Self-test prime dimensional EPR pairs with constant alphabet

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*

February 3, 2019

1 Introduction

Self-testing is a unique phenomenon of quantum mechanics. It has many applications in quantum delegated computation [RUV13, CGJV17] and device independent quantum cryptography [MPA11, VV14, MS16, FM18, AFDF+18].

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 [BP15]. and self-test many copies of the 2-dimensional EPR pair in parallel [McK16, Col17]. Self-testing general d-dimensional EPR pairs is a harder task. Recently, a remarkable result by Coladangelo *et al.*[CGS17] has shown that the maximally entangled state with arbitrary local dimension can be self-tested with constant question alphabet but answer alphabet growing with the local dimension. Then Coladangelo and Stark [CS17] further showed that by playing many instances of the generalized Magic Square game and Magic Pentagram game, one can robustly self-test N copies of the maximally entangled state with local dimension d for any $d, N \geq 2$. Both the Magic Square game and the Magic Pentagram game have constant question alphabet but answer alphabet of size d.

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 (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 $\{2, 3, 5\}$ [Mur88], so the set D has infinitely many elements. To prove Theorem 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 to enforce the eigenvalue of some unknown operator which is the product of two binary observables used in the weighted CHSH test. I think this game might have been defined in APM12, and if so we need to cite it. -Carl 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 [Slo17]. Our work is heavily based on [Slo17], 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 all 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 5.

2 Preliminaries and notations

We use $[n]$ to denote the set $\{0, 1 \dots n-1\}$ and $[n]+1$ for the set $\{1, 2 \dots n\}$.

We use some basic number theory in our work. A primitive root of a prime number d is an integer r such that $r \pmod d$ has multiplicative order $d-1$. Equivalently, r is the generator of the multiplicative group of integers modulo d, $\mathbb{Z}_d^\times$ where $\mathbb{Z}_d^\times = \{1, 2 \dots d-1\}$. When we say d has primitive root r, we always mean that r is the smallest primitive root. I'm not clear on what that last sentence means. You just defined what a primitive root is — are you changing the definition now? -Carl

The EPR pair of local dimension $d-1$ for some prime d is denoted by

$$|EPR^{(d-1)}\rangle = \frac{1}{\sqrt{d-1}} \sum_{i=0}^{d-2} |ii\rangle. \tag{1}$$

The superscript $(d-1)$ stresses the local dimension and we follow this convention through this paper.

We self-test $|EPR^{(d-1)}\rangle$ by verify that Alice and Bob has operators

$$X = \sum_{k=1}^{d-1} \omega_d^k |k\rangle\langle k| \qquad\qquad U = \sum_{k=1}^{d-1} |k/r\rangle\langle k|, \tag{2}$$

where $\omega_d = e^{i2\pi/d}$ is the primitive d-th root of unity, and r is the primitive root of d. That sentence is jumping ahead — you should just stick to definitions and notation for now. -Carl Through out this paper, any operation on the label of the eigenvector $|k\rangle$ is taken modulo d, unless otherwise specified, for example, the k/r above.

The weighted CHSH inequality [AMP12]. The first building-block of our result is a robust self-testing result based on the weighted CHSH inequality. In the CHSH scenario, Alice and Bob

has binary observables A_x, B_y for $x, y = 1, 2$ on Hilbert space $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. It needs to be clearer who Alice and Bob are — perhaps you should put this subsection after the "Nonlocal games" subsection. -Carl A binary observable is a Hermitian matrix whose only eigenvalues are either $+1$ or -1. The weighted CHSH inequality is given by

$$\mathcal{I}_\alpha = \alpha\langle A_1 B_1\rangle + \alpha\langle A_1 B_2\rangle + \langle A_2 B_1\rangle - \langle A_2 B_2\rangle \leq 2\alpha, \tag{3}$$

where $\langle A_x B_y\rangle$ is the expectation value of of the observables. Same here – it's not clear enough what this means until the reader has seen the "Nonlocal games" subsection. -Carl If Alice and Bob share product state $|\phi\rangle = |\phi_A\rangle \otimes |\phi_B\rangle$, they cannot violate the weighted CHSH inequality. However, If they share entangled state $|\psi\rangle$, the maximal violation is

$$\mathcal{I}_\alpha \leq 2\sqrt{1+\alpha^2}. \tag{4}$$

Definition 2 (Ideal strategy for $\mathcal{I}_\alpha$)**.** *Define $\mu = \arctan(1/\alpha)$. The ideal strategy for weighted CHSH with parameter α (i.e. achieving maximal violation in eq. (4)) consists of the joint state $|EPR^{(2)}\rangle$ and observables $A_1 = \sigma_z$, $A_2 = \sigma_x$, $B_1 = \cos(\mu)\sigma_z + \sin(\mu)\sigma_x$ and $B_2 = \cos(\mu)\sigma_z - \sin(\mu)\sigma_x$.*

Just to be complete, you should say what σ_z and σ_x denote. -Carl An interesting observation of the weighted CHSH inequality is that its maximal violation can certify the shared states and the measurements up to isometry, which is a phenomenon referred as a self-test. We give formal statement of this self-testing result in Section 4.

Nonlocal games. The two players of a nonlocal game are Alice and Bob. Each of them is requested to give an answer for a randomly chosen question. We denote Alice's question set by $\mathcal{X}$ and 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 $V : \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, which is the scoring function. Such games are nonlocal because Alice and Bob cannot communicate after getting their questions but they may share some strategy before the start of the game. Note that in the literature, the typical scoring function of a nonlocal game maps the input-output pair to $\{0, 1\}$ which corresponds to losing and winning. Allowing the score to be any real number is the key ingredient to our new result.

A quantum strategy of a game G consists of projective measurements $\{\{A_x^a\}_a\}$ on Alice's side, $\{\{B_y^b\}_b\}_y$ on Bob's side, and a shared state $|\psi\rangle$, where a projective measurement P satisfies the condition that $P^2 = P = P^\dagger$. You need to put explicit labels on the quantum systems used by Alice and Bob. (Later on you seem to use "A" and "B" to refer to those systems, but those are also the same letters you use for Alice's and Bob's measurements. That's confusing.) -Carl Then Alice and Bob's quantum strategy produces the correlation

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

where $(A_x^a)^2 = A_x^a = (A_x^a)^\dagger$ and $(B_y^b)^2 = B_y^b = (B_y^b)^\dagger$.

The main contribution of our work is the construction of a correlation, which is the combination of the optimal correlation of a linear system game with constant-sized $\mathcal{X}$,$\mathcal{Y}$,$\mathcal{A}$ and $\mathcal{B}$ and the optimal correlation to violate the weighted CHSH inequality. The reader does not know yet what a linear system game is, so this comment is more confusing than helpful. -Carl The special property of the correlation is that it can self-test $|EPR^{(d-1)}\rangle$ for arbitrary d with certain primitive root. In the next part, we introduce the definition of self-testing, the definition of linear system game, and its operator solution, following definitions given in Ref. [CS17, Slo17].

Before the next definition, you should define what a "correlation" is. -Carl

Definition 3 (Self-testing). *We say that a correlation self-tests a quantum state $|\psi\rangle$, if the correlation is produced by a strategy with shared state $|\psi\rangle$, and for any quantum strategy $S = (\{\tilde{A}_x\}, \{\tilde{B}_y\}, |\tilde{\psi}\rangle)$ that produces the same correlation, there exists local isometries Φ_A and Φ_B on Alice and Bob's side and a state $|junk\rangle$ such that*

$$\Phi_A \otimes \Phi_B |\tilde{\psi}\rangle = |\psi\rangle \otimes |junk\rangle.$$

Definition 4 (Linear system game). *Let $Hx = c$ be an $m \times n$ system of linear equations over $\mathbb{Z}_2$, where H is an m-by-n matrix with entries in $\mathbb{Z}_2 = \{0, 1\}$ and 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 $i \in \mathcal{X} = [m] + 1$ and replies with $a \in \mathcal{A} = \mathbb{Z}_2^{\times n}$, and Bob is given a variable number $j \in \mathcal{Y} = [n] + 1$ and replies with an assignment $b \in \mathcal{B} = \mathbb{Z}_2$. The winning condition is that Alice's assignment to the variables should satisfy equation i and Alice's j-th assignment $a(j)$ should match b. Formally, the winning condition is*

$$\begin{aligned} a(j) &= b && \text{(Consistency)} \\ \sum_{k=1}^{n} H(i,k)a(k) &\equiv c(i) \pmod 2. && \text{(Constraint satisfaction),} \end{aligned}$$

for all $(i, j) \in \mathcal{X} \times \mathcal{Y}$.

The scoring function of linear system games always maps an input-output pair to $\{0, 1\}$, so later when we say a quantum strategy wins a linear system game perfectly, we mean that $V(a, b, x, y) = 0$ implies that $P(ab|xy) = 0$. Make this clear in the definition itself. -Carl We focus on quantum strategies for the linear system game presented in terms of binary observables.

Definition 5 (Quantum strategy of a linear system game). *A quantum strategy for the linear system game $(Hx = c)$ consists of*

1. *a pair of finite Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$;*
2. *a collection of binary observables B_j, $1 \le j \le n$, on $\mathcal{H}_B$ such that $B_j^2 = \mathbb{1}$ for every $1 \le j \le n$;*
3. *a collection of binary observables A_{ij}, $1 \le i \le m$, $1 \le j \le n$ on $\mathcal{H}_A$ such that*
   (a) *$A_{ij}^2 = \mathbb{1}$ for every i, j,*
   (b) *$\Pi_j A_{ij}^{H(i,j)} = (-\mathbb{1})^{c(i)}$ for every i, and*
   (c) *$A_{il}A_{ik} = A_{ik}A_{il}$ 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, and vice versa. This is not entirely true. A "quantum strategy presented in terms of projective measurement" allows Alice to give answers that violate the "Constraint satisfaction" condition, while a "quantum strategy presented in terms of binary observables" does not. -Carl

It has been shown in Ref. [**?**] that the linear system game has a perfect strategy satisfying conditions in Definition 5 if and only if the linear system has a finite-dimensional operator solution in the following sense. There's a missing reference here? -Carl

Definition 6 (Operator solution of a linear system). *An operator solution to a linear system $Hx = c$ over $\mathbb{Z}_2$ is a sequence of bounded Hermitian operators $A_1, A_2, \ldots A_n$ on a Hilbert space $\mathcal{H}$ such that (What does the word "bounded" mean here? It might be best to just state the definition for finite dimensions only.) -Carl*

1. *$A_i^2 = \mathbb{1}$, i.e. A_i is a binary observable, for all $1 \leq i \leq n$;*
2. *If x_l and x_k appear in the same equation, then A_l and A_k commute;*
3. *for all $1 \leq i \leq m$,*

$$\Pi_{k=1}^{n} A_k^{H(i,k)} = (-1)^{c(i)} \mathbb{1}.$$

A finite dimensional operator solution to a linear system $Hx = c$ over $\mathbb{Z}_2$ is an operator solution in which the Hilbert space is finite-dimensional.

Perfect quantum strategies can be extracted from operator solutions and vice versa. Another angle to look at the linear system game ($Hx = c$) is a finitely presented group over $\mathbb{Z}_2$, which is called the solution group.

Definition 7 (Solution group of a linear system game). *Let $Hx = c$ be an $m \times n$ linear system. The solution group of this system is the group*

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

Don't you also need a commutativity condition for J with $\{x_i\}$? -Carl For concepts about group presentations, we refer to Sec. 2 of Ref. [Slo17]. Combining Definition 6 and Definition 7, we know that an operator solution associated to $Hx = c$ is a finite-dimensional representation of $\Gamma(H,c)$ that maps J to $-\mathbb{1}$.

Section 2 needs to be cleaned up in some places, but overall it is pretty good. -Carl

3 The linear system game

The goal of this section and the following two sections is to present a correlation that can self-test $|EPR^{(d-1)}\rangle$, where d is prime and has primitive root $r \in \{2,3,5\}$, and this correlation is denoted by $C(d^{(r)})$. Be more precise — we are constructing such a correlation for any d which has 2, 3, or 5 as a primitive root. -Carl The correlation we constructed is the optimal correlation for a game $G(d^{(r)})$, which is the combination of two tests. In this section, we introduce the linear system game which tests a relation that should be satisfied by Alice and Bob's observables.

Slofstra's seminal work [Slo17] draws our attention to the relation $xyx^{-1} = y^2$. That sentence is too vague – either make it a rigorous statement or leave it out. -Carl The main component of $G(d^{(r)})$ is a linear system game LS, whose solution group is the embedding of the following group

$$\mathcal{P}_r = \langle u, x : uxu^{-1} = x^r \rangle, \text{ for } r \in \{2,3,5\}. \tag{6}$$

In other words, relations in the presentation of the solution group of LS_r can be combined to derive the relation $uxu^{-1} = x^r$.

Proposition 8. *The group $\mathcal{P}_r$ can be embedded into a linear system game LS_r over $\mathbb{Z}_2$, with constant numbers of variables and equations, where each equation involves 3 variables. You need to make this a more precise statement. (It sounds as if you're saying that the numbers of variables and equations is independent of r, which I'm pretty sure is not true.) -Carl*

The process of embedding $\mathcal{P}_r$ into the solution group of LS follows the recipe given in the proofs of Proposition 4.8, Lemma 4.4 and Proposition 4.2 of Ref. [Slo17]. The order of applying the results is the reverse of the order they are presented in Ref. [Slo17]. We give details of this embedding process in Appendix B for $r = 2$, Appendix C for $r = 3$, and Appendix D for $r = 5$.

In the later parts of the paper, we use n_r and m_r to refer to the size of LS_r. The key feature of this embedding is that LS_r has constant-sized input-output alphabet even if the generators u and x have arbitrary order. The product of the generators x_1 and x_2 from the solution group of LS_r plays the role of the generator x in $\mathcal{P}_r$ and the product of the generators x_3 and x_4 plays the role of the generator u in $\mathcal{P}_r$. It's reasonable to make a couple informal statements at the beginning of a section, but you shouldn't be making such vague statements at this point. Everything said at this point in the section should have a precise mathematical meaning. -Carl As shown in Appendix B, Appendix C and Appendix D, x_1,x_2,x_3 and x_4 satisfy the relation

$$(x_3x_4)(x_1x_2)(x_4x_3) = (x_1x_2)^r. \tag{7}$$

Hence the perfect quantum strategy $(|\psi\rangle, \{A_i, B_i\}_{i=1}^n)$ should satisfy the key relation that

$$(A_3A_4)(A_1A_2)(A_4A_3)|\psi\rangle = (A_1A_2)^r|\psi\rangle, \tag{8}$$

and similarly on Bob's side. If we denote the unitary corresponding to the generator u of $\mathcal{P}_r$ by U and the unitary corresponding to x by X, then $U = A_3A_4$, $X = A_1A_2$, and $UXU^\dagger|\psi\rangle = X^r|\psi\rangle$. The special property of U and X is summarized in the following lemma.

Lemma 9. *Suppose there exist unitaries U and X which have the following forms*

$$X = \sum_{i=1}^{d-1} \omega_d^i |i\rangle\langle i| \qquad\qquad U = \sum_{i=1}^{d-1} |i/r\rangle\langle i|, \tag{9}$$

where d is an odd prime number with primitive root r, then the set $\{U^kX^l\}$ for $k = 0, 1 \ldots d-2$ and $l = 1, 2 \ldots d-1$ forms a basis of the ring of $(d-1) \times (d-1)$ matrices over $\mathbb{C}$.

Note that the unitaries U and X defined in the lemma above satisfy the condition $UXU^\dagger = X^r$. In the self-test proof, this lemma will be a critical step.

Proof. We are going to show the $(d-1)^2$ matrices from the set $\{U^kX^l\}_{k\in[d-1],l\in[d-1]+1}$ 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

$$M = \sum_{k=0}^{d-2} \sum_{l=1}^{d-1} x_{k,l} U^k X^l = 0. \tag{10}$$

We further assume that there exists 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: $|x_{k_1}\rangle, |x_{k_2}\rangle \ldots |x_{k_{d-1}}\rangle$, where $|x_{k_i}\rangle = (x_{k_i,1}, x_{k_i,2} \ldots x_{k_i,d-1})^\intercal$. 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

$$\langle 1|M|1\rangle = \sum_{k=0}^{d-2}\sum_{l=1}^{d-1}\sum_{i=1}^{d-1} x_{k,l}\omega_d^{il}\langle 1|i/r^k\rangle\langle i|1\rangle. \tag{11}$$

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

$$\langle 1|M|1\rangle = \sum_{l=1}^{d-1} x_{k_1,l}\omega_d^{l} = 0. \tag{12}$$

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

$$\langle j|M|j\rangle = \sum_{k=0}^{d-2}\sum_{l=1}^{d-1}\sum_{i=1}^{d-1} x_{k,l}\omega_d^{il}\langle j|i/r^k\rangle\langle i|j\rangle = \sum_{l=1}^{d-1} x_{k_1,l}\omega_d^{jl} = 0. \tag{13}$$

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

$$W|x_{k_1}\rangle = 0, \tag{14}$$

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

$$\tilde{W} = \begin{pmatrix} 1 & 1 \\ 1 & W \end{pmatrix}. \tag{15}$$

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

$$\tilde{W}|\tilde{x}_{k_1}\rangle = 0, \tag{16}$$

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

$$0 = \sum_{j=1}^{d-1}\langle j|M|j\rangle = \sum_{j=1}^{d-1}\sum_{l=1}^{d-1} x_{k_1,l}\omega_d^{jl} = \sum_{l=1}^{d-1} x_{k_1,l}(\sum_{j=1}^{d-1}\omega_d^{jl}) = \sum_{l=1}^{d-1} -x_{k_1,l} \tag{17}$$

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

$$0 = \langle j|M|aj\rangle = \sum_{l=1}^{d-1} x_{k_a,l}\omega_d^{ajl} \tag{18}$$

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 $\mathbb{C}$. □

Note that the primitive root r in the lemma is not restricted to the set $\{2, 3, 5\}$, so this lemma works for any odd prime number d.

4 The extended weighted CHSH test

Before introducing the extended weighted CHSH test, we fill in some background about the weighted CHSH inequality, namely, that its maximal violation can self-test 2-dimensional EPR pair and rotated Pauli operators. The robust self-testing result is summarized in the following theorem.

Theorem 10. *Suppose the quantum strategy* $(|\tilde{\psi}\rangle, \{\tilde{A}_x\}_{x\in[2]+1}, \{\tilde{B}_y\}_{y\in[2]+1})$ *achieves the violation at least* $2\sqrt{1+\alpha^2} - \epsilon$ *for some* ϵ*, then there exists a local isometry* $\Phi = \Phi_A \otimes \Phi_B$ *and an auxiliary state* $|aux\rangle$ *such that*

$$\|\Phi(\tilde{A}_x \otimes \tilde{B}_y|\psi\rangle) - |junk\rangle \otimes (A_x \otimes B_y)|EPR^{(2)}\rangle\| = O((\alpha + \frac{1}{\alpha})\sqrt{\epsilon}) \tag{19}$$

for $x, y \in \{0,1,2\}$ *where the subscript* 0 *refers to the identity operator and* A_x, B_y *are defined in Definition 2.*

Our approach is very similar to the one used in Ref. [BP15] so we defer the proof of Theorem 14 till Appendix A. After proving the robust self-testing result, we take one step further and observe an interesting property of the product of Bob's observables.

Proposition 11. *Let* $\mu = \arctan(1/\alpha)$ *for some* α*. Suppose a quantum strategy* $(|\tilde{\psi}\rangle, \{\tilde{A}_x\}_{x\in[2]+1}, \{\tilde{B}_y\}_{y\in[2]+1})$ *achieves the maximal violation of* $\mathcal{I}_\alpha$*, then* $\tilde{B}_0\tilde{B}_1$ *has eigenvalues* $e^{i2\mu}$ *and* $e^{-i2\mu}$*.*

Proof of Proposition 11. Following the argument in the proof of Theorem 14 in Appendix A, we know $\tilde{Z}_B = (\tilde{B}_1 + \tilde{B}_2)/2\cos(\mu)$ is a binary observable. HF: Do we need to justify it? Expanding $\tilde{Z}_B^2$ we get

$$\mathbb{1} = \tilde{Z}_B^2 = \frac{\tilde{B}_1^2 + \tilde{B}_1\tilde{B}_2 + \tilde{B}_2\tilde{B}_1 + \tilde{B}_2^2}{4\cos(\mu)^2} = \frac{2\mathbb{1} + \tilde{B}_1\tilde{B}_2 + \tilde{B}_2\tilde{B}_1}{4\cos^2(\mu)},$$

so we can derive that

$$\tilde{B}_1\tilde{B}_2 + \tilde{B}_2\tilde{B}_1 = 2(2\cos^2(\mu) - 1)\mathbb{1} = 2\cos(2\mu)\mathbb{1}. \tag{20}$$

Observe that $(\tilde{B}_1\tilde{B}_2)^\dagger = \tilde{B}_2\tilde{B}_1$ and $(\tilde{B}_2\tilde{B}_1)(\tilde{B}_1\tilde{B}_2) = \tilde{B}_2\mathbb{1}\tilde{B}_2 = \mathbb{1}$, then we can conclude that $\tilde{B}_1\tilde{B}_2$ is unitary. Suppose the eigen-decomposition of $\tilde{B}_1\tilde{B}_2$ is

$$\tilde{B}_1\tilde{B}_2 = \sum_{i=1}^{m} \lambda_i |i\rangle\langle i| \tag{21}$$

for some orthonormal set of eigenvectors $\{|i\rangle\}_{i=1}^m$ and $\|\lambda_i\| = 1$ for $1 \le i \le m$. Eq. 20 tells us that for each $1 \le i \le m$

$$\lambda_i + \lambda_i^{-1} = 2\,\mathrm{Re}(\lambda_i) = 2\cos(2\mu),$$

so we can conclude that $\lambda_i = e^{i2\mu}$ or $e^{-i2\mu}$.

□

The extended weighted CHSH test is added to make sure that the operator X extracted from Alice and Bob's operator solution of LS has eigenvalues ω_d and ω_d^{d-1}. In Section 5, we will reason why showing these two eigenvalues is enough to guarantee that X has the eigen-structure required by Lemma 9. We denote this game that enforces the eigenvalues of the observable X by $CHSH_X^{(d)}$, where the superscript d emphasizes that the scoring rules of this game depend on d. We remark

that for the extended weighted CHSH test, the primitive root of d is irrelevant, so we drop the superscript (r).

In this test, Alice and Bob each gets a question x,$y \in \{1,2,*\}$ and they answer with $a,b \in \{0,1,\diamond,\perp\}$. The correlation $C(d^{(2)})$ is the optimal correlation of this test. Before presenting the optimal correlation, we give intuitions about how the players should behave.

- **Case 1:** when $x = y = *$, Alice and Bob should answer with $a,b \in \{\diamond,\perp\}$ and their answer should agree;
- **Case 2a:** when $x,y \in \{1,2\}$ and if they answer with $a,b \in \{0,1\}$, then their answers are scored according to $I_{\cot(-\pi/d)}$;
- **Case 2b:** when $x,y \in \{1,2\}$ and if Alice answers with $\perp$, then all Bob's answers are irrelevant;
- **Case 3:** when $x \in 1,2, y = *$, if Bob answers $\diamond$, Alice should answer with $\{0.1\}$ but not $\perp$, if Bob answers $\perp$, Alice should answer $\perp$ too.

The ideal strategy and ideal correlation. Alice and Bob share the state $|\psi\rangle = \frac{1}{\sqrt{d-1}} \sum_{i=1}^{d-1} |u_i\rangle|u_i\rangle$. We define two subspaces $V = \text{span}\{|u_1\rangle, |u_{d-1}\rangle\}$ and $V^\perp = \mathbb{C}^d \setminus \text{span}\{|u_1\rangle, |u_{d-1}\rangle\}$ and define Π_V and $\Pi_V^\perp$ to be the corresponding projectors. Note that V is the subspace on which they should maximize $\langle I_{\cot(-\pi/d)}\rangle$.

For completeness, we show the ideal correlation in the following three charts and then give the projectors. Note that we don't explicitly calculate the conditional probabilities of the form $P(\perp 0|xy)$ for all possible x,y because they are irrelevant.

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

Table 1: Alice and Bob's behaviour when $x = y = *$.

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

Table 2: Alice and Bob's behaviour when $x,y \in [2]$.

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

Table 3: Alice and Bob's behaviour when $x \in [2]$ and $y = *$.

Alice's projectors are

$$A_*^\diamond = \Pi_V, A_*^\perp = \Pi_V^\perp$$
$$A_1^0 = |u_1\rangle\langle u_1|, A_1^1 = |u_{d-1}\rangle\langle u_{d-1}|, A_1^\perp = \Pi_V^\perp$$
$$A_2^0 = \frac{1}{2}(|u_1\rangle + |u_{d-1}\rangle)(\langle u_1| + \langle u_{d-1}|), A_2^1 = \frac{1}{2}(|u_1\rangle - |u_{d-1}\rangle)(\langle u_1| - \langle u_{d-1}|), A_2^\perp = \Pi_V^\perp.$$

Bob's projectors are

$$\begin{aligned}
B_*^\diamond &= \Pi_V, B_*^\perp = \Pi_V^\perp \\
B_1^0|_V &= \left(\cos(\frac{\pi}{2d})|u_1\rangle + \sin(\frac{\pi}{2d})|u_{d-1}\rangle\right)\left(\cos(\frac{\pi}{2d})\langle u_1| + \sin(\frac{\pi}{2d})\langle u_{d-1}|\right) \\
B_1^1|_V &= \left(\sin(\frac{\pi}{2d})|u_1\rangle - \cos(\frac{\pi}{2d})|u_{d-1}\rangle\right)\left(\sin(\frac{\pi}{2d})\langle u_1| - \cos(\frac{\pi}{2d})\langle u_{d-1}|\right) \\
B_2^0|_V &= \left(\cos(\frac{\pi}{2d})|u_1\rangle - \sin(\frac{\pi}{2d})|u_{d-1}\rangle\right)\left(\cos(\frac{\pi}{2d})\langle u_1| - \sin(\frac{\pi}{2d})\langle u_{d-1}|\right) \\
B_2^1|_V &= \left(\sin(\frac{\pi}{2d})|u_1\rangle + \cos(\frac{\pi}{2d})|u_{d-1}\rangle\right)\left(\sin(\frac{\pi}{2d})\langle u_1| + \cos(\frac{\pi}{2d})\langle u_{d-1}|\right).
\end{aligned}$$

About Bob's projectors for input $y \in [2]+1$, we are only interested in their actions when restricted to the subspace V. Their actions on the subspace $V^\perp$ is irrelevant in this game.

The self-testing property of this correlation is summarized in the following lemma.

Lemma 12. *Suppose a quantum strategy* $\left(\{\{\tilde{A}_x^a\}_a\}_x, \{\{\tilde{B}_y^b\}_b\}_y, |\tilde{\psi}\rangle\right)$ *achieves the optimal correlation of the weighted* $CHSH^{(d)}$ *test, and let* $|\tilde{\psi}'\rangle = \tilde{A}_*^\diamond \otimes \tilde{B}_*^\diamond|\tilde{\psi}\rangle / \|\tilde{A}_*^\diamond \otimes \tilde{B}_*^\diamond|\tilde{\psi}\rangle\|^2$*, then there exists isometries* Φ_A *and* Φ_B *and a quantum state* $|junk\rangle$ *such that*

$$\begin{aligned}
&\Phi_A \otimes \Phi_B\left(|\tilde{\psi}'\rangle\right) = |EPR^{(2)}\rangle \otimes |junk\rangle \\
&\Phi_A \otimes \Phi_B\left[\left(\mathbb{1} \otimes \frac{\tilde{B}_1 + \tilde{B}_2}{2\cos(\pi/d)}\right)|\tilde{\psi}'\rangle\right] = [(\mathbb{1} \otimes \sigma_z)|EPR^{(2)}\rangle] \otimes |junk\rangle \\
&\Phi_A \otimes \Phi_B\left[\left(\mathbb{1} \otimes \frac{\tilde{B}_1 - \tilde{B}_2}{-2\sin(\pi/d)}\right)|\tilde{\psi}'\rangle\right] = [(\mathbb{1} \otimes \sigma_x)|EPR^{(2)}\rangle] \otimes |junk\rangle
\end{aligned}$$

where $\tilde{B}_1 = \tilde{B}_1^0 - \tilde{B}_1^1$ *and* $\tilde{B}_2 = \tilde{B}_2^0 - \tilde{B}_2^1$.

Proof. Note that $\tilde{A}_x^a \tilde{B}_y^b$ means $\tilde{A}_x^a \otimes \tilde{B}_y^b$ in the following proof.

From the marginal distribution $P_B(\diamond|*) = P_A(0|1) + P_A(1|1) = 2/(d-1)$, we know $\|\tilde{B}_*^\diamond|\tilde{\psi}\rangle\| = \|(\tilde{A}_1^0 + \tilde{A}_1^1)|\tilde{\psi}\rangle\| = \sqrt{2/d-1}$. Since $P(0\diamond|1*) + P(1\diamond|1*) = 2/(d-1)$, we find that

$$\frac{\langle\tilde{\psi}|\tilde{B}_*^\diamond(\tilde{A}_1^0 + \tilde{A}_1^1)\tilde{B}_*^\diamond|\tilde{\psi}\rangle}{\|\tilde{B}_*^\diamond|\tilde{\psi}\rangle\|^2} = 1,$$

which means that

$$(\tilde{A}_1^0 + \tilde{A}_1^1)\tilde{B}_*^\diamond|\tilde{\psi}\rangle = \tilde{B}_*^\diamond|\tilde{\psi}\rangle. \tag{22}$$

Using the commutation relation between $(\tilde{A}_1^0 + \tilde{A}_1^1)$ and $\tilde{B}_*^\diamond$, we get

$$\frac{\langle\tilde{\psi}|(\tilde{A}_1^0 + \tilde{A}_1^1)\tilde{B}_*^\diamond(\tilde{A}_1^0 + \tilde{A}_1^1)|\tilde{\psi}\rangle}{\|(\tilde{A}_1^0 + \tilde{A}_1^1)|\tilde{\psi}\rangle\|^2} = 1,$$

Similar argument gives us that

$$\tilde{B}_*^\diamond(\tilde{A}_1^0 + \tilde{A}_1^1)|\tilde{\psi}\rangle = (\tilde{A}_1^0 + \tilde{A}_1^1)|\tilde{\psi}\rangle. \tag{23}$$

The two equations above can be chained by commutativity to reach the conclusion that

$$(\tilde{A}_1^0 + \tilde{A}_1^1)|\tilde{\psi}\rangle = \tilde{B}_*^\diamond|\tilde{\psi}\rangle. \tag{24}$$

Following the same line of argument, we can conclude that

$$\tilde{B}_*^\diamond|\tilde{\psi}\rangle = \tilde{A}_*^\diamond|\tilde{\psi}\rangle = (\tilde{A}_1^0 + \tilde{A}_1^1)|\tilde{\psi}\rangle = (\tilde{A}_2^0 + \tilde{A}_2^1)|\tilde{\psi}\rangle. \tag{25}$$

Looking at the marginal distribution when Alice and Bob output $\perp$, we conclude that

$$\tilde{B}_*^\perp|\tilde{\psi}\rangle = \tilde{A}_*^\perp|\tilde{\psi}\rangle = \tilde{A}_1^\perp|\tilde{\psi}\rangle = \tilde{A}_2^\perp|\tilde{\psi}\rangle, \tag{26}$$

with similar arguments.

Next we examine the CHSH-type correlation when $x, y \in [2] + 1$,

$$\begin{aligned}
\langle\tilde{\psi}|\tilde{A}_1^0\tilde{B}_1^0|\tilde{\psi}\rangle =& \langle\tilde{\psi}|(\tilde{A}_*^\diamond + \tilde{A}_*^\perp)\tilde{A}_1^0\tilde{B}_1^0(\tilde{A}_*^\diamond + \tilde{A}_*^\perp)|\tilde{\psi}\rangle \\
=& \langle\tilde{\psi}|\tilde{A}_*^\diamond\tilde{A}_1^0\tilde{B}_1^0\tilde{A}_*^\diamond|\tilde{\psi}\rangle + \langle\tilde{\psi}|\tilde{A}_*^\diamond\tilde{A}_1^0\tilde{B}_1^0\tilde{A}_*^\perp|\tilde{\psi}\rangle \\
&+ \langle\tilde{\psi}|\tilde{A}_*^\perp\tilde{A}_1^0\tilde{B}_1^0\tilde{A}_*^\diamond|\tilde{\psi}\rangle + \langle\tilde{\psi}|\tilde{A}_*^\perp\tilde{A}_1^0\tilde{B}_1^0\tilde{A}_*^\perp|\tilde{\psi}\rangle \\
=& \langle\tilde{\psi}|\tilde{B}_*^\diamond\tilde{A}_1^0\tilde{B}_1^0\tilde{B}_*^\diamond|\tilde{\psi}\rangle + \langle\tilde{\psi}|\tilde{A}_*^\diamond\tilde{A}_1^0\tilde{B}_1^0\tilde{A}_0^\perp|\tilde{\psi}\rangle \\
&+ \langle\tilde{\psi}|\tilde{A}_0^\perp\tilde{A}_1^0\tilde{B}_1^0\tilde{A}_*^\diamond|\tilde{\psi}\rangle + \langle\tilde{\psi}|\tilde{A}_0^\perp\tilde{A}_1^0\tilde{B}_1^0\tilde{A}_0^\perp|\tilde{\psi}\rangle \\
=& \langle\tilde{\psi}|\tilde{B}_*^\diamond\tilde{A}_1^0\tilde{B}_1^0\tilde{B}_*^\diamond|\tilde{\psi}\rangle,
\end{aligned}$$

where we use the facts that $\tilde{B}_*^\diamond|\tilde{\psi}\rangle = \tilde{A}_*^\diamond|\tilde{\psi}\rangle$, $\tilde{A}_*^\perp|\tilde{\psi}\rangle = \tilde{A}_1^\perp|\tilde{\psi}\rangle$ and that $\mathrm{span}(\tilde{A}_1^0) \cap \mathrm{span}(\tilde{A}_0^\perp) = \emptyset$. This means that if Alice and Bob share the state $\tilde{B}_*^\diamond|\tilde{\psi}\rangle / \|\tilde{B}_*^\diamond|\tilde{\psi}\rangle\|$ and apply $\tilde{A}_1^0\tilde{B}_1^0$, the conditional probability is

$$\frac{\langle\tilde{\psi}|\tilde{B}_*^\diamond\tilde{A}_1^0\tilde{B}_1^0\tilde{B}_*^\diamond|\tilde{\psi}\rangle}{\langle\tilde{\psi}|\tilde{B}_*^\diamond\tilde{B}_*^\diamond|\tilde{\psi}\rangle} = \frac{\cos^2(\pi/2d)}{2}. \tag{27}$$

We can re-normalize the other correlations of $a, b \in [2]$ when $x, y \in [2+1]$ similarly, and get a new set of correlations which achieves the maximal value of $\langle \mathbb{1}_{-\cot(\pi/2d)} \rangle$. The conclusion of Lemma 12 follows the application of Theorem 14 on the state $\tilde{A}_*^\diamond\tilde{B}_*^\diamond|\tilde{\psi}\rangle / \|\tilde{A}_*^\diamond\tilde{B}_*^\diamond|\tilde{\psi}\rangle\|$ as $\tilde{A}_*^\diamond\tilde{B}_*^\diamond|\tilde{\psi}\rangle = \tilde{B}_*^\diamond|\tilde{\psi}\rangle$. □

Note that the combination of Lemma 12 with Proposition 11 gives us that $\tilde{B}_1\tilde{B}_2$ has eigenvalues ω_d and ω_d^{-1}.

5 Main result

In this section, we introduce the correlation $C(d^{(r)})$, which is the ideal correlation of $G(d^{(r)})$, and then prove that it can self-test the state $|EPR^{(d-1)}\rangle$.

5.1 The correlation $C(d^{(r)})$

Recall that the linear system game LS_r has n_r variables and m_r equations. In game $G(d^{(r)})$, Alice receives $x \in \{1, \dots m_r + 3\}$ and Bob receives $y \in \{1, \dots, n_r + 1\}$. We follow the previous structure by first give intuition about how they should behave in this game. The correlation can be easily extracted from the behaviour list below.

- When $x \in \{1, \dots m_r\}$ and $y \in \{1, \dots n_r\}$, they should win the linear system game LS_r perfectly;
- when $x \in \{m_r + 1, m_r + 2, m_r + 3\}$ and $y \in \{1, 2, n_r + 1\}$, they should follow the optimal correlation of the test $CHSH_X^{(d)}$, where

$$*_A = m_r + 1, \quad 0_A = m_r + 2, \quad 1_A = m_r + 3, \tag{28}$$
$$*_B = n_r + 1, \quad 0_B = 1, \quad 1_A = 2, \tag{29}$$

  are the inputs for the game $CHSH_X^{(d)}$ (The intuition behind is that $B_1 B_2 = X$.);
- otherwise, their behaviour is irrelevant.

Note that the dimension $d^{(r)}$ is defined in the rules of $CHSH_X^{(d)}$.

Proposition 13. *The correlation $C(d^{(r)})$ can be realized by a quantum strategy.*

We prove this proposition by giving the ideal strategy.

Proof. The construction start with the ideal strategy for the weighted $CHSH_X^{(d)}$ test. We choose $|\psi\rangle = \frac{1}{\sqrt{d-1}} \sum_{i=1}^{d-1} |ii\rangle$ and $V = \text{span}(|1\rangle, |d-1\rangle)$. Alice's projectors and observables for the weighted $CHSH_X^{(d)}$ test are

$$A_*^\diamond = |1\rangle\langle 1| + |d-1\rangle\langle d-1|, \qquad A_*^\perp = \mathbb{1} - A_*^\diamond,$$
$$A_1 = A_1^0 - A_1^1 = |1\rangle\langle 1| - |d-1\rangle\langle d-1|, \qquad A_1^\perp = A_*^\perp,$$
$$A_2 = A_2^0 - A_2^1 = |1\rangle\langle d-1| + |d-1\rangle\langle 1|, \qquad A_2^\perp = A_*^\perp.$$

Bob's observables restricted to the subspace V is

$$B_1|_V = \cos(\frac{\pi}{d})(|1\rangle\langle 1| - |d-1\rangle\langle d-1|) - \sin(\frac{\pi}{d})(|1\rangle\langle d-1| + |d-1\rangle\langle 1|),$$
$$B_2|_V = \cos(\frac{\pi}{d})(|1\rangle\langle 1| - |d-1\rangle\langle d-1|) + \sin(\frac{\pi}{d})(|1\rangle\langle d-1| + |d-1\rangle\langle 1|).$$

To make sure operator B_2 commutes with the operator U that we will construct later, we pick a special basis of $X = B_1 B_2$, which is

$$|x_1\rangle = \frac{-1}{\sqrt{2}}(|1\rangle + i|d-1\rangle), \qquad |x_{d-1}\rangle = \frac{-e^{i\pi/d}}{\sqrt{2}}(|1\rangle - i|d-1\rangle). \tag{30}$$

It can be checked that $X|x_1\rangle = \omega_d|x_1\rangle$ and $X|x_{d-1}\rangle = \omega_d^{d-1}|x_{d-1}\rangle$. In this basis, B_1 and B_2 are expressed as

$$B_1|_V = \omega_d|x_1\rangle\langle x_{d-1}| + \omega_d^{-1}|x_{d-1}\rangle\langle x_1|, \tag{31}$$

$$B_2|_V = |x_1\rangle\langle x_{d-1}| + |x_{d-1}\rangle\langle x_1|. \tag{32}$$

Since Bob gets the same symbols 1 and 2 in different sub-tests, the observables B_1 and B_2 are extended from $B_1|_V$ and $B_2|_V$ as follows

$$B_1 = \sum_{k=1}^{(d-1)/2} \left(\omega_d^k|x_k\rangle\langle x_{d-k}| + \omega_d^{-k}|x_{d-k}\rangle\langle x_k|\right) \tag{33}$$

$$B_2 = \sum_{k=1}^{(d-1)/2} \left(|x_{d-k}\rangle\langle x_k| + |x_k\rangle\langle x_{d-k}|\right). \tag{34}$$

It can be checked that $B_1B_2 = \sum_{i=1}^{d-1} \omega_d^i|x_i\rangle\langle x_i|$ and the operator U is

$$U = \sum_{i=1}^{d-1} |x_{i/r \pmod d}\rangle\langle x_i|. \tag{35}$$

The first step of the embedding procedure in the proof of Proposition 8 requires the commutativity between U and B_2, which can be verified as follows,

$$UB_2U^\dagger = \sum_{i=1}^{d-1} U|x_i\rangle\langle x_{d-i}|U^\dagger = \sum_{i=1}^{d-1} |x_{i/r \pmod d}\rangle\langle x_{(d-i)/r \pmod d}|. \tag{36}$$

Let $i_1 = i/r \pmod d$ and $i_2 = (d-i)/r \pmod d$, then

$$\begin{aligned} i_1 + i_2 \equiv& \frac{i}{r} + \frac{d-i}{r} \pmod d \\ \equiv& \frac{d}{r} \pmod d. \end{aligned}$$

Suppose $d/r \pmod d \equiv x$ for some integer x, then

$$rx \equiv d \pmod d. \tag{37}$$

Since d is a odd prime number and $r < d$, then r cannot divide d, so $x = 0$ or d. Considering the fact that $1 \le i_1 \le d-1$ and $1 \le i_2 \le d-1$, we know $x = i_1 + i_2 = d$. Hence, conjugation by U will transform $|x_i\rangle\langle x_{d-i}|$ into another term of the same form. Now suppose for $1 \le i < i' \le d-1$ and $U|x_i\rangle\langle x_{d-i}|U^\dagger = U|x_i'\rangle\langle x_{d-i'}|U^\dagger$, it means that $i/r \equiv i'/r \pmod d$, which further implies that $i = i'$. To conclude, conjugation by U only permutes the terms of B_2, so

$$UB_2U^\dagger = B_2. \tag{38}$$

The decomposition of U into B_3 and B_4 is similar to the decomposition of X. Suppose the eigen-decomposition of U is

$$U = \sum_{i=0}^{d-1} \omega_{d-1}^i|u_i\rangle\langle u_i|, \tag{39}$$

then the decomposition is the following

$$B_3 = |u_0\rangle\langle u_0| + \omega_{d-1}^{(d-1)/2}|u_{(d-1)/2}\rangle\langle u_{(d-1)/2}| + \sum_{k=1}^{(d-3)/2}\left(\omega_d^k|u_k\rangle\langle u_{d-1-k}| + \omega_d^{-k}|x_{d-1-k}\rangle\langle x_k|\right) \quad (40)$$

$$B_4 = |u_0\rangle\langle u_0| + |u_{(d-1)/2}\rangle\langle u_{(d-1)/2}| + \sum_{k=1}^{(d-3)/2}\left(|u_{d-1-k}\rangle\langle u_k| + |u_k\rangle\langle u_{d-1-k}|\right). \quad (41)$$

In the next step of embedding, for each x_i, we find y_i ,z_i, w_i and f such that $x_i = y_i z_i = f w_i$. We demonstrate how this splitting is done by pick x_1 as an example, which is mapped to B_1 in the previous step. Now we map x_1 to B_1' which is

$$B_1' = \begin{pmatrix} B_1 & 0 \\ 0 & B_1 \end{pmatrix}, \quad (42)$$

then y_1, z_1, w_1 and f are mapped to

$$y_1 \to \begin{pmatrix} B_1 & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \qquad z_1 \to \begin{pmatrix} \mathbb{1} & 0 \\ 0 & B_1 \end{pmatrix}, \quad (43)$$

$$w_1 \to \begin{pmatrix} 0 & B_1 \\ B_1 & 0 \end{pmatrix}, \qquad f \to \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \quad (44)$$

The last step of embedding follows the same manner as we double the dimension of the operators again. For example, y_1 is mapped to $B_1 \oplus \mathbb{1} \oplus B_1 \oplus \mathbb{1}$ and z_1 is mapped to $\mathbb{1} \oplus B_1 \oplus \mathbb{1} \oplus B_1$, so effectively, x_1 is mapped to $B_1^{\oplus 4}$. [1]

In the end, we determine the shared state. The optimal correlation for $CHSH_X^{(d)}$ uses state $|\psi\rangle = \frac{1}{\sqrt{d-1}}\sum_{i=1}^{d-1}|u_i\rangle|u_i\rangle$, so the shared state for G is $\frac{1}{2}(|\psi\rangle^{\oplus 4})$. □

Next we are going to prove that the correlation $C(d^{(r)})$, which is produced by the strategy winning this game optimally, can self-test $d-1$-dimensional EPR pair.

5.2 Self-test

This subsection needs to be formalized. We give a formal version of **??** first and then prove it.

Theorem 14. *If a quantum strategy using the shared state $|\psi\rangle$ achieves the ideal 2-party correlation $C(d^{(r)})$ where d is an odd prime number with primitive root $r \in \{2,3,5\}$, then there exist local isometries Φ_A and Φ_B, and a state $|junk\rangle$ such that $\Phi_A \otimes \Phi_B|\psi\rangle = |EPR^{(d-1)}\rangle \otimes |junk\rangle$.*

Proof. Suppose Alice and Bob achieve the optimal correlation with the quantum strategy $(|\psi\rangle, \{A_x\}, \{B_y\})$ for all $x, y \in \mathcal{X} \times \mathcal{Y}$. The observables A_x and B_y and the shared state $|\psi\rangle$ shall not be confused with the ones used in the optimal strategy. By Lemma 4.3 of Ref. [CS17], we can extract an operator solution from the perfect winning strategy of the linear system game LS_r. For each variable $\{x_i\}_{i=1}^{n_r}$, Alice and Bob has operators A_i and B_i respectively. The condition that they agree with assignment to variables means that

$$\langle\psi|A_i \otimes \overline{B_i}|\psi\rangle = 1 \Rightarrow A_i \otimes \overline{B_i}|\psi\rangle = |\psi\rangle \text{ for } 1 \le i \le n_r \quad (45)$$

[1] Details of this step of embedding can be found in proof of Proposition 4.2 in Ref. [Slo17].

and the condition that Alice's assignments satisfy the constraint means that

$$\mathrm{Tr}(\rho_A \Pi_{j:H(i,j)\neq 0} A_j) = \mathrm{Tr}(\rho_A) \text{ for all } 1 \leq i \leq m_r \tag{46}$$

where $\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|)$. Similarly we define $\rho_B = \mathrm{Tr}_A(|\psi\rangle\langle\psi|)$. For any $|v\rangle \in \mathrm{supp}(\rho_A)$, we have

$$\Pi_{j:H(i,j)\neq 0} A_j |v\rangle = |v\rangle \text{ for all } 1 \leq i \leq m_r. \tag{47}$$

Since the relation $uxu^{-1} = x^r$, where r is the primitive root of d, is embedded in this linear system game, we know

$$A_3A_4A_1A_2(A_3A_4)^\dagger |v\rangle = (A_1A_2)^r |v\rangle \text{ for all } |v\rangle \in \mathrm{supp}(\rho_A). \tag{48}$$

For simplicity, we define $X_A = A_1A_2$ and $U_A = A_3A_4$ such that the condition is equivalent to

$$U_A X_A U_A^\dagger |v\rangle = X_A^r |v\rangle \text{ for all } |v\rangle \in \mathrm{supp}(\rho_A). \tag{49}$$

We will come back to the implication of this condition later.

Suppose $A_{m_r+1} = A_*^\diamond - A_*^\perp = \Pi_{V_A} - \Pi_{V_A^\perp}$m, where V_A is a $2m$-dimensional vector space and Π_{V_A} is the projector onto it. The reason why it has dimension $2m$ will be clear shortly. On Bob's side, we also have $B_{n_r+1} = B_*^\diamond - B_*^\perp = \Pi_{V_B} - \Pi_{V_B^\perp}$. Recall that from the extended weighted CHSH test, we know

$$A_*^\diamond |\psi\rangle = B_*^\diamond |\psi\rangle = A_*^\diamond \otimes B_*^\diamond |\psi\rangle, \tag{50}$$

which implies that with an appropriate change of basis, we can get $V_A = V_B$, so in the rest of the proof we drop the subscript of V. By Lemma 12 and Proposition 11, we know V consists of ω_d-eigenvectors and ω_d^{-1} eigenvectors of X_A, so we can write

$$\Pi_V = \Pi_{V_1} + \Pi_{V_{d-1}}, \tag{51}$$

where Π_{V_1} is the projector onto the ω_d-eigenspace of X_A and $\Pi_{V_{d-1}}$ is the projector onto the ω_d^{-1}-eigenspace of X_A. Suppose $|x_1\rangle \in V_1$, then $X_A|x_1\rangle = \omega_d |x_1\rangle$. By eq. (49) we can calculate that

$$X_A U_A^\dagger |x_1\rangle = U_A^\dagger X_A^r |x_1\rangle = \omega_d^r U_A^\dagger |x_1\rangle, \tag{52}$$

so $U_A^\dagger |x_1\rangle$ is an eigenvector of X_A with eigenvalue ω_d^r. By induction, we know $X_A (U_A^\dagger)^i |x_1\rangle = \omega_d^{ri} (U_A^\dagger)^i |x_1\rangle$. From the set $\{(U_A^\dagger)^i \Pi_{V_1} (U_A)^i\}_{i=0}^{d-2}$, we can identify Π_{V_i} for $i = 1 \ldots d-1$ such that Π_{V_i} is the projector onto the ω_d^i-eigenspace of X_A, and

$$\cup_{i\in[d-1]+1} V_i \subset \mathrm{supp}(\rho_A). \tag{53}$$

Since unitary transformation does not change the rank of a matrix, we know $\mathrm{rank}(\Pi_{V_1}) = \mathrm{rank}(\Pi_{V_i}) = N$ for $i = 1 \ldots d-1$. We pick a basis for V_1 such that

$$\Pi_{V_1} = \sum_{j=1}^{m} |x_{1,j}\rangle\langle x_{1,j}|, \tag{54}$$

then we can construct

$$\Pi_{V_i} = \sum_{j=1}^{m} |x_{i,j}\rangle\langle x_{i,j}|, \tag{55}$$

where $|x_{i,j}\rangle = (U^\dagger)^{k_i}|x_{1,j}\rangle$ for $r^{k_i} \equiv i \pmod d$ and $1 \le j \le m$. By eq. (49) we also know that $U_A|x_{i,j}\rangle = |x_{i/2,j}\rangle$ for $i = 1, 2 \dots d-1$. In order to apply Lemma 9, we construct m subspaces $\{W_j\}_{j=1}^m$ where

$$W_j = \text{span}(\{|x_{i,j}\rangle\}_{i=1}^{d-1}) \tag{56}$$

The subspace W_j is orthogonal to $W_{j'}$ for $j \neq j'$, and U_A and X_A satisfy the condition of Lemma 9 when their actions are restricted to each W_j. Similar argument also applies to operator X_B and U_B on Bob's side.

With an appropriate change of basis, we assume that $|\psi\rangle = \text{vec}(\tau)$ for some $\tau \in L(\text{supp}(\rho_A))$ The consistency condition is equivalent to

$$A_i \tau B_i^\dagger = \tau. \tag{57}$$

Substituting $i = 1, 2$ into eq. (57), we get

$$X_A \tau X_A^\dagger = A_1 A_2 \tau B_2^\dagger B_1^\dagger = A_1 \tau B_1^\dagger = \tau. \tag{58}$$

Similar argument gives us that

$$U_A \tau U_A^\dagger = \tau. \tag{59}$$

Then we can conclude that for any $k \in \{0, 1 \dots d-2\}$ and $l \in \{1, 2 \dots d-1\}$

$$U_A^k X_A^l \tau (U_A^k X_A^l)^\dagger = \tau. \tag{60}$$

Let Π_{W_j} be the projector onto W_j. By Lemma 9, $\Pi_{W_j} \tau \Pi_{W_j}$ commutes with all the $(d-1) \times (d-1)$ matrices, which means that

$$\Pi_{W_j} \tau \Pi_{W_j} = c_j \mathbb{1}_{W_j} \text{ for } j = 1 \dots m. \tag{61}$$

In the vector form, we know

$$\sum_{j=1}^{N} \Pi_{W_j} |\psi\rangle = \sum_{i=1}^{d-1} \sum_{j=1}^{m} c_j |x_{i,j}\rangle |x_{i,j}\rangle. \tag{62}$$

Recalling the fact that

$$\|\Pi_{V_1} + \Pi_{V_{d-1}}|\psi\rangle\|^2 = \|\Pi_{V_1}|\psi\rangle\|^2 + \|\Pi_{V_{d-1}}|\psi\rangle\|^2 = \frac{2}{d-1}, \tag{63}$$

it means that

$$2 \sum_{j=1}^{m} \|c_j\|^2 = \frac{2}{d-1}. \tag{64}$$

Then we can calculate the norm of $\sum_{j=1}^m \Pi_{W_j}|\psi\rangle$, which is

$$\|\sum_{j=1}^{m} \Pi_{W_j}|\psi\rangle\|^2 = \sum_{i=1}^{d-1} \sum_{j=1}^{m} \|c_j\|^2 = 1 = \||\psi\rangle\|^2. \tag{65}$$

We can conclude that $\operatorname{supp}\rho_A = \cup_{i=1}^{d-1} V_i$ and

$$|\psi\rangle = \sum_{i=1}^{d-1}\sum_{j=1}^{m} c_j |x_{i,j}\rangle|x_{i,j}\rangle. \tag{66}$$

The last step is to construct the local isometries Φ_A and Φ_B that can produce $|EPR^{(d-1)}\rangle$ from $|\psi\rangle$. We give steps about how Φ_A and Φ_B works. Since the subspaces W_j's are orthogonal, Alice and Bob can map state $|\psi\rangle$ to the state

$$|\psi'\rangle = \sum_{i=1}^{d-1}\sum_{j=1}^{m} c_j |x_{i,j}\rangle_A |x_{i,j}\rangle_B |j\rangle_A |j\rangle_B. \tag{67}$$

Note that we added subscript A and B to stress the subsystems holden by Alice and Bob respectively. Since $|x_{i,j}\rangle$ is orthogonal to $|x_{i',j'}\rangle$ for $i \neq i'$ or $j \neq j'$, we further assume that there exists a unitary O such that $O|x_{i,j}\rangle = O|x_{i,j+1}\rangle$ for all i and j, but when $j = N$, $O|x_{i,N}\rangle = |x_{i,1}\rangle$. Then controlled by the appended register, Alice and Bob can apply $(O^\dagger)^j$ to the original system and $|\psi'\rangle$ is mapped to

$$\sum_{i=1}^{d-1}\sum_{j=1}^{m} c_j [(O^\dagger)^j |x_{i,j}\rangle_A] \otimes [(O^\dagger)^j |x_{i,j}\rangle_B] \otimes |j\rangle_A |j\rangle_B \tag{68}$$

$$= \sum_{i=1}^{d-1}\sum_{j=1}^{m} c_j |x_{i,1}\rangle_A |x_{i,1}\rangle_B |j\rangle_A |j\rangle_B \tag{69}$$

$$= \frac{1}{\sqrt{d}} \sum_{i=1}^{d-1} |x_{i,1}\rangle_A |x_{i,1}\rangle_B \otimes \sum_{j=1}^{m} \sqrt{d} c_j |j\rangle_A |j\rangle_B \tag{70}$$

After relabelling $|x_{i,1}\rangle$ as $|i\rangle$, we factor $|EPR^{(d-1)}\rangle$ out of $\Phi_A \otimes \Phi_B |\psi\rangle$. □

In summary, for any odd prime number d whose primitive root is 2,3 or 5, Proposition 13 tells us that the correlation $C(d)$ is achievable by a quantum strategy and Theorem 14 tells us that the correlation $C(d)$ can self-test the EPR pair of local dimension $d-1$. Moreover, there are infinitely many prime numbers whose primitive root is in the set $\{2, 3, 5\}$ [Mur88], Hence, our main result is the following theorem.

Theorem 15. *There exists an infinity-sized set D of prime numbers such that each $d \in D$ has primitive root 2, 3 or 5 and there exists a constant-sized correlation $C(d)$ that can self-test the EPR pair of local dimension $d-1$.*

We remark that our proof works for any odd prime number with primitive root r. However, since there is no upper-bound of r for a general prime number d, we cannot claim the corresponding correlation $C(d)$ is constant-sized. On the other hand, since r is usually much smaller than d, our result implies a more efficient way to test EPR pairs of prime local dimension than the one proposed in Ref. [CGS17]. The size of our correlation grows linearly in r whereas the Coladangelo *et. al.*'s method uses correlation grows linearly in d.

References

[AFDF+18] Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick. Practical device-independent quantum cryptography via entropy accumulation. *Nature communications*, 9(1):459, 2018.

[AMP12] Antonio Acín, Serge Massar, and Stefano Pironio. Randomness versus nonlocality and entanglement. *Physical review letters*, 108(10):100402, 2012.

[BP15] Cédric Bamps and Stefano Pironio. Sum-of-squares decompositions for a family of clauser-horne-shimony-holt-like inequalities and their application to self-testing. *Physical Review A*, 91(5):052111, 2015.

[CGJV17] Andrea Coladangelo, Alex Grilo, Stacey Jeffery, and Thomas Vidick. Verifier-on-a-leash: new schemes for verifiable delegated quantum computation, with quasilinear resources. *arXiv preprint arXiv:1708.07359*, 2017.

[CGS17] Andrea Coladangelo, Koon Tong Goh, and Valerio Scarani. All pure bipartite entangled states can be self-tested. *Nature communications*, 8:15485, 2017.

[Col17] Andrea Coladangelo. Parallel self-testing of (tilted) epr pairs via copies of (tilted) chsh and the magic square game. *Quantum Information and Computation*, 17(9-10):831–865, 2017.

[CS17] Andrea Coladangelo and Jalex Stark. Robust self-testing for linear constraint system games. *arXiv preprint arXiv:1709.09267*, 2017.

[FM18] Honghao Fu and Carl A Miller. Local randomness: Examples and application. *Physical Review A*, 97(3):032324, 2018.

[McK16] Matthew McKague. Self-testing in parallel. *New Journal of Physics*, 18(4):045013, 2016.

[MPA11] Lluís Masanes, Stefano Pironio, and Antonio Acín. Secure device-independent quantum key distribution with causally independent measurement devices. *Nature communications*, 2:238, 2011.

[MS16] Carl A Miller and Yaoyun Shi. Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices. *Journal of the ACM (JACM)*, 63(4):33, 2016.

[Mur88] M Ram Murty. Artin's conjecture for primitive roots. *The Mathematical Intelligencer*, 10(4):59–67, 1988.

[RUV13] Ben W Reichardt, Falk Unger, and Umesh Vazirani. Classical command of quantum systems. *Nature*, 496(7446):456, 2013.

[Slo17] William Slofstra. The set of quantum correlations is not closed. *arXiv preprint arXiv:1703.08618*, 2017.

[VV14] Umesh Vazirani and Thomas Vidick. Fully device-independent quantum key distribution. *Physical review letters*, 113(14):140501, 2014.

A Proof of Theorem 14

Proof. Following the techniques developed in Ref. [BP15], the first step is to find a sum-of-square decomposition of

$$\bar{\mathcal{I}}_\alpha = 2\sqrt{\alpha^2+1}\mathbb{1} - \mathcal{I}_\alpha = \frac{2}{\sin(\mu)}\mathbb{1} - \frac{\cos(\mu)}{\sin(\mu)}(A_1B_1 + A_1B_2) - A_2B_1 + A_2B_2. \tag{71}$$

With the following notation

$$Z_A = A_1 \quad X_A = A_2$$
$$Z_B = \frac{B_1 + B_2}{2\cos(\mu)} \quad X_B = \frac{B_1 - B_2}{2\sin(\mu)},$$

the two SOS decompositions that we use are

$$\bar{\mathcal{I}}_\alpha = \frac{\sin(\mu)\bar{\mathcal{I}}_\alpha^2 + 4\sin(\mu)\cos(\mu)^2(Z_A X_B + X_A Z_B)^2}{4}, \tag{72}$$

$$\bar{\mathcal{I}}_\alpha = \frac{\cos^2(\mu)}{\sin(\mu)}(Z_A - Z_B)^2 + \sin(\mu)(X_A - X_B)^2. \tag{73}$$

The verification is omitted here.

Suppose the quantum strategy $(|\psi\rangle, \{\tilde{A}_x\}_{x\in[2]}, \{\tilde{B}_{y\in[2]}\}$ achieves that $\langle\psi|\bar{\mathcal{I}}_\alpha|\psi\rangle \leq \epsilon$. The second step is to establish bounds of the following form

$$\|(\tilde{Z}_A - \tilde{Z}_B)|\psi\rangle\| \leq c_1\sqrt{\epsilon} \tag{74}$$
$$\|(\tilde{X}_A(\mathbb{1} + \tilde{Z}_B) - \tilde{X}_B(\mathbb{1} - \tilde{Z}_A))|\psi\rangle\| \leq c_2\sqrt{\epsilon} \tag{75}$$
$$\|(\tilde{X}_A - \tilde{X}_B)|\psi\rangle\| \leq c_3\sqrt{\epsilon} \tag{76}$$
$$\|(\tilde{Z}_A\tilde{X}_A + \tilde{X}_A\tilde{Z}_A)|\psi\rangle\| \leq c_4\sqrt{\epsilon}. \tag{77}$$

Now we write $s = \sin(\mu)$, $c = \cos(\mu)$ and define

$$S_1 = \frac{\sqrt{s}}{2}\bar{\mathcal{I}}_\alpha, \quad S_2 = \sqrt{s}c(\tilde{Z}_A\tilde{X}_B + \tilde{X}_A\tilde{Z}_B),$$
$$S_3 = \frac{c}{\sqrt{s}}(\tilde{Z}_A - \tilde{Z}_B), \quad S_4 = \sqrt{s}(\tilde{X}_A - \tilde{X}_B)$$

then $\bar{\mathcal{I}}_\alpha = S_1^2 + S_2^2 = S_3^2 + S_4^2$ and $\langle\psi|\bar{\mathcal{I}}_\alpha|\psi\rangle \leq \epsilon$ implies that $\langle\psi|S_i^2|\psi\rangle \leq \epsilon$ and $\|S_i|\psi\rangle\| \leq \sqrt{\epsilon}$ for $i = 1, 2, 3, 4$. We can easily check that

$$c_1 = \frac{\sqrt{s}}{c}, \quad c_2 = \frac{1}{\sqrt{s}} + \frac{1}{c\sqrt{s}}, \quad c_3 = \frac{1}{\sqrt{s}}.$$

where we use the relation that $\tilde{X}_A(\mathbb{1} + \tilde{Z}_B) - \tilde{X}_B(\mathbb{1} - \tilde{Z}_A) = S_4/s^{1/2} + S_2/(cs^{1/2})$. To calculate c_4, we use the relation

$$\tilde{Z}_A\tilde{X}_A + \tilde{X}_A\tilde{Z}_A = \frac{S_2}{c\sqrt{s}} + \frac{\sqrt{s}\tilde{X}_A S_3}{c} + \frac{\tilde{Z}_A S_4}{\sqrt{s}} \tag{78}$$

and reach the conclusion that

$$c_4 = \frac{1 + c + s}{c\sqrt{s}} \tag{79}$$

where we use that fact that $\tilde{Z}_A, \tilde{X}_A$ are unitaries.

With appropriate substitutions, the rest of the proof follows the same derivation as that in Appendix A of Ref. [BP15], so we omit it here. The dependence of the error term on α comes from the fact that

$$\frac{1}{\cos(\arctan(1/\alpha))\sin^{1/2}(\arctan(1/\alpha))} = \frac{1}{\alpha} + \frac{3}{4}\alpha - O(\alpha^3). \tag{80}$$

□

B Equations to embed $uxu^{-1} = x^2$

We have order-2 generators $\{x_i, w_i, y_i, j_i\}_{i=1}^7 \cup \{f\} \cup \{g_{I_i}\}_{i=1}^5$, where each I_i represent a tuple of the form (i, j, k) and it means that $x_i x_j x_i = x_k$. We have

$$I_1 = (1,2,5), \qquad I_2 = (3,5,7), \qquad I_3 = (4,1,7), \qquad I_4 = (4,2,6), \qquad I_5 = (3,2,6) \tag{81}$$

and we denote $C = \{I_i\}_{i=1}^5$. The linear relations involving the generators listed above are

$$x_i y_i z_i = e \qquad \text{for} \quad i = 1 \dots 7 \tag{82}$$
$$x_i f w_i = e \qquad \text{for} \quad i = 1 \dots 7 \tag{83}$$
$$g_{I_1} y_2 z_5 = e \tag{84}$$
$$g_{I_2} y_5 z_7 = e \tag{85}$$
$$g_{I_3} y_1 z_7 = e \tag{86}$$
$$g_{I_4} y_2 z_6 = e \tag{87}$$
$$g_{I_5} y_2 z_6 = e \tag{88}$$

For $j = 1 \dots 7$ we introduce a set of order-2 generators $\{y_{F_j,k}\}_{k=1}^6$ and a group of linear relations of the form

$$f y_{F_j,1} y_{F_j,2} = e, \qquad f y_{F_j,5} y_{F_j,6} = e \tag{89}$$
$$y_j y_{F_j,2} y_{F_j,3} = e, \qquad z_j y_{F_j,6} y_{F_j,7} = e \tag{90}$$
$$y_{F_j,1} y_{F_j,4} y_{F_j,7} = e, \qquad y_{F_j,3} y_{F_j,4} y_{F_j,5} = e. \tag{91}$$

We refer to each group of linear relations of the from above as (F_j).

In the end, for each $K \in C$, we introduce a set of order-2 generators $\{y_{K,k}\}_{k=1}^6$ and linear relations. Suppose $K = (k_1, k_2, k_3)$, then the group of linear relations are

$$w_{k_1} y_{K,1} y_{K,2} = e, \qquad w_{k_1} y_{K,5} y_{K,6} = e \tag{92}$$
$$y_{k_2} y_{K,2} y_{K,3} = e, \qquad z_{k_3} y_{K,6} y_{K,7} = e \tag{93}$$
$$y_{K,1} y_{K,4} y_{K,7} = e, \qquad y_{K,3} y_{K,4} y_{K,5} = e, \tag{94}$$

We refer to each group of linear relations as (K). In summary, we have 118 binary generators, which are

$$\{x_i, w_i, y_i, j_i\}_{i=1}^7 \cup \{f\} \cup \{g_{I_i}\}_{i=1}^5 \cup \{\{y_{F_j,k}\}_{k=1}^6\}_{j=1}^7 \cup \{\{y_{K,k}\}_{k=1}^6\}_{K \in C}. \tag{95}$$

In total, there are 91 linear relations. Converting a linear relation to a linear equation is trivial, so we omit it here.

From the group of linear relations (F_j), we can deduce that for $j = 1 \dots 7$

$$f y_j f = (y_{F_j,1} y_{F_j,2})(y_{F_j,2} y_{F_j,3})(y_{F_j,5} y_{F_j,6}) \tag{96}$$
$$= y_{F_j,1} (y_{F_j,3} y_{F_j,5}) y_{F_j,6} \tag{97}$$
$$= (y_{F_j,1} y_{F_j,4}) y_{F_j,6} \tag{98}$$
$$= y_{F_j,7} y_{F_j,6} \tag{99}$$
$$= z_j. \tag{100}$$

The immediate implication of the group of linear relations (K) is that

$$\begin{aligned} w_{k_1} y_{k_2} w_{k_1} =& (y_{K,1} y_{K,2})(y_{K,2} y_{K,3})(y_{K,5} y_{K,6}) && (101)\\ =& y_{K,1}(y_{K,3} y_{K,5}) y_{K,6} && (102)\\ =& (y_{K,1} y_{K,4}) y_{K,6} && (103)\\ =& y_{K,7} y_{K,6} && (104)\\ =& z_{k_3}. && (105) \end{aligned}$$

From the conjugacy relation above, we can first deduce that

$$w_{k_1} z_{k_2} w_{k_1} = w_{k_1}(f y_{k_2} f) w_{k_1} = f(w_{k_1} y_{k_2} w_{k_1}) f = f z_{k_3} f = y_{k_3} \tag{106}$$

where we use the fact that $(f w_{k_1})^2 = x_{k_1}^2 = e$ from eq. (128). Then we can reason the relation between x_{k_1}, x_{k_2} and x_{k_3} as follows

$$x_{k_1} x_{k_2} x_{k_1} = f w_{k_1} y_{k_2} z_{k_2} f w_{k_1} = (f w_{k_1} y_{k_2} w_{k_1} f)(f w_{k_1} z_{k_2} w_{k_1} f) = (f z_{k_3} f)(f y_{k_3} f) = y_{k_3} z_{k_3} = x_{k_3}, \tag{107}$$

where we repeated use eq. (127) and eq. (128). In the end, we show why I_1,I_2,I_3,I_4 and I_5 imply that

$$\begin{aligned} x_3 x_4 x_1 x_2 x_4 x_3 =& x_3(x_4 x_1 x_4)(x_4 x_2 x_4) x_3 && (108)\\ =& (x_3 x_7 x_3)(x_3 x_6 x_3) && (109)\\ =& x_5 x_2 && (110)\\ =& x_1 x_2 x_1 x_2 && (111) \end{aligned}$$

where we used the fact that x_i's are of order-2. If we treat $x_3 x_4$ as u and $x_1 x_2$ as x, then we have derived that $uxu^{-1} = x^2$.

C Equations to embed $uxu^{-1} = x^3$

We have order-2 generators $\{x_i, w_i, y_i, j_i\}_{i=1}^{8} \cup \{f\} \cup \{g_{I_i}\}_{i=1}^{6}$, where each I_i represent a tuple of the form (i, j, k) and it means that $x_i x_j x_i = x_k$. We have

$$\begin{aligned} &I_1 = (2,1,5), && I_2 = (1,3,6), && I_3 = (4,2,7),\\ &I_4 = (3,2,7), && I_5 = (4,1,8), && I_6 = (3,6,8) \end{aligned}$$

and we collect them in a set $C = \{I_i\}_{i=1}^{6}$. The linear relations involving the generators listed above are

$$\begin{aligned} x_i y_i z_i &= e && \text{for} \quad i = 1 \ldots 8 && (112)\\ x_i f w_i &= e && \text{for} \quad i = 1 \ldots 8 && (113)\\ g_{I_1} y_1 z_5 &= e && && (114)\\ g_{I_2} y_3 z_6 &= e && && (115)\\ g_{I_3} y_2 z_7 &= e && && (116)\\ g_{I_4} y_2 z_7 &= e && && (117)\\ g_{I_5} y_1 z_8 &= e && && (118)\\ g_{I_6} y_6 z_8 &= e && && (119) \end{aligned}$$

To introduce the following linear relations, we introduce another set of generators $\{\{y_{F_j,k}\}_{k=1}^{6}\}_{j=1}^{8}$. For $j = 1 \dots 8$ we have a group of linear relations of the form

$$fy_{F_j,1}y_{F_j,2} = e, \qquad fy_{F_j,5}y_{F_j,6} = e \tag{120}$$

$$y_j y_{F_j,2} y_{F_j,3} = e, \qquad z_j y_{F_j,6} y_{F_j,7} = e \tag{121}$$

$$y_{F_j,1}y_{F_j,4}y_{F_j,7} = e, \qquad y_{F_j,3}y_{F_j,4}y_{F_j,5} = e, \tag{122}$$

We refer to each group of linear relations of the from above as (F_j).

The last set of linear relations comes from each $K \in C$, so we introduce another set of generators $\{\{y_{K,k}\}_{k=1}^{6}\}_{K\in C}$. We have a group of linear relations for each $K \in C$. Suppose $K = (k_1, k_2, k_3)$, then the group of linear relations are

$$w_{k_1}y_{K,1}y_{K,2} = e, \qquad w_{k_1}y_{K,5}y_{K,6} = e \tag{123}$$

$$y_{k_2}y_{K,2}y_{K,3} = e, \qquad z_{k_3}y_{K,6}y_{K,7} = e \tag{124}$$

$$y_{K,1}y_{K,4}y_{K,7} = e, \qquad y_{K,3}y_{K,4}y_{K,5} = e, \tag{125}$$

We refer to each group of linear relations as (K). In summary, we have 123 binary generators, which are

$$\{x_i, w_i, y_i, j_i\}_{i=1}^{8} \cup \{f\} \cup \{g_{I_i}\}_{i=1}^{6} \cup \{\{y_{F_j,k}\}_{k=1}^{6}\}_{j=1}^{8} \cup \{\{y_{K,k}\}_{k=1}^{6}\}_{K\in C}. \tag{126}$$

There are 106 linear relations. Converting a linear relation to a linear equation is trivial, so we omit it here.

The derivation of $fy_jf = z_j$ from the group of linear relations $(F_j)$for $j = 1 \dots 8$ is the same as the previous case, so we omit it here. We also omit the derivation of $x_{k_1}x_{k_2}x_{k_1} = x_{k_3}$, from the group of linear relations (K), for each $K \in C$. In the end, we show the implication of the conjugacy relations in C as follows

$$\begin{aligned} x_3x_4x_1x_2x_4x_3 =& x_3(x_4x_1x_4)(x_4x_2x_4)x_3 \\ =& (x_3x_8x_3)(x_3x_7x_3) \\ =& x_6x_2 \\ =& x_1x_5x_1x_2 \\ =& x_1x_2x_1x_2x_1x_2 \end{aligned}$$

where we used the fact that x_i's are of order-2. If we treat x_3x_4 as u and x_1x_2 as x, then we have derived that $uxu^{-1} = x^3$.

D Equations to embed $uxu^{-1} = x^5$

We have order-2 generators $\{x_i, w_i, y_i, j_i\}_{i=1}^{10} \cup \{f\} \cup \{g_{I_i}\}_{i=1}^{8}$, where each I_i represent a tuple of the form (i, j, k) and it means that $x_ix_jx_i = x_k$. We have

$$I_1 = (2,1,5), \quad I_2 = (1,3,6), \quad I_3 = (2,6,7), \quad I_4 = (1,7,8),$$
$$I_5 = (4,1,9), \quad I_6 = (3,8,9), \quad I_7 = (4,2,10) \quad I_8 = (3,2,10)$$

and we collect them in a set $C = \{I_i\}_{i=1}^{8}$. The linear relations involving the generators listed above are

$$x_i y_i z_i = e \qquad \text{for} \quad i = 1 \dots 10 \tag{127}$$
$$x_i f w_i = e \qquad \text{for} \quad i = 1 \dots 10 \tag{128}$$
$$g_{I_1} y_1 z_5 = e \tag{129}$$
$$g_{I_2} y_3 z_6 = e \tag{130}$$
$$g_{I_3} y_6 z_7 = e \tag{131}$$
$$g_{I_4} y_7 z_8 = e \tag{132}$$
$$g_{I_5} y_1 z_9 = e \tag{133}$$
$$g_{I_6} y_8 z_9 = e \tag{134}$$
$$g_{I_7} y_2 z_{10} = e \tag{135}$$
$$g_{I_8} y_2 z_{10} = e \tag{136}$$

To introduce the following linear relations, we introduce another set of generators $\{\{y_{F_j,k}\}_{k=1}^{6}\}_{j=1}^{10}$. For $j = 1 \dots 10$ we have a group of linear relations of the form

$$f y_{F_j,1} y_{F_j,2} = e, \qquad f y_{F_j,5} y_{F_j,6} = e \tag{137}$$
$$y_j y_{F_j,2} y_{F_j,3} = e, \qquad z_j y_{F_j,6} y_{F_j,7} = e \tag{138}$$
$$y_{F_j,1} y_{F_j,4} y_{F_j,7} = e, \qquad y_{F_j,3} y_{F_j,4} y_{F_j,5} = e, \tag{139}$$

We refer to each group of linear relations of the from above as (F_j).

The last set of linear relations comes from each $K \in C$, so we introduce another set of generators $\{\{y_{K,k}\}_{k=1}^{6}\}_{K \in C}$. We have a group of linear relations for each $K \in C$. Suppose $K = (k_1, k_2, k_3)$, then the group of linear relations are

$$w_{k_1} y_{K,1} y_{K,2} = e, \qquad w_{k_1} y_{K,5} y_{K,6} = e \tag{140}$$
$$y_{k_2} y_{K,2} y_{K,3} = e, \qquad z_{k_3} y_{K,6} y_{K,7} = e \tag{141}$$
$$y_{K,1} y_{K,4} y_{K,7} = e, \qquad y_{K,3} y_{K,4} y_{K,5} = e, \tag{142}$$

We refer to each group of linear relations as (K). In summary, we have 157 binary generators, which are

$$\{x_i, w_i, y_i, j_i\}_{i=1}^{10} \cup \{f\} \cup \{g_{I_i}\}_{i=1}^{8} \cup \{\{y_{F_j,k}\}_{k=1}^{6}\}_{j=1}^{10} \cup \{\{y_{K,k}\}_{k=1}^{6}\}_{K \in C}. \tag{143}$$

There are 136 linear relations. Converting a linear relation to a linear equation is trivial, so we omit it here.

The derivation of $f y_j f = z_j$ from the group of linear relations $(F_j)$for $j = 1 \dots 10$ is the same as the previous case, so we omit it here. We also omit the derivation of $x_{k_1} x_{k_2} x_{k_1} = x_{k_3}$, from the group of linear relations (K), for each $K \in C$. In the end, we show the implication of the conjugacy

relations in C as follows

$$\begin{aligned} x_3x_4x_1x_2x_4x_3 =& x_3(x_4x_1x_4)(x_4x_2x_4)x_3 \\ =& (x_3x_9x_3)(x_3x_{10}x_3) \\ =& x_8x_2 \\ =& (x_1x_7x_1)x_2 \\ =& x_1(x_2x_6x_2)x_1x_2 \\ =& x_1x_2(x_1x_5x_1)x_2x_1x_2 \\ =& x_1x_2x_1(x_2x_1x_2)x_1x_2x_1x_2 \end{aligned}$$

where we used the fact that x_i's are of order-2. If we treat x_3x_4 as u and x_1x_2 as x, then we have derived that $uxu^{-1} = x^5$.