{\pi}{2d})|1\rangle + \sin(-\frac{\pi}{2d})|d-1\rangle & |(d-1)_+\rangle &= \sin(-\frac{\pi}{2d})|1\rangle - \cos(-\frac{\pi}{2d})|d-1\rangle\\ |1_-\rangle &= \cos(-\frac{\pi}{2d})|1\rangle - \sin(-\frac{\pi}{2d})|d-1\rangle & |(d-1)_-\rangle &= \sin(-\frac{\pi}{2d})|1\rangle + \cos(-\frac{\pi}{2d})|d-1\rangle\\ |1_\times\rangle &= \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|d-1\rangle & |(d-1)_\times\rangle &= \frac{1}{\sqrt{2}}|1\rangle - \frac{1}{\sqrt{2}}|d-1\rangle.\end{aligned}$$

Then the measurements are

$$\begin{aligned}M_*^{\diamond} &= (\mathbb{1}_{W_2})^{\otimes 2} \otimes \Pi_V & M_*^{\perp} &= M_{n+1}^{\perp} = M_{n+2}^{\perp} = (\mathbb{1}_{W_2})^{\otimes 2} \otimes \Pi_{V^{\perp}}\\ N_*^{\diamond} &= (\mathbb{1}_{W_2})^{\otimes 2} \otimes \Pi_V & N_*^{\perp} &= N_{n+1}^{\perp} = N_{n+2}^{\perp} = (\mathbb{1}_{W_2})^{\otimes 2} \otimes \Pi_{V^{\perp}}\\ M_{n+1}^0 &= N_{n+1}^0 = (\mathbb{1}_{W_2})^{\otimes 2} \otimes |1\rangle\langle 1| & M_{n+1}^1 &= N_{n+1}^1 = (\mathbb{1}_{W_2})^{\otimes 2} \otimes |d-1\rangle\langle d-1|\\ M_{n+2}^0 &= N_{n+2}^0 = (\mathbb{1}_{W_2})^{\otimes 2} \otimes |1_\times\rangle\langle 1_\times| & M_{n+2}^1 &= N_{n+2}^1 = (\mathbb{1}_{W_2})^{\otimes 2} \otimes |(d-1)_\times\rangle\langle (d-1)_\times|\\ M_1^0 &= N_1^0 = (\mathbb{1}_{W_2})^{\otimes 2} \otimes |1_+\rangle\langle 1_+| & M_1^1 &= N_1^1 = (\mathbb{1}_{W_2})^{\otimes 2} \otimes |(d-1)_+\rangle\langle (d-1)_+|\\ M_2^0 &= N_2^0 = (\mathbb{1}_{W_2})^{\otimes 2} \otimes |1_-\rangle\langle 1_-| & M_1^1 &= N_1^1 = (\mathbb{1}_{W_2})^{\otimes 2} \otimes |(d-1)_-\rangle\langle (d-1)_-|.\end{aligned}$$

The observables for $LS(\Gamma_r)$ are

$$M_i = N_i = \Psi(x_i), \tag{19}$$

where the representation Ψ is defined in Definition 3.12. Note that Alice uses the same observable for the same variable appearing in different equations.

To make sure the strategy is correct, we need to check a few properties of it.

Proposition 5.1. *Let $M_1 = N_1 := \Psi(a_1)$, $M_2 = N_2 := \Psi(a_2)$, $M_3 = N_3 := \Psi(a_3)$ and $M_4 = N_4 := \Psi(a_4)$, then*

$$
\begin{aligned}
&M_iN_i|\psi\rangle = |\psi\rangle && \textit{for } i = 1\dots 4, && (20)\\
&(M_3M_4)M_2(M_4M_3) = M_2, && && (21)
\end{aligned}
$$

and similar commutation relation holds on Bob's side too.

Proof. Since all the operators involved in this proposition act trivially on $W_2^{\otimes 2}$, we focus on their actions on W_{d-1}. In the later proof, by $|\psi\rangle$ we mean the part of $|\psi\rangle$ on W_{d-1}.

We prove eq. (21) first. It is simple to check that

$$
M_1M_2 = \sum_{j=1}^{d-1} \omega_d^j |x_j\rangle\langle x_j|, \qquad M_3M_4 = \sum_{j=0}^{d-2} |x_{r^{j-1}}\rangle\langle x_{r^j}|,
$$

and that

$$
\begin{aligned}
M_3M_4M_2M_4M_3 &= \sum_{j=0}^{(d-3)/2} M_3M_4(|x_{r^j}\rangle\langle x_{r^{d-1-j}}| + |x_{r^{d-1-j}}\rangle\langle x_{r^j}|)M_4M_3\\
&= \sum_{j=0}^{(d-3)/2} |x_{r^{j-1}}\rangle\langle x_{r^{d-2-j}}| + |x_{r^{d-2-j}}\rangle\langle x_{r^{j-1}}| = M_2.
\end{aligned}
$$

Next we prove eq. (20). It is straightforward to see that $M_iN_i|\psi\rangle = |\psi\rangle$ for $i = 1, 2$. Before checking $i = 3$ and $i = 4$, we first observe that

$$
M_3M_4N_3N_4|\psi\rangle = |\psi\rangle.
$$

The eigenvectors of M_3M_4 and N_3N_4 are

$$
|u_k\rangle = \frac{1}{\sqrt{d-1}} \sum_{j=0}^{d-2} \omega_{d-1}^{jk} |x_{r^j}\rangle.
$$

Hence we can express $|x_{r^j}\rangle$ in terms of $|u_k\rangle$ as

$$
|x_{r^j}\rangle = \frac{1}{\sqrt{d-1}} \sum_{k=0}^{d-2} \omega_{d-1}^{-jk} |u_k\rangle.
$$

Then M_4 acts on $|x_{r^j}\rangle$ by

$$
\begin{aligned}
M_4|x_{r^j}\rangle =& \frac{1}{\sqrt{d-1}} M_4 \left(|u_0\rangle + \omega_{d-1}^{-j(d-1)/2} |u_{(d-1)/2}\rangle + \sum_{k=1}^{(d-3)/2} (\omega_{d-1}^{-jk}|u_k\rangle + \omega_{d-1}^{jk}|u_{d-1-k}\rangle) \right)\\
=& \frac{1}{\sqrt{d-1}} \left(|u_0\rangle + (-1)^j |u_{(d-1)/2}\rangle + \sum_{k=1}^{(d-3)/2} (\omega_{d-1}^{-jk}|u_{d-1-k}\rangle + \omega_{d-1}^{jk}|u_k\rangle) \right)\\
=& |x_{r^{d-1-j}}\rangle.
\end{aligned}
$$

Then we show that $M_4N_4|\psi\rangle = |\psi\rangle$ as

$$
M_4N_4|\psi\rangle = \frac{1}{\sqrt{d-1}} \sum_{j=0}^{d-2} M_4N_4 |x_{r^j}\rangle|x_{r^{d-1-j}}\rangle = \frac{1}{\sqrt{d-1}} \sum_{j=0}^{d-2} |x_{r^{d-1-j}}\rangle|x_{r^j}\rangle = |\psi\rangle,
$$

Combining it with the fact that $M_3M_4N_3N_4|\psi\rangle = |\psi\rangle$, we can conclude that $M_3N_3|\psi\rangle = |\psi\rangle$. □

The similar result, i.e., $M_jN_j|\psi\rangle = |\psi\rangle$ for $j = 1 \dots n(r)$, are easy to check because all the other observables are defined by M_j, N_j for $j = 1,2,3,4$. An interesting property of $U_A := M_3M_4$ and $O_A := M_1M_2$ is summarized in the proposition below.

Lemma 5.2. *The unitaries O and U of the form*

$$O = \sum_{j=1}^{d-1} \omega_d^j |x_j\rangle\langle x_j| \qquad\qquad U = \sum_{j=1}^{d-1} |x_{jr^{-1}}\rangle\langle x_j|, \tag{22}$$

where d is an odd prime number with primitive root r, generates the ring of $(d-1)\times(d-1)$ matrices over $\mathbb{C}$.

The proof of this lemma is provided in Appendix B. Our self-test argument certifies that Alice and Bob has such O and U in their hands, so it provides another angle to justify our self-testing argument.

5.2 The correlation $\mathfrak{C}(d,r)$

The full correlation is generated by the ideal strategy presented in the previous subsection. To help readers understanding the proofs in Section 6 of this paper better, parts of the full correlation is given in the tables below.

		$x=*$	
		$a=\diamond$	$a=\perp$
$y=*$	$b=\diamond$	2/(d-1)	0
	$b=\perp$	0	(d-3)/(d-1)

Table 1: The correlation for $x = y = *$.

		$x=n(r)+1$			$x=n(r)+2$		
		$a=0$	$a=1$	$a=\perp$	$a=0$	$a=1$	$a=\perp$
$y=1$	$b=0$	$\frac{\cos^2(\pi/2d)}{d-1}$	$\frac{\sin^2(\pi/2d)}{d-1}$	$P(\perp 0\|00)$	$\frac{1+\sin(\pi/d)}{2(d-1)}$	$\frac{1-\sin(\pi/d)}{2(d-1)}$	$P(\perp 0\|10)$
	$b=1$	$\frac{\sin^2(\pi/2d)}{d-1}$	$\frac{\cos^2(\pi/2d)}{d-1}$	$\frac{d-3}{d-1} - P(\perp 0\|00)$	$\frac{1-\sin(\pi/d)}{2(d-1)}$	$\frac{1+\sin(\pi/d)}{2(d-1)}$	$\frac{d-3}{d-1} - P(\perp 0\|10)$
$y=2$	$b=0$	$\frac{\cos^2(\pi/2d)}{d-1}$	$\frac{\sin^2(\pi/2d)}{d-1}$	$P(\perp 0\|01)$	$\frac{1-\sin(\pi/d)}{2(d-1)}$	$\frac{1+\sin(\pi/d)}{2(d-1)}$	$P(\perp 0\|11)$
	$b=1$	$\frac{\sin^2(\pi/2d)}{d-1}$	$\frac{\cos^2(\pi/2d)}{d-1}$	$\frac{d-3}{d-1} - P(\perp 0\|01)$	$\frac{1+\sin(\pi/d)}{2(d-1)}$	$\frac{1-\sin(\pi/d)}{2(d-1)}$	$\frac{d-3}{d-1} - P(\perp 0\|11)$

Table 2: The correlation for $x \in \{n+1, n+2\}$ and $y \in \{1,2\}$.

We don't explicitly calculate the conditional probabilities of the form $P(\perp 0|xy)$ for all possible x,y because they are irrelevant. Note that when $x \in \{1,2\}$ and $y \in \{n(r)+1, n(r)+2\}$, the correlation table is the transpose of Table 2, so we omit it here.

		$x = n(r)+1$			$x = n(r)+2$			$x = *$	
		$a = 0$	$a = 1$	$a = \perp$	$a = 0$	$a = 1$	$a = \perp$	$a = \diamond$	$a = \perp$
$y = n(r)+1$	$b = 0$	$\frac{1}{d-1}$	0	0	$\frac{1}{2d-2}$	$\frac{1}{2d-2}$	0	$\frac{1}{d-1}$	0
	$b = 1$	0	$\frac{1}{d-1}$	0	$\frac{1}{2d-2}$	$\frac{1}{2d-2}$	0	$\frac{1}{d-1}$	0
	$b = \perp$	0	0	$\frac{d-3}{d-1}$	0	0	$\frac{d-3}{d-1}$	0	$\frac{d-3}{d-1}$
$y = n(r)+2$	$b = 0$	$\frac{1}{2d-2}$	$\frac{1}{2d-2}$	0	$\frac{1}{d-1}$	0	0	$\frac{1}{d-1}$	0
	$b = 1$	$\frac{1}{2d-2}$	$\frac{1}{2d-2}$	0	0	$\frac{1}{d-1}$	0	$\frac{1}{d-1}$	0
	$b = \perp$	0	0	$\frac{d-3}{d-1}$	0	0	$\frac{d-3}{d-1}$	0	$\frac{d-3}{d-1}$
$y = *$	$b = \diamond$	$\frac{1}{d-1}$	$\frac{1}{d-1}$	0	$\frac{1}{d-1}$	$\frac{1}{d-1}$	0	$\frac{2}{d-1}$	0
	$b = \perp$	0	0	$\frac{d-3}{d-1}$	0	0	$\frac{d-3}{d-1}$	0	$\frac{d-3}{d-1}$

Table 3: The correlation for $x, y \in \{n(r)+1, n(r)+2, *\}$.

		$y = (n(r)+1, r+6)$					
		$b = 00$	$b = 01$	$b = 10$	$b = 11$	$b = \perp 0$	$b = \perp 1$
$x = n(r)+1$	$a = 0$	$\frac{1}{2d-2}$	$\frac{1}{2d-2}$	0	0	0	0
	$a = 1$	0	0	$\frac{1}{2d-2}$	$\frac{1}{2d-2}$	0	0
	$b = \perp$	0	0	0	0	$\frac{d-3}{2d-2}$	$\frac{d-3}{2d-2}$
$x = r+6$	$a = 0$	$\frac{1}{2d-2}$	0	$\frac{1}{2d-2}$	0	$\frac{d-3}{2d-2}$	0
	$a = 1$	0	$\frac{1}{2d-2}$	0	$\frac{1}{2d-2}$	0	$\frac{d-3}{2d-2}$

Table 4: The correlation for the commutation test between M_{r+6} and $M_{n(r)+1}$.

The correlation table for the other 7 commutation tests are similar.

6 Robust self-testing with $\mathfrak{C}(d, r)$

In this section, we examine the properties of a strategy S that can reproduce the correlation $\mathfrak{C}(d, r)$ approximately and prove that the strategy S must contain the state $|\Sigma^{(2)}\rangle^{\otimes 2} \otimes |\Sigma^{(d-1)}\rangle$ up to some local isometry.

Let us denote the full correlation by $\mathfrak{C}(d, r) = (\tilde{P}(ab|xy))$, then the formal statement of an approximate strategy is given below.

Definition 6.1. *For a correlation $\tilde{C} = (\tilde{P}(ab|xy))$, we say a quantum strategy $S = (|\psi\rangle, \{M_x^a\}, \{N_y^b\})$ is an ϵ-approximate quantum strategy of $\tilde{C}$, if it produces the correlation $(P(ab|xy) = \langle\psi|M_x^a N_y^b|\psi\rangle)$ such that*

$$E_{(x,y)} \sum_{ab} |P(ab|xy) - \tilde{P}(ab|xy)| \le \epsilon. \tag{23}$$

In our full test, the probability distribution over all the question pairs (x, y) is uniform. Since the number of all the possible question pairs is $\Theta(r^2)$ and the number of possible output pairs is constant for each question pair (x, y), the approximate condition directly implies that, if S is an ϵ-approximate strategy of $\mathfrak{C}(d, r)$, then $P(ab|xy) \simeq_{O(r^2\epsilon)} \tilde{P}(ab|xy)$, for all (a, b, x, y).

In the next three subsections, we analyze the implication of an approximate strategy to pass the three subtests given in Section 5.

6.1 Implication of passing $LS(\Gamma_r)$

In this subsection, we use (i,j) to denote the question pair so that Alice gets equation number i, sometimes with a variable label j, and Bob gets a variable label j. The way we conduct the linear constraint test forces Alice to have binary observables $\{M_{ij}\}$ that satisfy Definition 2.4.

Proposition 6.2. *Let S be an ϵ-approximate strategy of $\mathfrak{C}(d,r)$, then there exist binary observables $\{M_j\}_{j=1}^{n(r)}$, such that*

$$M_j N_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} |\psi\rangle \tag{24}$$

$$\Pi_{j\in I_i} M_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} (-1)^{\underline{c}(i)}|\psi\rangle. \tag{25}$$

Proof. The condition that $\tilde{P}(\underline{x}b|ij)$ from $\mathfrak{C}(d,r)$ wins the linear system game optimally is equivalent to

$$\sum_{\underline{x},b:\underline{x}\in X_i,\underline{x}(j)=b} \tilde{P}(\underline{x}b|ij) = 1$$

for all $i = 1\ldots m(r)$. It has been shown in Ref. [**?**] that

$$\langle\psi|M_{ij}N_j|\psi\rangle = 2\sum_{\underline{x},b:\underline{x}\in X_i,\underline{x}(j)=b} P(\underline{x}b|ij) - 1 \simeq_{O(r^2\epsilon)} 1.$$

Since M_{ij} and N_j are unitaries, we know that

$$M_{ij}|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N_j|\psi\rangle \quad \text{for all } i \leq m(r) \text{ and } j \leq n(r).$$

We define $M_j := M_{ij}$ for the smallest i such that $j \in I_i$ then

$$M_j N_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_{ij}N_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} |\psi\rangle \quad \text{for all } j \leq n(r).$$

The definition of M_{ij} tells us that $\langle\psi|\Pi_{j\in I_i}M_{ij}|\psi\rangle = (-1)^{\underline{c}(i)}$ or equivalently,

$$\Pi_{j\in I_i}M_{ij}|\psi\rangle = (-1)^{b_i}|\psi\rangle \quad \text{for all } i \leq m(r).$$

If we replace each M_{ij} with M_j in the equation above, we demonstrate the effect on a general case that $I_i = \{j_1, j_2, j_3\}$,

$$\begin{aligned}
&\|M_{ij_1}M_{ij_2}M_{ij_3}|\psi\rangle - M_{j_1}M_{j_2}M_{j_3}|\psi\rangle\| \\
\leq&\|M_{ij_1}M_{ij_2}M_{ij_3}|\psi\rangle - M_{j_1}M_{ij_2}M_{ij_3}|\psi\rangle\| + \|M_{j_1}M_{ij_2}M_{ij_3}|\psi\rangle - M_{j_1}M_{j_2}M_{ij_3}|\psi\rangle\| + \\
&\|M_{j_1}M_{j_2}M_{ij_3}|\psi\rangle - M_{j_1}M_{j_2}M_{j_3}|\psi\rangle\| \\
\leq&3O(r\sqrt{\epsilon}) = O(r\sqrt{\epsilon}).
\end{aligned}$$

Hence we can conclude that

$$\Pi_{j\in I_i}M_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} (-1)^{\underline{c}(i)}|\psi\rangle \quad \text{for all } i \leq m(r).$$

□

Note that with the same technique, we can prove that

$$\Pi_{j\in I_i} N_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} (-1)^{\underline{c}(i)}|\psi\rangle \quad \text{for all } i \leq m(r).$$

Following the steps of the embedding of $(a_3a_4)^l a_1a_2(a_4a_3)^l = (a_1a_2)^{r^l}$ into Γ_r, we can get the following proposition.

Proposition 6.3. *Let S be an ϵ-approximate strategy of $\mathfrak{C}(d,r)$ and $\{M_j, N_j\}_{j\in n(r)}$ be the observables defined in Proposition 6.2, then*

$$M_1M_2(M_4M_3)^l|\psi\rangle \simeq_{O(r^{l+1}\sqrt{\epsilon})} (M_4M_3)^l(M_1M_2)^{r^l}|\psi\rangle, \tag{26}$$

$$N_1N_2(N_4N_3)^l|\psi\rangle \simeq_{O(r^{l+1}\sqrt{\epsilon})} (N_4N_3)^l(N_1N_2)^{r^l}|\psi\rangle, \tag{27}$$

for all $l \geq 1$.

Proof. The $l = 1$ case follows from the fact that the linear system game embeds $O(r)$ conjugacy relations and all the $O(r)$ conjugacy relations are used in the derivation of $(a_3a_4)a_1a_2(a_4a_3) = (a_1a_2)^r$, so we have

$$M_1M_2(M_4M_3)|\psi\rangle \simeq_{O(r^2\sqrt{\epsilon})} (M_4M_3)(M_1M_2)^r|\psi\rangle. \tag{28}$$

Assume the statement is true for $l = n$ and consider the case that $l = n + 1$,

$$\begin{aligned} M_1M_2(M_4M_3)^{n+1}|\psi\rangle &\simeq_{O(r\sqrt{\epsilon})} (N_3N_4)M_1M_2(M_4M_3)^n|\psi\rangle \\ &\simeq_{O(r^{n+1}\sqrt{\epsilon})} (N_3N_4)(M_4M_3)^n(M_1M_2)^{r^n}|\psi\rangle \\ &\simeq_{O(r\sqrt{\epsilon})} (M_4M_3)^n(M_1M_2)^{r^n}(M_4M_3)|\psi\rangle \\ &\simeq_{O(r^{n+1}\sqrt{\epsilon})} (M_4M_3)^{n+1}(M_1M_2)^{r^{n+1}}|\psi\rangle, \end{aligned}$$

where from the second to the last line we repeat the process of applying eq. (28) and the relation $M_1M_2|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N_1N_2|\psi\rangle$, r^n times. In summary, we have

$$M_1M_2(M_4M_3)^{n+1}|\psi\rangle \simeq_{O(r^{n+2}\sqrt{\epsilon})} (M_4M_3)^{n+1}(M_1M_2)^{r^{n+1}}|\psi\rangle, \tag{29}$$

because $O(r^{n+1}\sqrt{\epsilon}) + O(r^{n+1}\sqrt{\epsilon}) = O(r^{n+2}\sqrt{\epsilon})$. By the principal of proof by induction, the proof is complete. The proof for Bob's observables $\{N_j\}_{j=1}^{n(r)}$ is similar. □

Recall $O_A = M_1M_2$ and $U_A = M_3M_4$, then the conclusion of Proposition 6.3 can be rewritten as

$$U_A^n O_A (U_A^\dagger)^n|\psi\rangle \simeq_{O(r^{n+1}\sqrt{\epsilon})} O_A^{r^n}|\psi\rangle.$$

On Bob's side, we also define $O_B = N_1N_2$ and $U_B = N_3N_4$ and similar relations hold.

The other relations that are embedded in Γ_r are

$$[f_0, a_i] = [f_2, a_i] = [g_0, a_i] = [g_2, a_i] = e \text{ for all } i = 1 \ldots r+5 \tag{30}$$

$$f_0g_0 = Jg_0f_0 \tag{31}$$

$$f_2g_2 = Jg_2f_2. \tag{32}$$

We summarize the implications in the following proposition.

Proposition 6.4. *Let S, $\{M_j, N_j\}_{j=1}^{n(r)}$ be the strategy and observables defined in Proposition 6.2, then*

$$M_{r+6}M_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_jM_{r+6}|\psi\rangle \qquad M_{r+8}M_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_jM_{r+8}|\psi\rangle \tag{33}$$

$$M_{r+9}M_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_jM_{r+9}|\psi\rangle \qquad M_{r+11}M_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_jM_{r+11}|\psi\rangle \tag{34}$$

for all $j = 1 \dots r+5$, and

$$M_{r+6}M_{r+9}|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} -M_{r+9}M_{r+6}|\psi\rangle \tag{35}$$

$$M_{r+8}M_{r+11}|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} -M_{r+11}M_{r+8}|\psi\rangle. \tag{36}$$

Recall all the relations in eqs. (30) to (32) are embedded in constant number of linear relations in Γ_r, so this proposition can be proved using the same technique used in the proof of Proposition 6.2.

6.2 Implication of passing the extended weighted CHSH test

Proposition 6.5. *Let S be an ϵ-approximate quantum strategy, then the strategy S satisfy the following relations*

$$M_*^\diamond N_*^\diamond|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_*^\diamond|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} (M_{n(r)+1}^0 + M_{n(r)+1}^1)|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} (M_{n(r)+2}^0 + M_{n(r)+2}^1)|\psi\rangle \tag{37}$$

$$\simeq_{O(r\sqrt{\epsilon})} N_*^\diamond|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} (N_{n(r)+1}^0 + N_{n(r)+1}^1)|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} (N_{n(r)+2}^0 + N_{n(r)+2}^1)|\psi\rangle \tag{38}$$

$$M_*^\perp N_*^\perp|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_*^\perp|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_{n(r)+1}^\perp|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_{n(r)+2}^\perp|\psi\rangle \tag{39}$$

$$\simeq_{O(r\sqrt{\epsilon})} N_*^\perp|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N_{n(r)+1}^\perp|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N_{n(r)+2}^\perp|\psi\rangle. \tag{40}$$

and

$$M_{n(r)+1}^0|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N_{n(r)+1}^0|\psi\rangle \qquad M_{n(r)+1}^1|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N_{n(r)+1}^1|\psi\rangle \tag{41}$$

$$M_{n(r)+2}^0|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N_{n(r)+2}^0|\psi\rangle \qquad M_{n(r)+2}^1|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N_{n(r)+2}^1|\psi\rangle. \tag{42}$$

Proof. Recall that the implication of S being an ϵ-approximate strategy is that $P(ab|xy) \simeq_{O(r^2\epsilon)} \tilde{P}(ab|xy)$. We prove $(M_{n(r)+1}^0 + M_{n(r)+1}^1)|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N_*^\diamond|\psi\rangle$ as an example. The rest of the relations follow the same line of arguments. From the correlation, we can observe that

$$\langle\psi|(M_{n(r)+1}^0 + M_{n(r)+1}^1)N_*^\diamond|\psi\rangle = \frac{2}{d-1} - O(r^2\epsilon).$$

On the other hand, the correlation also tells us that

$$\langle\psi|N_*^\diamond|\psi\rangle \simeq_{O(r^2\epsilon)} \langle\psi|(M_{n(r)+1}^0 + M_{n(r)+1}^1)|\psi\rangle \simeq_{O(r^2\epsilon)} \frac{2}{d-1},$$

which implies that

$$\begin{aligned} &\|(M_{n(r)+1}^0 + M_{n(r)+1}^1)|\psi\rangle - N_*^\diamond|\psi\rangle\|^2 \\ =&\langle\psi|N_*^\diamond|\psi\rangle + \langle\psi|(M_{n(r)+1}^0 + M_{n(r)+1}^1)|\psi\rangle - 2\langle\psi|N_*^\diamond(M_{n(r)+1}^0 + M_{n(r)+1}^1)N_*^\diamond|\psi\rangle \\ =&O(r^2\epsilon), \end{aligned}$$

where we used the fact that $(M_{n(r)+1}^0 + M_{n(r)+1}^1)^2 = M_{n(r)+1}^0 + M_{n(r)+1}^1$. Hence we can conclude that $(M_{n(r)+1}^0 + M_{n(r)+1}^1)N_*^\diamond|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N_*^\diamond|\psi\rangle$. □

Proposition 6.6. *Let S be an ϵ-approximate strategy. We define*

$$M_{n(r)+1} = M^0_{n(r)+1} - M^1_{n(r)+1} \qquad M_{n(r)+2} = M^0_{n(r)+2} - M^1_{n(r)+2}, \tag{43}$$

$$N_{n(r)+1} = N^0_{n(r)+1} - N^1_{n(r)+1} \qquad N_{n(r)+2} = N^0_{n(r)+2} - N^1_{n(r)+2}. \tag{44}$$

Then S induces two strategies that achieve $\langle \mathcal{I}_{\cot(-\pi/d)} \rangle_{\max} - O(dr\sqrt{\epsilon})$, which are

$$S_1 = \left(M_{n(r)+1}, M_{n(r)+2}, N_1, N_2, \frac{M_*^{\diamond}|\psi\rangle}{\|M_*^{\diamond}|\psi\rangle\|} \right), \tag{45}$$

$$S_2 = \left(N_{n(r)+1}, N_{n(r)+2}, M_1, M_2, \frac{M_*^{\diamond}|\psi\rangle}{\|M_*^{\diamond}|\psi\rangle\|} \right). \tag{46}$$

Proof. We will prove that $\langle\psi|M^0_{n(r)+1}N^0_1|\psi\rangle \simeq_{O(\sqrt{d}r\sqrt{\epsilon})} \langle\psi|M_*^{\diamond}M^0_{n(r)+1}N^0_1M_*^{\diamond}|\psi\rangle$ as an example,

$$\begin{aligned}
\langle\psi|M^0_{n(r)+1}N^0_1|\psi\rangle =& \langle\psi|(M_*^{\diamond} + M_*^{\perp})M^0_{n(r)+1}N^0_1(M_*^{\diamond} + M_*^{\perp})|\psi\rangle \\
=& \langle\psi|M_*^{\diamond}M^0_{n(r)+1}N^0_1M_*^{\diamond}|\psi\rangle + \langle\psi|M_*^{\diamond}M^0_{n(r)+1}N^0_1M_*^{\perp}|\psi\rangle \\
&+ \langle\psi|M_*^{\perp}M^0_{n(r)+1}N^0_1M_*^{\diamond}|\psi\rangle + \langle\psi|M_*^{\perp}M^0_{n(r)+1}N^0_1M_*^{\perp}|\psi\rangle \\
\simeq_{O(r\sqrt{\epsilon})}& \langle\psi|M_*^{\diamond}M^0_{n(r)+1}N^0_1M_*^{\diamond}|\psi\rangle + \langle\psi|M_*^{\diamond}M^0_{n(r)+1}N^0_1M^{\perp}_{n(r)+1}|\psi\rangle \\
&+ \langle\psi|M^{\perp}_{n(r)+1}M^0_{n(r)+1}N^0_1M_*^{\diamond}|\psi\rangle + \langle\psi|M^{\perp}_{n(r)+1}M^0_{n(r)+1}N^0_1M^{\perp}_{n(r)+1}|\psi\rangle \\
=& \langle\psi|M_*^{\diamond}M^0_{n(r)+1}N^0_1M_*^{\diamond}|\psi\rangle,
\end{aligned}$$

where we use the facts that $M_*^{\perp}|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M^{\perp}_{n(r)+1}|\psi\rangle$ and that $\mathrm{supp}(M^0_{n(r)+1}) \cap \mathrm{supp}(M^{\perp}_{n(r)+1}) = \emptyset$. This means that if Alice and Bob share state $M_*^{\diamond}|\psi\rangle / \|M_*^{\diamond}|\psi\rangle\|$ and apply $M^0_{n(r)+1}N^0_1$ we get conditional probability

$$\frac{\langle\psi|M_*^{\diamond}M^0_{n(r)+1}N^0_1M_*^{\diamond}|\psi\rangle}{\langle\psi|M_*^{\diamond}|\psi\rangle} \simeq_{O(dr\sqrt{\epsilon})} \frac{\langle\psi|M^0_{n(r)+1}N^0_1|\psi\rangle}{\langle\psi|M_*^{\diamond}|\psi\rangle} \simeq_{O(dr^2\epsilon)} \frac{\cos^2(-\pi/2d)}{2}.$$

because

$$\frac{\|\langle\psi|M_*^{\diamond}M^0_{n(r)+1}N^0_1M_*^{\diamond}|\psi\rangle - \langle\psi|M^0_{n(r)+1}N^0_1|\psi\rangle\|}{\langle\psi|M_*^{\diamond}|\psi\rangle} \leq \frac{O(r\sqrt{\epsilon})}{\frac{2}{d-1} - O(r^2\epsilon)} = O(dr\sqrt{\epsilon}).$$

Using the fact that $O(dr\sqrt{\epsilon})$ dominates $O(dr^2\epsilon)$, we can prove that the correlation generated by S_1 is $O(dr\sqrt{\epsilon})$-close to the ideal correlation that achieves $\langle \mathcal{I}_{\cot(-\pi/d)} \rangle_{\max}$. Hence, the strategy S_1 achieves the value $\langle \mathcal{I}_{\cot(-\pi/d)} \rangle_{\max} - O(dr\sqrt{\epsilon})$. The proof for S_2 is very similar so we omit it here. □

Based on Propositions 6.5 and 6.6, we identify a special component of $|\psi\rangle$, whose special properties are summarized in the proposition below. In the next lemma, we establish that this special component can be used to approximate $|\psi\rangle$.

Proposition 6.7. *Let S be an ϵ-approximate strategy and $M_{n(r)+1}$, $M_{n(r)+2}$, $N_{n(r)+1}$ and $N_{n(r)+2}$ as defined in Proposition 6.6. We also define*

$$|\psi_1\rangle = 1/2(M^0_{n(r)+1} + iM_{n(r)+2}M^1_{n(r)+1} - iM_{n(r)+2}M^0_{n(r)+1} + M^1_{n(r)+1})|\psi\rangle \tag{47}$$

$$|\psi_2\rangle = 1/2(M^1_{n(r)+1} + iM_{n(r)+2}M^0_{n(r)+1} - iM_{n(r)+2}M^1_{n(r)+1} + M^0_{n(r)+1})|\psi\rangle \tag{48}$$

$$|\psi_3\rangle = 1/2(N^0_{n(r)+1} - iN_{n(r)+2}N^1_{n(r)+1} + iN_{n(r)+2}N^0_{n(r)+1} + N^1_{n(r)+1})|\psi\rangle \tag{49}$$

$$|\psi_4\rangle = 1/2(N^1_{n(r)+1} + iN_{n(r)+2}N^1_{n(r)+1} - iN_{n(r)+2}N^0_{n(r)+1} + N^1_{n(r)+1})|\psi\rangle \tag{50}$$

then

$$
\begin{aligned}
&|\psi_1\rangle \simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} |\psi_3\rangle \textit{ and } |\psi_2\rangle \simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} |\psi_4\rangle, && (51)\\
&M_*^\diamond|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} |\psi_1\rangle + |\psi_2\rangle, && (52)\\
&\langle\psi_2|\psi_1\rangle \simeq_{O(r\sqrt{\epsilon})} 0, && (53)\\
&\||\psi_i\rangle\|^2 \simeq_{O(r\sqrt{\epsilon})} 1/(d-1) \quad \textit{for } i = 1,2,3,4, && (54)\\
&N_1N_2|\psi_1\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} \omega_d|\psi_1\rangle, && (55)\\
&M_1M_2|\psi_1\rangle \simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} \omega_d^{-1}|\psi_1\rangle. && (56)
\end{aligned}
$$

In this proof, we use the notation $c := \cos(\pi/d)$ and $s := \sin(\pi/d)$.

Proof. We prove eq. (51) to eq. (56) one by one.

To prove eq. (51), recall that $M^0_{n(r)+1}|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N^0_{n(r)+1}|\psi\rangle$ and $M^1_{n(r)+1}|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} N^1_{n(r)+1}|\psi\rangle$ proved in Proposition 6.5, so it suffices to show $(M_{n(r)+2}M^1_{n(r)+1} - M_{n(r)+2}M^0_{n(r)+1})|\psi\rangle$ is close to $(-N_{n(r)+2}N^1_{n(r)+1} + N_{n(r)+2}N^0_{n(r)+1})|\psi\rangle$, or equivalently, $M_{n(r)+2}M_{n(r)+1}|\psi\rangle$ is close to $-N_{n(r)+2}N_{n(r)+1}|\psi\rangle$.

At this point, we use the anticommutation relation induced by S_1 (eq. (17)) and conclude that

$$M_{n(r)+1}M_{n(r)+2}M_*^\diamond|\psi\rangle \simeq_{\frac{\||M_*^\diamond|\psi\rangle\rangle\|}{c\sqrt{s}}O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} -M_{n(r)+1}M_{n(r)+2}M_*^\diamond|\psi\rangle. \tag{57}$$

If we use the fact $c = 1 - \pi^2/(2d^2) + O(1/d^4)$ and $s = \pi/d + O(1/d^3)$, the distance above is of order $O(\sqrt{d}\sqrt{r}\epsilon^{1/4})$. Then we can show that

$$
\begin{aligned}
M_{n(r)+2}M_{n(r)+1}|\psi\rangle &= M_{n(r)+2}M_{n(r)+1}(M^0_{n(r)+1} + M^1_{n(r)+1})|\psi\rangle\\
&\simeq_{O(r\sqrt{\epsilon})} M_{n(r)+2}M_{n(r)+1}M_*^\diamond|\psi\rangle\\
&\simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} -M_{n(r)+1}M_{n(r)+2}M_*^\diamond|\psi\rangle\\
&\simeq_{O(r\sqrt{\epsilon})} -M_{n(r)+1}M_{n(r)+2}|\psi\rangle\\
&\simeq_{O(r\sqrt{\epsilon})} -M_{n(r)+1}N_{n(r)+2}|\psi\rangle\\
&\simeq_{O(r\sqrt{\epsilon})} -N_{n(r)+2}N_{n(r)+1}|\psi\rangle.
\end{aligned}
$$

The proof of $|\psi_2\rangle \simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} |\psi_4\rangle$ is similar.

To prove eq. (52), we use the definition of $|\psi_1\rangle$ and $|\psi_2\rangle$

$$|\psi_2\rangle + |\psi_1\rangle = 1/2(2M^0_{n(r)+1} + 2M^1_{n(r)+1})|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_*^\diamond|\psi\rangle. \tag{58}$$

To prove eq. (53), we expand the inner product

$$
\begin{aligned}
&\langle\psi_1|\psi_2\rangle\\
=&1/4\langle\psi|(M^0_{n(r)+1} - iM^1_{n(r)+1}M_{n(r)+2} + iM^0_{n(r)+1}M_{n(r)+2} + M^1_{n(r)+1})\\
&(M^1_{n(r)+1} + iM_{n(r)+2}M^0_{n(r)+1} - iM_{n(r)+2}M^1_{n(r)+1} + M^0_{n(r)+1})|\psi\rangle\\
=&1/2i\langle\psi|(M^0_{n(r)+1}M_{n(r)+2}M^0_{n(r)+1} - M^1_{n(r)+1}M_{n(r)+2}M^1_{n(r)+1}|\psi\rangle.
\end{aligned}
$$

Then we will show that $\langle\psi|M^0_{n(r)+1}M_{n(r)+2}M^0_{n(r)+1}|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} 0$ as follows

$$
\begin{aligned}
\langle\psi|(M^0_{n(r)+1}M_{n(r)+2}M^0_{n(r)+1}|\psi\rangle &\simeq_{O(r\sqrt{\epsilon})} \langle\psi|(N^0_{n(r)+1}M_{n(r)+2}N^0_{n(r)+1})|\psi\rangle\\
&= \langle\psi|M_{n(r)+2}N^0_{n(r)+1})|\psi\rangle\\
&\simeq_{O(r^2\epsilon)} 0,
\end{aligned}
$$

where the last approximation comes from the correlation. Similarly, we have

$$\langle\psi|(M^1_{n(r)+1}M_{n(r)+2}M^1_{n(r)+1}|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} \langle\psi|M_{n(r)+2}N^1_{n(r)+1}|\psi\rangle \simeq_{O(r^2\epsilon)} 0.$$

Hence, $\langle\psi_2|\psi_1\rangle \simeq_{O(r\sqrt{\epsilon})} 0$.

To prove eq. (54), we use $\||\psi_1\rangle\|^2$ as an example. The norms of the other 3 vectors can be derived with previous results. We start by expanding the inner product

$$\begin{aligned}
&\langle\psi_1|\psi_1\rangle\\
=&1/4\langle\psi|(M^0_{n+1} - iM^1_{n+1}M_{n+2} + iM^0_{n+1}M_{n+2} + M^1_{n+1})\\
&(M^0_{n+1} + iM_{n+2}M^1_{n+1} - iM_{n+2}M^0_{n+1} + M^1_{n+1})|\psi\rangle\\
=&1/2\langle\psi|(M^0_{n+1} + M^1_{n+1} - iM^1_{n+1}M_{n+2}M^0_{n+1} + iM^0_{n+1}M_{n+2}M^1_{n+1}|\psi\rangle
\end{aligned}$$

Here we apply the same trick to flip M^1_{n+1} to Bob's side and get

$$\begin{aligned}
\langle\psi|M^1_{n(r)+1}M_{n(r)+2}M^0_{n(r)+1}|\psi\rangle &\simeq_{O(r\sqrt{\epsilon})} \langle\psi|N^1_{n(r)+1}M_{n(r)+2}N^0_{n(r)+1}|\psi\rangle\\
&= \langle\psi|N^1_{n(r)+1}N^0_{n(r)+1}M_{n(r)+2}|\psi\rangle = 0.
\end{aligned}$$

Similarly, we have $\langle\psi|M^0_{n(r)+1}M_{n(r)+2}M^1_{n(r)+1}|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} 0$. Hence, we get

$$\langle\psi_1|\psi_1\rangle \simeq_{O(r\sqrt{\epsilon})} 1/2\langle\psi|(M^0_{n(r)+1} + M^1_{n(r)+1}|\psi\rangle \simeq_{O(r^2\epsilon)} \frac{1}{d-1}. \tag{59}$$

To prove eq. (55), we need to dive deeper into the implication of the fact that S_1 almost achieves $\langle\mathcal{I}_{\cot(-\pi/d)}\rangle_{\max}$. We define

$$Z_A := M_{n(r)+1} \qquad X_A := M_{n(r)+2} \tag{60}$$

$$Z_B := \frac{N_1+N_2}{2c} \qquad X_B := \frac{N_1-N_2}{-2s}. \tag{61}$$

Note that Z_A should be defined by $M^0_{n(r)+1} - M^1_{n(r)+1} + M^\perp_{n(r)+1}$, but $M^\perp_{n(r)+1}$ is almost orthogonal to $M^\diamond_*$, so we omit the $M^\perp_{n(r)+1}$ term here. In the definition of X_A, we omit $M^\perp_{n(r)+2}$ too.

Recalling eq. (13), which states that

$$\|Z_B M^\diamond_*|\psi\rangle - M_{n(r)+1}M^\diamond_*|\psi\rangle\| \le \sqrt{s}\|M^\diamond_*|\psi\rangle\|O(\sqrt{d}\sqrt{r}\epsilon^{1/4}) = O(\sqrt{r}\epsilon^{1/4}/\sqrt{d}).$$

Since $M^\diamond_*|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} (M^0_{n(r)+1} + M^1_{n(r)+1})|\psi\rangle$, the equation above can be rewritten as

$$Z_B(M^0_{n(r)+1} + M^1_{n(r)+1})|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} (M^0_{n(r)+1} - M^1_{n(r)+1})|\psi\rangle.$$

On the other hand, using the fact $M^0_{n(r)+1}M^1_{n(r)+1} = 0$, we can get that

$$\begin{aligned}
O(r\sqrt{\epsilon}/d) \ge&\|Z_B(M^0_{n(r)+1} + M^1_{n(r)+1})|\psi\rangle - (M^0_{n(r)+1} - M^1_{n(r)+1})|\psi\rangle\|^2\\
=&\|Z_B M^0_{n(r)+1}|\psi\rangle - M^0_{n(r)+1}|\psi\rangle\|^2 + \|Z_B M^1_{n(r)+1}|\psi\rangle + M^1_{n(r)+1}|\psi\rangle\|^2,
\end{aligned}$$

which immediately gives us that

$$\frac{N_1+N_2}{2c}M^0_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} M^0_{n(r)+1}|\psi\rangle, \tag{62}$$

$$\frac{N_1+N_2}{2c}M^1_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} -M^1_{n(r)+1}|\psi\rangle. \tag{63}$$

Substituting the appropriate operators and states into eqs. (15) and (16), we get

$$\frac{N_1 - N_2}{-2s} M^0_{n(r)+1} M^\diamond_* |\psi\rangle \simeq_{\|M^\diamond_*|\psi\rangle\|/\sqrt{s}O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} M_{n(r)+2} M^1_{n(r)+1} M^\diamond_* |\psi\rangle$$
$$\frac{N_1 - N_2}{-2s} M^1_{n(r)+1} M^\diamond_* |\psi\rangle \simeq_{\|M^\diamond_*|\psi\rangle\|/\sqrt{s}O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} M_{n(r)+2} M^0_{n(r)+1} M^\diamond_* |\psi\rangle.$$

Using the facts that $M^\diamond_*|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} (M^0_{n(r)+1} + M^1_{n(r)+1})|\psi\rangle$ and $O(\sqrt{d}\sqrt{r}\epsilon^{1/4})$ dominates $O(r\sqrt{\epsilon})$, we know

$$(N_1 - N_2) M^0_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} -2s M_{n(r)+2} M^1_{n+1}|\psi\rangle, \tag{64}$$
$$(N_1 - N_2) M^1_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} -2s M_{n(r)+2} M^0_{n+1}|\psi\rangle. \tag{65}$$

In summary, we have shown that

$$N_1 M^0_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} c M^0_{n(r)+1}|\psi\rangle - s M_{n(r)+2} M^1_{n(r)+1}|\psi\rangle, \tag{66}$$
$$N_1 M^1_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} -c M^1_{n(r)+1}|\psi\rangle - s M_{n(r)+2} M^0_{n(r)+1}|\psi\rangle, \tag{67}$$
$$N_2 M^0_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} c M^0_{n(r)+1}|\psi\rangle + s M_{n(r)+2} M^1_{n(r)+1}|\psi\rangle, \tag{68}$$
$$N_2 M^1_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} -c M^1_{n(r)+1}|\psi\rangle + s M_{n(r)+2} M^0_{n(r)+1}|\psi\rangle. \tag{69}$$

Then we can calculate that

$$\begin{aligned}
&N_1 N_2 M^0_{n(r)+1}|\psi\rangle \\
\simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})}\ &(c N_1 M^0_{n(r)+1} + s M_{n(r)+2} N_1 M^1_{n(r)+1})|\psi\rangle \\
\simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})}\ &(c^2 M^0_{n(r)+1} - cs M_{n(r)+2} M^1_{n(r)+1} - cs M_{n(r)+2} M^1_{n(r)+1} - s^2 M^2_{n(r)+2} M^0_{n(r)+1})|\psi\rangle \\
=\ &(\cos(2\pi/d) M^0_{n(r)+1} - \sin(2\pi/d) M_{n(r)+2} M^1_{n(r)+1})|\psi\rangle.
\end{aligned}$$

and

$$\begin{aligned}
&N_1 N_2 M^1_{n(r)+1}|\psi\rangle \\
\simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})}\ &(-c N_1 M^1_{n(r)+1} + s M_{n(r)+2} N_1 M^0_{n(r)+1})|\psi\rangle \\
\simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})}\ &(c^2 M^1_{n(r)+1} + cs M_{n(r)+2} M^0_{n(r)+1} + cs M_{n(r)+2} M^0_{n(r)+1} - s^2 M^2_{n(r)+2} M^1_{n(r)+1})|\psi\rangle \\
=\ &(\cos(2\pi/d) M^1_{n(r)+1} + \sin(2\pi/d) M_{n(r)+2} M^0_{n(r)+1})|\psi\rangle,
\end{aligned}$$

or equivalently,

$$N_1 N_2 M_{n(r)+2} M^1_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} \sin(\frac{2\pi}{d}) M^0_{n(r)+1}|\psi\rangle + \cos(\frac{2\pi}{d}) M_{n(r)+2} M^1_{n(r)+1}|\psi\rangle.$$

Combining the two results above we know

$$N_1 N_2 (M^0_{n(r)+1} + i M_{n(r)+2} M^1_{n(r)+1})|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} \omega_d (M^0_{n(r)+1} + i M_{n(r)+2} M^1_{n(r)+1})|\psi\rangle. \tag{70}$$

Multiplying both sides by $1/2(\mathbb{1} - i M_{n(r)+2})$, we get

$$N_1 N_2 |\psi_1\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} \omega_d |\psi_1\rangle. \tag{71}$$

Note that here we use another form of $|\psi_1\rangle$ which is $|\psi_1\rangle = 1/2(\mathbb{1} - iM_{n(r)+2})(M^0_{n(r)+1} + iM_{n(r)+2}M^1_{n(r)+1})|\psi\rangle$.

To prove eq. (56), we first show that $M_1M_2|\psi_3\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} \omega_d^{-1}|\psi_3\rangle$, then we use the fact that $|\psi_3\rangle \simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} |\psi_1\rangle$ to draw the conclusion.

From the strategy S_2, we can summarize that

$$(M_1 + M_2)N^0_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} 2cN^0_{n(r)+1}|\psi\rangle. \tag{72}$$

$$(M_1 + M_2)N^1_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} -2cN^1_{n(r)+1}|\psi\rangle, \tag{73}$$

$$(M_1 - M_2)N^0_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} -2sN_{n(r)+2}N^1_{n(r)+1}|\psi\rangle \tag{74}$$

$$(M_1 - M_2)N^1_{n(r)+1}|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} -2sN_{n(r)+2}N^0_{n(r)+1}|\psi\rangle. \tag{75}$$

With similar calculation as in the previous case, we can get

$$M_1M_2(N^0_{n(r)+1} - iN_{n(r)+2}N^1_{n(r)+1})|\psi\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} \omega_d^{-1}(N^0_{n(r)+1} - iN_{n(r)+2}N^1_{n(r)+1})|\psi\rangle. \tag{76}$$

Multiplying both sides of the equation above by $1/2(\mathbb{1} + iN_{n(r)+2})$ we get

$$M_1M_2|\psi_3\rangle \simeq_{O(\sqrt{r}\epsilon^{1/4}/\sqrt{d})} \omega_d^{-1}|\psi_3\rangle, \tag{77}$$

where we use the fact that $|\psi_3\rangle = 1/2(\mathbb{1} + iN_{n(r)+2})(N^0_{n(r)+1} - iN_{n(r)+2}N^1_{n(r)+1})|\psi\rangle$ In the end, the relation $|\psi_1\rangle \simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} |\psi_3\rangle$ gives us that

$$M_1M_2|\psi_1\rangle \simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} \omega_d^{-1}|\psi_1\rangle. \tag{78}$$

□

Next lemma is the major result of this subsection, which gives us a decomposition of $|\psi\rangle$. Before proving the next lemma, we make one important observation. Recall that Proposition 6.2 establishes that there exist observables $\{M_i, N_i\}$ such that $M_iN_i|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} |\psi\rangle$ for $i = 1 \dots n(r)$. Then we can deduce that

$$M_3M_4 \otimes N_3N_4|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_3 \otimes N_3|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} |\psi\rangle, \tag{79}$$

or equivalently,

$$U_A \otimes U_B|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} |\psi\rangle. \tag{80}$$

Intuitively, it means that U_AU_B only permutes the Schmidt basis of $|\psi\rangle$. Similarly, we also have $U_A^\dagger U_B^\dagger|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} |\psi\rangle$. We use this observation in the proof of the following proposition.

Lemma 6.8. *Let S be an ϵ-approximate strategy and define $|\psi'\rangle := \sum_{j=1}^{(d-1)}(U_A^\dagger U_B^\dagger)^{\log_r j}|\psi_1\rangle$, then*

$$|\psi\rangle \simeq_{O(d^{7/4}r^{1/4}\epsilon^{1/8})} |\psi'\rangle, \tag{81}$$

where $\log_r j$ is the discrete log.

Proof. We first prove that $\||\psi'\rangle\|^2 \simeq_{O(d^4\sqrt{r}\epsilon^{1/8})} 1$. Observe that

$$\langle\psi_1|(U_AU_B)^{\log_r j}(U_A^\dagger U_B^\dagger)^{\log_r j}|\psi_1\rangle = \langle\psi_1|\psi_1\rangle \simeq_{O(r\sqrt{\epsilon})} \frac{1}{d-1}.$$

Using the relation $O_A(U_A^\dagger)^j|\psi\rangle \simeq_{O(r^j\sqrt{\epsilon})} (U_A^\dagger)^j O_A^{r^j}|\psi\rangle$ and $|\psi_1\rangle \simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} |\psi_3\rangle$, we can see that

$$\begin{aligned}
O_A(U_A^\dagger)^j|\psi_1\rangle &\simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} O_A(U_A^\dagger)^j|\psi_3\rangle \\
&= 1/2(\mathbb{1} + iN_{n(r)+2})(N^0_{n(r)+1} - iN_{n(r)+2}N^1_{n(r)+1})O_A(U_A^\dagger)^j|\psi\rangle \\
&\simeq_{O(r^j\sqrt{\epsilon})} 1/2(\mathbb{1} + iN_{n(r)+2})(N^0_{n(r)+1} - iN_{n(r)+2}N^1_{n(r)+1})(U_A^\dagger)^j O_A^{r^j}|\psi\rangle \\
&= (U_A^\dagger)^j O_A^{r^j}|\psi_3\rangle \\
&\simeq_{O(r^j\sqrt{r}\epsilon^{1/4}/\sqrt{d})} \omega_d^{-r^j}(U_A^\dagger)^j|\psi_3\rangle \\
&\simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} \omega_d^{-r^j}(U_A^\dagger)^j|\psi_1\rangle.
\end{aligned}$$

In short, we have shown that

$$O_A(U_A^\dagger)^j|\psi_1\rangle \simeq_{O(\sqrt{d}r^{j+1/2}\epsilon^{1/4})} \omega_d^{-r^j}(U_A^\dagger)^j|\psi_1\rangle. \tag{82}$$

Hence for $j \neq j'$, Proposition 2.9 tells us that

$$\langle\psi_1|(U_AU_B)^j(U_A^\dagger U_B^\dagger)^{j'}|\psi_1\rangle \simeq_{O(d^{3/2}(r^j+r^{j'})\sqrt{r}\epsilon^{1/4})} 0. \tag{83}$$

We can calculate the norm of $|\psi'\rangle$ as

$$\begin{aligned}
\langle\psi'|\psi'\rangle &= \sum_{j,j'=1}^{d-1}\langle\psi_1|(U_AU_B)^{\log_r j}(U_A^\dagger U_B^\dagger)^{\log_r j'}|\psi_1\rangle \\
&\simeq_{O(d^{5/2}r^{d+1/2}\epsilon^{1/4})} \sum_{j=1}^{d-1}\langle\psi_1|(U_AU_B)^{\log_r j}(U_A^\dagger U_B^\dagger)^{\log_r j}|\psi_1\rangle \\
&\simeq_{O(dr\sqrt{\epsilon})} (d-1)\frac{1}{d-1} = 1.
\end{aligned}$$

Next we calculate $\langle\psi|\psi'\rangle$ which is

$$\langle\psi|\psi'\rangle = \sum_{j=1}^{d-1}\langle\psi|(U_A^\dagger U_B^\dagger)^{\log_r j}|\psi_1\rangle \simeq_{O(dr\sqrt{\epsilon})} (d-1)\langle\psi|\psi_1\rangle.$$

The problem is reduced to calculate $\langle\psi|\psi_1\rangle$, which is

$$\begin{aligned}
\langle\psi|\psi_1\rangle &= \frac{1}{2}\langle\psi|(M^0_{n(r)+1} + iM_{n(r)+2}M^1_{n(r)+1} - iM_{n(r)+2}M^0_{n(r)+1} + M^1_{n(r)+1})|\psi\rangle \\
&= \frac{1}{2}\left(\langle\psi|(M^0_{n(r)+1} + M^1_{n(r)+1})|\psi\rangle - i\langle\psi|M_{n(r)+2}M_{n(r)+1}|\psi\rangle\right) \\
&\simeq_{O(r\sqrt{\epsilon})} \frac{1}{2}\left(\langle\psi|(M^0_{n(r)+1} + M^1_{n(r)+1})|\psi\rangle - i\langle\psi|N_{n(r)+2}M_{n(r)+1}|\psi\rangle\right) \\
&\simeq_{O(r^2\epsilon)} \frac{1}{2}\frac{2}{d-1} = \frac{1}{d-1},
\end{aligned}$$

where the last approximation comes from the correlation. Hence, we know $\langle\psi|\psi'\rangle \simeq_{O(dr\sqrt{\epsilon})} 1$. In the end, we apply Proposition 2.10 to get

$$|\psi\rangle \simeq_{O(d^{5/4}r^{d/2+1/4}\epsilon^{1/8})} |\psi'\rangle. \tag{84}$$

□

6.3 Implications of passing the commutation test

The commutation test is introduced to certify the operators used for the Magic Square test and the operators for the extended weighted CHSH test commute. For details of the general commutation test, we refer to Appendix A.2 of [?]. We use the similar technique in our proof but our proof also bases on the correlation given in Table 4.

The implication of passing the commutation test is summarized in the following proposition.

Proposition 6.9. *Let S be an ϵ-approximate strategy , then*

$$M^0_{n(r)+1}M_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_jM^0_{n(r)+1}|\psi\rangle, \qquad M^1_{n(r)+1}M_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_jM^1_{n(r)+1}|\psi\rangle, \tag{85}$$

$$M_{n(r)+2}M_j|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_jM_{n(r)+2}|\psi\rangle, \tag{86}$$

for $j = r+6, r+8, r+9, r+11$.

Note that M_j's tested in the commutation test are also used in the Magic Square test, which will be used to construct the isometry later.

Proof. The case of $M_jM_{n(r)+2}$ is an application of Lemma 28 of [?] with $\sqrt{\epsilon}$ replaced by $O(r\sqrt{\epsilon})$.

For the other relations, we prove $M^0_{n(r)+1}M_{r+6}|\psi\rangle \simeq_{O(r\sqrt{\epsilon})} M_{r+6}M_{n(r)+1}|\psi\rangle$ to demonstrate the techniques involved. We denote Bob's projectors by $\{N^{b_1b_2}_{(n(r)+1,r+6)}\}$ where $b_1 \in \{0,1,\perp\}$ and $b_2 \in \{0,1\}$. From the correlation, we can get that

$$\begin{aligned} M^0_{n(r)+1}|\psi\rangle &\simeq_{O(r\sqrt{\epsilon})} (N^{00}_{(n(r)+1,r+6)} + N^{01}_{(n(r)+1,r+6)})|\psi\rangle \\ M_{r+6}|\psi\rangle &\simeq_{O(r\sqrt{\epsilon})} (\sum_{b_1,b_2}(-1)^{b_2}N^{b_1b_2}_{(n(r)+1,r+6)})|\psi\rangle. \end{aligned}$$

Then we can show that

$$\begin{aligned} M^0_{n(r)+1}M_{r+6}|\psi\rangle &\simeq_{O(r\sqrt{\epsilon})} M^0_{n(r)+1}(\sum_{b_1,b_2}(-1)^{b_2}N^{b_1b_2}_{(n(r)+1,r+6)})|\psi\rangle \\ &\simeq_{O(r\sqrt{\epsilon})} (\sum_{b_1,b_2}(-1)^{b_2}N^{b_1b_2}_{(n(r)+1,r+6)})(N^{00}_{(n(r)+1,r+6)} + N^{01}_{(n(r)+1,r+6)})|\psi\rangle \\ &= (N^{00}_{(n(r)+1,r+6)} + N^{01}_{(n(r)+1,r+6)})(\sum_{b_1,b_2}(-1)^{b_2}N^{b_1b_2}_{(n(r)+1,r+6)})|\psi\rangle \\ &\simeq_{O(r\sqrt{\epsilon})} (N^{00}_{(n(r)+1,r+6)} + N^{01}_{(n(r)+1,r+6)})M_{r+6}|\psi\rangle \\ &= M_{r+6}(N^{00}_{(n(r)+1,r+6)} + N^{01}_{(n(r)+1,r+6)})|\psi\rangle \\ &\simeq_{O(r\sqrt{\epsilon})} M_{r+6}M^0_{n(r)+1}|\psi\rangle. \end{aligned}$$

□

6.4 An isometry

With the previous propositions and lemmas, we can prove the self-testing theorem.

Theorem 6.10. *Let S be an ϵ-approximate strategy. There exists an isometry $\Phi_A \otimes \Phi_B = (\Phi_{A,2} \otimes \Phi_{B,2})(\Phi_{A,1} \otimes \Phi_{B,2})$ and a state $|junk\rangle$ such that $\| |junk\rangle \|^2 \simeq_{O(dr\sqrt{\epsilon})} 1$ and*

$$\Phi_A \otimes \Phi_B(|\psi\rangle) \simeq_{O(d^{5/4}r^{2d}\epsilon^{1/8}))} |junk\rangle \otimes |\Sigma^{(2)}\rangle^{\otimes 2} \otimes |\Sigma^{(d-1)}\rangle \tag{87}$$

$$\Phi_A \otimes \Phi_B(O_A|\psi\rangle) \simeq_{O(d^{5/4}r^{2d}\epsilon^{1/8}))} |junk\rangle \otimes |\Sigma^{(2)}\rangle^{\otimes 2} \otimes \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} \omega_d^{d-j} |x_{d-j}\rangle_{A'} |x_j\rangle_{B'} \tag{88}$$

$$\Phi_A \otimes \Phi_B(O_B|\psi\rangle) \simeq_{O(d^{5/4}r^{2d}\epsilon^{1/8}))} |junk\rangle \otimes |\Sigma^{(2)}\rangle^{\otimes 2} \otimes \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} \omega_d^{j} |x_{d-j}\rangle_{A'} |x_j\rangle_{B'} \tag{89}$$

$$\Phi_A \otimes \Phi_B(U_A|\psi\rangle) \simeq_{O(d^{5/4}r^{2d}\epsilon^{1/8}))} |junk\rangle \otimes |\Sigma^{(2)}\rangle^{\otimes 2} \otimes \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} |x_{(d-j)r^{-1}}\rangle_{A'} |x_j\rangle_{B'} \tag{90}$$

$$\Phi_A \otimes \Phi_B(U_B|\psi\rangle) \simeq_{O(d^{5/4}r^{2d}\epsilon^{1/8}))} |junk\rangle \otimes |\Sigma^{(2)}\rangle^{\otimes 2} \otimes \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} |x_{d-j}\rangle_{A'} |x_{jr^{-1}}\rangle_{B'}. \tag{91}$$

Before proving Theorem 6.10, we state the consequence of it. It implies that, for any odd prime number d with primitive root r, there exists a correlation of size $\Theta(r^2)$ that can self-test the maximally entangled state of local dimension $4(d-1)$. Moreover, there are infinitely many prime numbers whose primitive roots are in the set $\{2,3,5\}$ [**?**], Hence, our main result is the following theorem.

Theorem 6.11. *There exists an infinity-sized set D of prime numbers such that for each $d \in D$, there exists a constant-sized correlation that can self-test the maximally entangled state of local dimension $4(d-1)$.*

The isometry used in the proof of Theorem 6.10 has two steps. We denote the isometry used in the first step and the second step by $\Phi_{A,1} \otimes \Phi_{B,1}$ and $\Phi_{A,2} \otimes \Phi_{B,2}$ respectively, which are illustrated in the two figures below. In the first step, we use U_A, O_A and U_B, O_B to distill a copy of $|\Sigma^{(d-1)}\rangle$. In the second step. we use the observables $M_{r+6}, M_{r+8}, M_{r+9}, M_{r+11}$ and $N_{r+6}, N_{r+8}, N_{r+9}, N_{r+11}$ to distill two copies of $|\Sigma^{(2)}\rangle$. Intuitively, M_{r+6} acts as $\mathbb{1} \otimes X$; M_{r+8} acts as $X \otimes \mathbb{1}$; M_{r+9} acts as $\mathbb{1} \otimes Z$ and M_{r+11} acts as $Z \otimes \mathbb{1}$; and similarly on Bob's side.

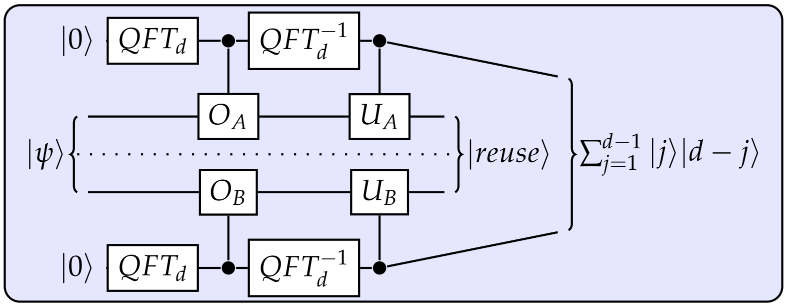

Figure 1: The isometries $\Phi_{A,1} \otimes \Phi_{B,1}$.

The first isometry has the following steps:

1. Append control register $|0\rangle_{A'}$ on Alice's side and $|0\rangle_{B'}$ on Bob's side;

2. Apply Quantum Fourier Transform (QFT_d) to Alice and Bob's control registers;

3. Apply Controlled-$O_{A/B}$ operations (i.e. if the control register is in state $|k\rangle_{A'/B'}$, apply $O^k_{A/B}$.);

4. Apply inverse Quantum Fourier Transform (QFT_d^{-1}) to the control registers;
5. Apply Controlled-$U_{A/B}$ operations (i.e. If Alice's control register is in state $|j\rangle$, she applies $U_A^{\log_r(d-j)}$. If Bob's control register is in state $|j\rangle$, he applies $(U_B^\dagger)^{\log_r j}$).

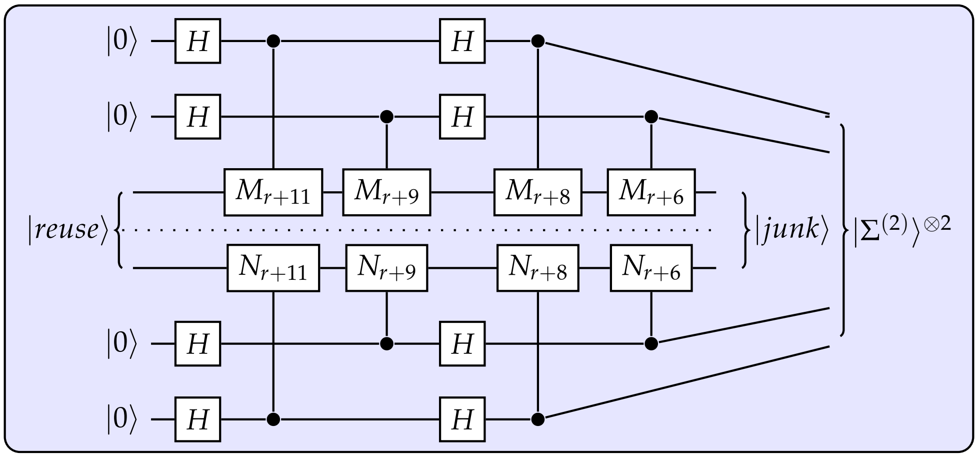

Figure 2: The isometries $\Phi_{A,2} \otimes \Phi_{B,2}$.

The second isometry is the standard isometry used in the self-testing result of the Magic Square game **[?]**. The only thing to note is that the state $|reuse\rangle$, which is on the same Hilbert space as $|\psi\rangle$ and produced by $\Phi_{A,1} \otimes \Phi_{B,1}$, is the input state to $\Phi_{A,2} \otimes \Phi_{B,2}$. In this sense, our isometry is a 2-step sequential procedure.

Proof of Theorem 6.10. The proof takes two steps. We first show that

$$\Phi_{A,1} \otimes \Phi_{B,1}(|\psi\rangle) \simeq_{O(d^{5/4}r^{2d+1/2}\epsilon^{1/8})} \sqrt{d-1}|\psi_1\rangle \otimes \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} |x_{d-j}\rangle_{A'} |x_j\rangle_{B'}. \tag{92}$$

Then we show that there exists a state $|junk\rangle$ such that

$$\Phi_{A,2} \otimes \Phi_{B,2}(\sqrt{d-1}|\psi_1\rangle) \simeq_{O(r\sqrt{\epsilon})} |junk\rangle \otimes |\Sigma^{(2)}\rangle^{\otimes 2}. \tag{93}$$

Combining the two equations above we get that

$$\begin{aligned}
&\Phi_{A,2} \otimes \Phi_{B,2}(\Phi_{A,1} \otimes \Phi_{B,1}(|\psi\rangle)) \\
&\simeq_{O(d^{5/2}\sqrt{r}\epsilon^{1/8})} \Phi_{A,2} \otimes \Phi_{B,2}(\sqrt{d-1}|\psi_1\rangle) \otimes \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} |x_{d-j}\rangle_{A'} |x_j\rangle_{B'} \\
&\simeq_{O(r\sqrt{\epsilon})} |junk\rangle \otimes |\Sigma^{(2)}\rangle^{\otimes 2} \otimes \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} |x_{d-j}\rangle_{A'} |x_j\rangle_{B'},
\end{aligned}$$

where we use the fact that $\Phi_{A,2} \otimes \Phi_{B,2}$ only acts on the state $|\psi_1\rangle$.

Lemma 6.8 implies that $\Phi_{A,1} \otimes \Phi_{B,1}(|\psi\rangle) \simeq_{O(d^{5/4}r^{d/2}r^{1/4}\epsilon^{1/8})} \Phi_{A,1} \otimes \Phi_{B,1}(|\psi'\rangle)$, so we focus on how the isometry evolves $|\psi'\rangle$. The evolution is summarized below.

$$\begin{aligned}
&\sum_{j=1}^{d-1}(U_A^\dagger U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_1\rangle_{A'}|x_1\rangle_{B'} && (94)\\
\xrightarrow{QFT_d}&\frac{1}{d}\sum_{k_1,k_2=0}^{d-1}\sum_{j=1}^{d-1}(U_A^\dagger U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_{k_1}\rangle_{A'}|x_{k_2}\rangle_{B'} && (95)\\
\xrightarrow{\text{Controlled-}O_{A/B}}&\frac{1}{d}\sum_{k_1,k_2=0}^{d-1}\sum_{j=1}^{d-1}O_A^{k_1}(U_A^\dagger)^{\log_r j}O_B^{k_2}(U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_{k_1}\rangle_{A'}|x_{k_2}\rangle_{B'} && (96)\\
\simeq_{O(\sqrt{d}r^{2d}\sqrt{r}\epsilon^{1/4})}&\frac{1}{d}\sum_{k_1,k_2=0}^{d-1}\sum_{j=1}^{d-1}(U_A^\dagger U_B^\dagger)^{\log_r j}\omega_d^{(k_2-k_1)j}|\psi_1\rangle|x_{k_1}\rangle_{A'}|x_{k_2}\rangle_{B'} && (97)\\
\xrightarrow{QFT_d^{-1}}&\frac{1}{d^2}\sum_{l_1,l_2=0}^{d-1}\sum_{j=1}^{d-1}(U_A^\dagger U_B^\dagger)^{\log_r j}\omega_d^{k_1(d-j-l_1)}\omega_d^{k_2(j-l_2)}|\psi_1\rangle|x_{l_1}\rangle_{A'}|x_{l_2}\rangle_{B'} && (98)\\
=&\sum_{j=1}^{d-1}(U_A^\dagger U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_{d-j}\rangle_{A'}|x_j\rangle_{B'} && (99)\\
\xrightarrow{\text{Controlled-}U_{A/B}}&\sum_{j=1}^{d-1}U_A^{\log_r j}(U_A^\dagger)^{\log_r j}U_B^{\log_r j}(U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_{d-j}\rangle_{A'}|x_j\rangle_{B'} && (100)\\
=&|\psi_1\rangle\otimes\sum_{j=1}^{d-1}|x_{d-j}\rangle_{A'}|x_j\rangle_{B'}, && (101)
\end{aligned}$$

In summary, we have shown that

$$\Phi_{A,1} \otimes \Phi_{B,1}(|\psi\rangle) \simeq_{O(d^{5/4}r^{2d+1/2}\epsilon^{1/8})} \sqrt{d-1}|\psi_1\rangle \otimes \frac{1}{\sqrt{d-1}}\sum_{j=1}^{d-1}|x_{d-j}\rangle_{A'}|x_j\rangle_{B'}, \qquad (102)$$

where we use the fact that $O(d^{5/4}r^{2d+1/2}\epsilon^{1/8})$ dominates both $O(d^{5/4}r^{d/2+1/4}\epsilon^{1/8})$ and $O(\sqrt{d}r^{2d}\sqrt{r}\epsilon^{1/4})$. Since $\||\psi_1\rangle\|^2 \simeq_{O(r\sqrt{\epsilon})} 1/(d-1)$, we know $\sqrt{d-1}\||\psi_1\rangle\| \simeq_{O(\sqrt{d}\sqrt{r}\epsilon^{1/4})} 1$.

To prove eq. (93), we first recall that $|\psi_1\rangle = \frac{1}{2}(M^0_{n(r)+1} + iM_{n(r)+2}M^1_{n(r)+1} - iM_{n(r)+2}M^0_{n(r)+1} + M^1_{n(r)+1})|\psi\rangle$. The other observation that we need is that

$$\begin{aligned}
&\Phi_{A,2}\otimes\Phi_{B,2}(|\psi_1\rangle) && (103)\\
=&\frac{1}{16}\sum_{\underline{a}\underline{b}\in\{0,1\}^4}M^{\underline{a}(4)}_{r+6}N^{\underline{b}(4)}_{r+6}M^{\underline{a}(3)}_{r+8}N^{\underline{b}(3)}_{r+8}M^{\underline{a}(2)}_{r+9}N^{\underline{b}(2)}_{r+9}M^{\underline{a}(1)}_{r+11}N^{\underline{b}(1)}_{r+11}|\psi_1\rangle|\underline{a}\rangle_{A''}|\underline{b}\rangle_{B''} && (104)
\end{aligned}$$

For simplicity we define $M^{\underline{a}} := M^{\underline{a}(4)}_{r+6}M^{\underline{a}(3)}_{r+8}M^{\underline{a}(2)}_{r+9}M^{\underline{a}(1)}_{r+11}$ and $N^{\underline{b}} := N^{\underline{b}(4)}_{r+6}N^{\underline{b}(3)}_{r+8}N^{\underline{b}(2)}_{r+9}N^{\underline{b}(1)}_{r+11}$. We would like to show that

$$M^{\underline{a}}|\psi_1\rangle \simeq_{O(r\sqrt{\epsilon})} \frac{1}{2}(M^0_{n(r)+1} + iM_{n(r)+2}M^1_{n(r)+1} - iM_{n(r)+2}M^0_{n(r)+1} + M^1_{n(r)+1})M^{\underline{a}}|\psi\rangle, \qquad (105)$$

for all $\underline{a} \in \{0,1\}^4$, which can be justified by Proposition 6.9. We show how to apply Proposition 6.9 to commute $M^{\underline{a}(1)}_{r+11}$ through $\frac{1}{2}(M^0_{n(r)+1} + iM_{n(r)+2}M^1_{n(r)+1} - iM_{n(r)+2}M^0_{n(r)+1} + M^1_{n(r)+1})$ and then similar process can be repeated for M_{r+9}, M_{r+8} and M_{r+6},

$$
\begin{aligned}
& M_{r+11}^{\underline{a}(1)}\frac{1}{2}(M_{n(r)+1}^0 + iM_{n(r)+2}M_{n(r)+1}^1 - iM_{n(r)+2}M_{n(r)+1}^0 + M_{n(r)+1}^1)|\psi\rangle \\
\simeq_{O(r\sqrt{\epsilon})} & \ \frac{1}{2}[(M_{n(r)+1}^0 + M_{n(r)+1}^1)M_{r+11}^{\underline{a}(1)} - iM_{r+11}^{\underline{a}(1)}M_{n(r)+2}N_{n(r)+1}]|\psi\rangle \\
\simeq_{O(r\sqrt{\epsilon})} & \ \frac{1}{2}[(M_{n(r)+1}^0 + M_{n(r)+1}^1)M_{r+11}^{\underline{a}(1)} - iM_{n(r)+2}M_{r+11}^{\underline{a}(1)}N_{n(r)+1}]|\psi\rangle \\
\simeq_{O(r\sqrt{\epsilon})} & \ \frac{1}{2}[(M_{n(r)+1}^0 + M_{n(r)+1}^1)M_{r+11}^{\underline{a}(1)} - iM_{n(r)+2}M_{r+11}^{\underline{a}(1)}M_{n(r)+1}]|\psi\rangle \\
\simeq_{O(r\sqrt{\epsilon})} & \ \frac{1}{2}[(M_{n(r)+1}^0 + M_{n(r)+1}^1)M_{r+11}^{\underline{a}(1)} - iM_{n(r)+2}M_{n(r)+1}M_{r+11}^{\underline{a}(1)}]|\psi\rangle.
\end{aligned}
$$

In summary, we have

$$
M^{\underline{a}}N^{\underline{b}}|\psi_1\rangle \simeq_{O(r\sqrt{\epsilon})} \frac{1}{2}(M_{n(r)+1}^0 + iM_{n(r)+2}M_{n(r)+1}^1 - iM_{n(r)+2}M_{n(r)+1}^0 + M_{n(r)+1}^1)M^{\underline{a}}N^{\underline{b}}|\psi\rangle \tag{106}
$$

for all $\underline{a}, \underline{b} \in \{0,1\}^4$, which implies that

$$
\Phi_{A,2}\otimes\Phi_{B,2}(|\psi_1\rangle) \simeq_{O(r\sqrt{\epsilon})} \tag{107}
$$

$$
\frac{1}{2}(M_{n(r)+1}^0 + iM_{n(r)+2}M_{n(r)+1}^1 - iM_{n(r)+2}M_{n(r)+1}^0 + M_{n(r)+1}^1)(\Phi_{A,2}\otimes\Phi_{B,2}(|\psi\rangle)). \tag{108}
$$

At this point we can apply Lemma C.1 of [?] to $\Phi_{A,2}\otimes\Phi_{B,2}(|\psi\rangle)$ and conclude that

$$
\Phi_{A,2}\otimes\Phi_{B,2}(\sqrt{d-1}|\psi_1\rangle) \simeq_{O(\sqrt{d}r\sqrt{\epsilon})} |junk\rangle \otimes |\Sigma^{(2)}\rangle^{\otimes 2}, \tag{109}
$$

for some state $|junk\rangle$ whose norm can be deduced from the norm of $|\psi_1\rangle$.

In the rest of the proof, we only show how $\Phi_{A,1}\otimes\Phi_{B,1}$ acts on $O_A|\psi\rangle$ and $U_A|\psi\rangle$.

If the initial state is $O_A|\psi\rangle$, we first use the fact that $\Phi_{A,1}\otimes\Phi_{B,1}(O_A|\psi\rangle) \simeq_{O(d^{5/4}r^{d/2+1/4}\epsilon^{1/8})} \Phi_{A,1}\otimes\Phi_{B,1}(O_A|\psi'\rangle)$, and then we calculate $\Phi_{A,1}\otimes\Phi_{B,1}(O_A|\psi'\rangle)$ as

$$
O_A|\psi'\rangle|0\rangle_{A'}|0\rangle_{B'} = \sum_{j=1}^{d-1} O_A(U_A^\dagger U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_1\rangle_{A'}|x_1\rangle_{B'} \tag{110}
$$

$$
\simeq_{O(\sqrt{d}r^{d+1/2}\epsilon^{1/4})} \sum_{j=1}^{d-1}(U_A^\dagger U_B^\dagger)^{\log_r j}\omega_d^{-j}|\psi_1\rangle|x_1\rangle_{A'}|x_1\rangle_{B'} \tag{111}
$$

$$
\xrightarrow{QFT_d} \frac{1}{d}\sum_{j=1}^{d-1}\sum_{k_1,k_2=0}^{d-1}(U_A^\dagger U_B^\dagger)^{\log_r j}\omega_d^{-j}|\psi_1\rangle|x_{k_1}\rangle_{A'}|x_{k_2}\rangle_{B'} \tag{112}
$$

$$
\xrightarrow{\text{Controlled-}O_{A/B}} \frac{1}{d}\sum_{j=1}^{d-1}\sum_{k_1,k_2=0}^{d-1} O_A^{k_1}(U_A^\dagger)^{\log_r j}O_B^{k_2}(U_B^\dagger)^{\log_r j}\omega_d^{-j}|\psi_1\rangle|x_{k_1}\rangle_{A'}|x_{k_2}\rangle_{B'} \tag{113}
$$

$$
\simeq_{O(\sqrt{d}r^{2d+1/2}\epsilon^{1/4})} \frac{1}{d}\sum_{k_1,k_2=0}^{d-1}\sum_{j=1}^{d-1}(U_A^\dagger U_B^\dagger)^{\log_r j}\omega_d^{-j}\omega_d^{(k_2-k_1)j}|\psi_1\rangle|x_{k_1}\rangle_{A'}|x_{k_2}\rangle_{B'} \tag{114}
$$

$$
\xrightarrow{QFT_d^{-1}} \sum_{j=1}^{d-1}(U_A^\dagger U_B^\dagger)^{\log_r j}\omega_d^{d-j}|\psi_1\rangle|x_{d-j}\rangle_{A'}|x_j\rangle_{B'} \tag{115}
$$

$$
\xrightarrow{\text{Controlled-}U_{A/B}} \sqrt{d-1}|\psi_1\rangle \otimes \frac{1}{\sqrt{d-1}}\sum_{j=1}^{d-1}\omega_d^{d-j}|x_{d-j}\rangle_{A'}|x_j\rangle_{B'}. \tag{116}
$$

The analysis for $\Phi_{A,1} \otimes \Phi_{B,1}(O_B|\psi\rangle)$ is very similar.

If the initial state is $U_A|\psi\rangle$, we first use the fact that $\Phi_{A,1} \otimes \Phi_{B,1}(U_A|\psi\rangle) \simeq_{O(d^{5/4}r^{d/2+1/4}\epsilon^{1/8})} \Phi_{A,1} \otimes \Phi_{B,1}(U_A|\psi'\rangle)$, and then we calculate $\Phi_{A,1} \otimes \Phi_{B,1} U_A|\psi'\rangle)$.

$$\begin{aligned}
U_A|\psi'\rangle|0\rangle_{A'}|0\rangle_{B'} &= \sum_{j=1}^{d-1} U_A(U_A^\dagger U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_1\rangle_{A'}|x_1\rangle_{B'} && (117)\\
&= \sum_{j=1}^{d-1} (U_A^\dagger)^{\log_r j-1}(U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_1\rangle_{A'}|x_1\rangle_{B'} && (118)\\
&\xrightarrow{QFT_d} \frac{1}{d}\sum_{j=1}^{d-1}\sum_{k_1,k_2=0}^{d-1} (U_A^\dagger)^{\log_r j-1}(U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_{k_1}\rangle_{A'}|x_{k_2}\rangle_{B'} && (119)\\
&\xrightarrow{\text{Controlled-}O_{A/B}} \frac{1}{d}\sum_{j=1}^{d-1}\sum_{k_1,k_2=0}^{d-1} O_A^{k_1}(U_A^\dagger)^{\log_r j-1}O_B^{k_2}(U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_{k_1}\rangle_{A'}|x_{k_2}\rangle_{B'} && (120)\\
&\simeq_{O(\sqrt{d}r^{2d+1/2}\epsilon^{1/4})} \frac{1}{d}\sum_{k_1,k_2=0}^{d-1}\sum_{j=1}^{d-1} (U_A^\dagger)^{\log_r j-1}(U_B^\dagger)^{\log_r j}\omega_d^{k_2 j-k_1 j r^{-1}}|\psi_1\rangle|x_{k_1}\rangle_{A'}|x_{k_2}\rangle_{B'} && (121)\\
&\xrightarrow{QFT_d^{-1}} \sum_{j=1}^{d-1} (U_A^\dagger)^{\log_r j-1}(U_B^\dagger)^{\log_r j}|\psi_1\rangle|x_{(d-j)r^{-1}}\rangle_{A'}|x_j\rangle_{B'} && (122)\\
&\xrightarrow{\text{Controlled-}U_{A/B}} \sqrt{d-1}|\psi_1\rangle \otimes \frac{1}{\sqrt{d-1}}\sum_{j=1}^{d-1}|x_{(d-j)r^{-1}}\rangle_{A'}|x_j\rangle_{B'}, && (123)
\end{aligned}$$

where we use the fact that $(d-j)r^{-1} \equiv d - jr^{-1} \pmod{d}$. The analysis for $\Phi_{A,1} \otimes \Phi_{B,1}(U_B|\psi\rangle)$ is very similar. □

References

A The proof of Theorem 4.1

This proof follows the same line of argument in Appendix A of Ref. [?]. We first find two sum-of-square decompositions of $2\sqrt{\alpha^2+1}\mathbb{1} - \mathcal{I}_\alpha$, where $\mathcal{I}_\alpha$ is expressed in terms of $\{M_x\}$ and $\{N_y\}$. The decompositions allow us to determine some key relations between Alice and Bob's observables and their shared state, which will be used to draw the conclusion.

Proof. The first step is to find a sum-of-square decomposition of the following Bell expression

$$\bar{\mathcal{I}}_\alpha = 2\sqrt{\alpha^2+1}\mathbb{1} - \mathcal{I}_\alpha = \frac{2}{\sin(\mu)}\mathbb{1} - \frac{\cos(\mu)}{\sin(\mu)}(M_1N_1 + M_1N_2) - M_2N_1 + M_2N_2. \tag{124}$$

With the notation $c := \cos(\mu)$, $s := \sin(\mu)$ and

$$\begin{aligned}
Z_A &= M_1 & X_A &= M_2\\
Z_B &= \frac{N_1+N_2}{2c} & X_B &= \frac{N_1-N_2}{2s},
\end{aligned}$$

the two SOS decompositions that we use are

$$\bar{\mathcal{I}}_\alpha = \frac{s\bar{\mathcal{I}}_\alpha^2 + 4sc^2(Z_A X_B + X_A Z_B)^2}{4}, \tag{125}$$

$$\bar{\mathcal{I}}_\alpha = \frac{c^2}{s}(Z_A - Z_B)^2 + s(X_A - X_B)^2. \tag{126}$$

The verification is omitted here. From the SOS decomposition, we establish eqs. (13) to (18). We define

$$S_1 = \frac{\sqrt{s}}{2}\bar{\mathcal{I}}_\alpha, \qquad S_2 = \sqrt{s}c(Z_A X_B + X_A Z_B),$$
$$S_3 = \frac{c}{\sqrt{s}}(Z_A - Z_B), \qquad S_4 = \sqrt{s}(X_A - X_B)$$

then

$$\bar{\mathcal{I}}_\alpha = S_1^2 + S_2^2 = S_3^2 + S_4^2 \tag{127}$$

The fact that the quantum strategy $(|\psi\rangle, \{M_x\}_{x=1,2}, \{N_y\}_{y=1,2}$ achieves that $\langle\bar{\mathcal{I}}_\alpha\rangle \leq \epsilon$ implies that $\langle\psi|S_i^2|\psi\rangle \leq \epsilon$, and equivalently, $\|S_i|\psi\rangle\| \leq \sqrt{\epsilon}$ for $i = 1, 2, 3, 4$. From the definitions of S_i's, we can get that

$$\|(X_A Z_B + X_B Z_A)|\psi\rangle\| \leq \frac{1}{c\sqrt{s}}\sqrt{\epsilon} \tag{128}$$

$$\|(Z_A - Z_B)|\psi\rangle\| \leq \frac{\sqrt{s}}{c}\sqrt{\epsilon} \tag{129}$$

$$\|(X_A - X_B)|\psi\rangle\| \leq \frac{1}{\sqrt{s}}\sqrt{\epsilon}. \tag{130}$$

The first and the second inequality give us that

$$\|[Z_A(\mathbb{1} + X_B) - (\mathbb{1} - X_A)Z_B]|\psi\rangle\| \leq \|(X_A Z_B + X_B Z_A)|\psi\rangle\| + \|(Z_A - Z_B)|\psi\rangle\| \leq \frac{s+1}{c\sqrt{s}}\sqrt{\epsilon}. \tag{131}$$

Similarly, the first and the third inequality give us that

$$\|[X_A(\mathbb{1} + Z_B) - X_B(\mathbb{1} - Z_A)]|\psi\rangle\| \leq \frac{c+1}{c\sqrt{s}}\sqrt{\epsilon}. \tag{132}$$

Since $Z_A X_A + X_A Z_A = \frac{S_2}{c\sqrt{s}} + \frac{\sqrt{s}\tilde{X}_A S_3}{c} + \frac{\tilde{Z}_A S_4}{\sqrt{s}}$, we can deduce that

$$\|(Z_A X_A + X_A Z_A)|\psi\rangle\| \leq \frac{1+c+s}{c\sqrt{s}}\sqrt{\epsilon}. \tag{133}$$

To prove eq. (18), we switch to the approximate relation form and derive that

$$X_A Z_A|\psi\rangle \simeq_{\frac{1+c+s}{c\sqrt{s}}\sqrt{\epsilon}} -Z_A X_A|\psi\rangle \tag{134}$$

$$\simeq_{\frac{\sqrt{s}}{c}\sqrt{\epsilon}} -Z_A X_B|\psi\rangle \tag{135}$$

$$\simeq_{\frac{1}{s^{3/2}}\sqrt{\epsilon}} -X_B Z_B|\psi\rangle, \tag{136}$$

where in the last line we use the fact $\|X_B\| \leq 1/s$.

Now we introduce the isometries Φ_A and Φ_B mentioned in the statement of the theorem. They are the same as the ones used in Ref. [**?**]. To construct Φ_A and Φ_B we need to regularize Z_B and X_B to make sure the corresponding operations are unitary in the isometries. We define Z_B^* to be the operator obtained from Z_B by changing all the 0-eigenvalues to 1 and

$$Z'_B := Z^*_B|Z^*_B|^{-1},$$

where $|Z_B^*|$ is obtained from Z_B^* by replacing all negative eigenvalues by its absolute value. In a similar way, we define X_B^* and X'_B. On Alice's side, since Z_A and X_A are unitaries already, we define $Z'_A := Z_A$ and $X'_A = X_A$. The isometries are illustrated in the figure below.

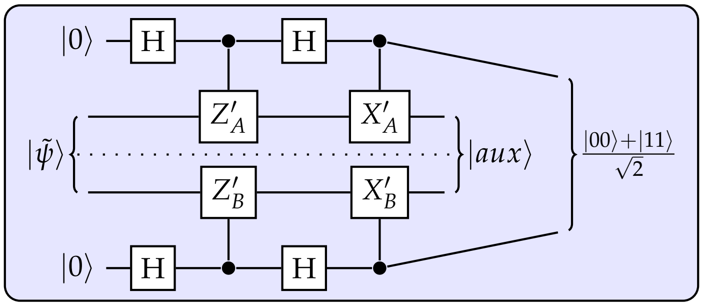

Figure 3: The isometries Φ_A and Φ_B.

To bound $e_{xy} := \|(\Phi_A \otimes \Phi_B)(\tilde{A}_x \otimes \tilde{B}_y)|\tilde{\psi}\rangle - |junk\rangle \otimes (A_x \otimes B_y)|\Sigma^{(2)}\rangle\|$, there are some intermediate steps. Since the derivations are the same as in Ref. [**?**], we only record the key relations here.

$$\begin{aligned}
&\|(Z'_B - Z_B)|\tilde{\psi}\rangle\| \leq \frac{\sqrt{s}}{c}\sqrt{\epsilon},\\
&\|(Z'_B - Z'_A)|\tilde{\psi}\rangle\| \leq 2\frac{\sqrt{s}}{c}\sqrt{\epsilon},\\
&\|(X'_B - X_B)|\tilde{\psi}\rangle\| \leq \frac{c+1}{s^{3/2}}\sqrt{\epsilon} := \delta_1\sqrt{\epsilon},\\
&\|(X'_B Z'_B + Z'_B X'_B)|\tilde{\psi}\rangle\| \leq [\frac{2\sqrt{s}}{c} + \frac{2}{\sqrt{s}} + 2\delta_1 + (\sqrt{2} + \frac{1}{c})(2\frac{\sqrt{s}}{c} + \frac{1+c+s}{c\sqrt{s}})]\sqrt{\epsilon} := \delta_2\sqrt{\epsilon}.
\end{aligned}$$

Then we can calculate that

$$\begin{aligned}
&e_{00} = e_{10} = 2\delta_2\sqrt{\epsilon}\\
&e_{20} = 2(\frac{1+c+s}{c\sqrt{s}} + \delta_2)\sqrt{\epsilon}\\
&e_{01} = e_{02} = e_{11} = e_{12} = [\sqrt{s} + s(2\frac{1+c+s}{c\sqrt{s}} + \delta_1) + 2\delta_2]\sqrt{\epsilon}\\
&e_{21} = e_{22} = [2\frac{1+c+s}{c\sqrt{s}} + \sqrt{s} + s(2\frac{1+c+s}{c\sqrt{s}} + \delta_1) + 2\delta_2]\sqrt{\epsilon},
\end{aligned}$$

so an upper bound of the error is that

$$\forall x, y \in \{0,1,2\}, e_{xy} \in O((\frac{1}{c^2 s^{1/2}} + \frac{1}{s^{3/2}})\sqrt{\epsilon}). \tag{137}$$

□

B The proof of Lemma 5.2

We are going to prove the set $\{U^k O^l\}$ for $k = 0, 1 \dots d-2$ and $l = 1, 2 \dots d-1$ forms a basis of the ring of $(d-1) \times (d-1)$ matrices over $\mathbb{C}$ Note that the unitaries U and O defined in the lemma above satisfy the condition $UOU^\dagger = O^r$. In this proof, we denote the set $\{0, 1, 2 \dots d-2\}$ by $[d-1]$ and the set $\{1, 2 \dots d-1\}$ by $[d-1]+1$.

Proof. We are going to show the $(d-1)^2$ matrices from the set $\{U^k O^l\}_{k\in[d-1], l\in[d-1]+1}$ are linearly independent. Suppose there exists a set of complex numbers $\{x_{k,l}\}_{k\in[d-1], l\in[d-1]+1}$ such that

$$M = \sum_{k=0}^{d-2} \sum_{l=1}^{d-1} x_{k,l} U^k O^l = 0. \tag{138}$$

We define a set of integers $\{k_i\}_{i=1}^{d-1}$ such that $r^{k_i} \equiv i \pmod d$. We shouldn't be making extra assumptions that are not stated in the lemma. This looks to me more like a definition of the variables $\{k_i\}$, rather than an assumption. -Carl The fact that r is a primitive root of d guarantees that k_i's are distinct. Then we can group $\{x_{k,l}\}$ into vectors: $|x_{k_1}\rangle, |x_{k_2}\rangle \dots |x_{k_{d-1}}\rangle$, where $|x_{k_i}\rangle = (x_{k_i,1}, x_{k_i,2} \dots x_{k_i,d-1})^\intercal$. Our goal is equivalent to proving that $|x_{k_i}\rangle = 0$ for all i.

We start with proving that $|x_{k_1}\rangle = 0$. Proving $|x_{k_i}\rangle = 0$ for other i follows a similar argument, so we briefly discuss about it in the end. The entry $\langle 1|M|1\rangle$ can be expressed as

$$\langle 1|M|1\rangle = \sum_{k=0}^{d-2} \sum_{l=1}^{d-1} \sum_{i=1}^{d-1} x_{k,l} \omega_d^{il} \langle 1|i(r^{-1})^k\rangle \langle i|1\rangle. \tag{139}$$

For the term $\langle 1|i(r^{-1})^k\rangle\langle i|1\rangle \neq 0$ we must have $i = 1$ and $r^k \equiv 1 \pmod d$, or equivalently, $k = k_1$. We can conclude that

$$\langle 1|M|1\rangle = \sum_{l=1}^{d-1} x_{k_1,l} \omega_d^l = 0. \tag{140}$$

Similarly we can determine that for all $j = 1, 2 \dots d-1$,

$$\langle j|M|j\rangle = \sum_{k=0}^{d-2} \sum_{l=1}^{d-1} \sum_{i=1}^{d-1} x_{k,l} \omega_d^{il} \langle j|i(r^{-1})^k\rangle \langle i|j\rangle = \sum_{l=1}^{d-1} x_{k_1,l} \omega_d^{jl} = 0. \tag{141}$$

Hence we get $d-1$ equations with $d-1$ variables, and the linear system is

$$W|x_{k_1}\rangle = 0, \tag{142}$$

where $W(m,n) = \omega_d^{mn}$. Then we define

$$\tilde{W} = \begin{pmatrix} 1 & 1 \\ 1 & W \end{pmatrix}. \tag{143}$$

First observe that $\tilde{W}$ is a Vandermonde matrix, hence it is non-singular. Next, we define $|\tilde{x}_{k_1}\rangle = (0, x_{k_1,1}, \dots x_{k_1,d-1})^\intercal$ and prove that it satisfies the condition that

$$\tilde{W}|\tilde{x}_{k_1}\rangle = 0, \tag{144}$$

which involves d equations. The last $d-1$ equations are given by the assumption and M. We only need to prove that $\sum_{l=1}^{d} x_{k_1,l} = 0$, which is required by the first row of $\tilde{W}$. It can be proved by summing the known $d-1$ equations as follows

$$0 = \sum_{j=1}^{d-1} \langle j|M|j\rangle = \sum_{j=1}^{d-1}\sum_{l=1}^{d-1} x_{k_1,l}\omega_d^{jl} = \sum_{l=1}^{d-1} x_{k_1,l}(\sum_{j=1}^{d-1}\omega_d^{jl}) = \sum_{l=1}^{d-1} -x_{k_1,l} \tag{145}$$

where we have used the fact that $\sum_{j=1}^{d-1}\omega_d^{jl} = -1$ for all $l = 1, 2 \ldots d-1$. Since $\tilde{W}$ is non-singluar, we know $|\tilde{x}_{k_1}\rangle = 0$ which implies that $|x_{k_1}\rangle = 0$.

For $|x_{k_a}\rangle$, we look at entries $\{\langle j|M|aj\rangle\}_{j=1}^{d-1}$ for $a = 2 \ldots d-1$ and get equations of the form

$$0 = \langle j|M|aj\rangle = \sum_{l=1}^{d-1} x_{k_a,l}\omega_d^{ajl} \tag{146}$$

The corresponding coefficient matrix has value ω_d^{amn} at coordinate (m, n), so it is also a submatrix of a Vandermonde matrix. Similar argument gives us that $|x_{k_a}\rangle = 0$.

To summarize, we have proven that $x_{k,l} = 0$ for all k and l, which implies that the elements of the set $\{U^k X^l\}$ are linearly independent and forms a basis for the ring of all the $(d-1)\times(d-1)$ matrices over $\mathbb{C}$. Nice proof. I like the use of the Vandermonde matrix. We can think a little about possible simplifications. -Carl □