Entanglement swapping in generalized probabilistic theories and iterated Clauser-Horne-Shimony-Holt games

Lionel J. Dmello,[*] Laurens T. Ligthart [ORCID],[†] and David Gross[‡]

Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany

While there exist theories that have states more strongly entangled than quantum theory, in the sense that they show Clauser-Horne-Shimony-Holt (CHSH) values above Tsirelson's bound, all known examples of such theories have a strictly smaller set of measurements. Therefore, in tasks that require both bipartite states and measurements, they do not perform better than quantum mechanics. One of the simplest information processing tasks involving both bipartite states and measurements is that of entanglement swapping. In this paper, we study entanglement swapping in generalized probabilistic theories (GPTs). In particular, we introduce the iterated CHSH game, which measures the power of a GPT to preserve nonclassical correlations, in terms of the largest CHSH value obtainable after n rounds of entanglement swapping. Our main result is the construction of a GPT that achieves a CHSH value of 4 after an arbitrary number of rounds. This addresses a question about the optimality of quantum theory for such games recently raised by Weilenmann and Colbeck. One challenge faced when treating this problem is that there seems to be no general framework for constructing GPTs in which entanglement swapping is a well-defined operation. Therefore, we introduce an algorithmic construction that turns a bipartite GPT into a multipartite GPT that supports entanglement swapping, if consistently possible.

I. INTRODUCTION

Finding a set of operationally motivated axioms that single out quantum mechanics (QM) has been a longstanding problem in the field of foundations of quantum mechanics. Such an undertaking requires a mathematical framework that allows us to compare QM with other theories, such as classical theory. Arguably the most general framework is that of generalized probabilistic theories (GPTs) [1–9].

Within this framework, a large number of axioms have been proposed over the years [10–16]. One of the main foci of many investigations is the behavior of bipartite theories. In particular, explaining why Clauser-Horne-Shimony-Holt (CHSH)-type experiments [17,18] in nature are bounded by $2\sqrt{2}$, which Tsirelson [19] famously showed to be the largest value allowed by quantum theory. But, as it stands, we still have no definitive axiom singling out quantum theory based on this property.

In their recent work [20], Weilenmann and Colbeck turn to multipartite theories to possibly explain $2\sqrt{2}$. The argument hinges on a well-known tension between the set of states and measurements in a theory: Expanding state space to include more correlations requires one to shrink the set of measurements in order to avoid the emergence of negative probabilities.

For example, bipartite boxworld [8,9,21] state space is the biggest possible bipartite state space one could make. But, as a consequence, the bipartite effect space (the space from which the measurements are constructed) has only product effects. This in turn, is the smallest possible bipartite effect space.

Therefore, by adding a step of entanglement swapping before playing the CHSH game, the set of possible correlations achievable in theories with such state spaces shrink (see Refs. [22,23]). Quantum theory seems to strike the optimum between these competing notions, in that the CHSH violation of $2\sqrt{2}$ is preserved under entanglement swapping. And indeed, Weilenmann and Colbeck show that, for the adaptive CHSH game [24], the winning probability of any theory whose unipartite state and effect spaces are characterized by regular polygons [25] is upper bounded by the winning probability of quantum theory.

Naturally, the question arises whether quantum theory is optimal, in the sense that no GPT can sustain a CHSH violation greater than $2\sqrt{2}$ through entanglement swapping. To investigate this, we introduce the iterated CHSH game, which is an extension of the adaptive CHSH game to multiple rounds. The main result of this paper is to answer this question in the negative, with the construction of a GPT, which sustains the maximal possible value of 4 over arbitrarily many rounds in the iterated CHSH game. Along the way, we also discuss a simpler construction, which sustains a CHSH violation of 4 only for a finite number of rounds.

A. Generalized probabilistic theories

The main idea of GPTs is to model experiments as a two-step process. A preparation step, which produces a state, and a measurement step, which probabilistically maps states

---

[*]Contact author: ldmello@thp.uni-koeln.de
[†]Contact author: ligthart@thp.uni-koeln.de
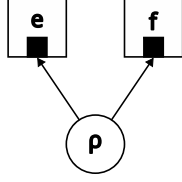[‡]Contact author: david.gross@thp.uni-koeln.de

FIG. 1. Diagrammatic representation of the measurement of a bipartite state by two local observers. Here, the bipartite state ρ is represented by a circle with two arrows emerging from it, one for each subsystem. The unipartite effects e and f on the other hand, are represented by boxes with one port each. The full diagram represents the probability that, on measuring ρ, Alice and Bob get the outcome corresponding to effect e and f, respectively.

to outcomes. Associated to each outcome of a measurement, comes an effect.

Mathematically, states are modeled as members of a convex set (convexity corresponding operationally to probabilistic mixtures). Effects can be viewed as positive linear functionals on states. The pairing between a state ρ and an effect e is interpreted as the probability of obtaining the outcome corresponding to the effect e, given that we prepared the state ρ.

It is useful to represent these operations diagrammatically. For example, the CHSH experiment consists of a bipartite state measured locally by two parties. The resulting contraction can be diagrammatically represented as in Fig. 1. This diagram corresponds to the joint probability obtained by contracting the two-tensor representing the state with the tensor product of the local unipartite effects, i.e., the pairing

$$\langle \rho_{12}, e_1 \otimes f_2 \rangle \tag{1}$$

(see also Sec. II).

B. Entanglement swapping

Contractions of the type depicted in Fig. 1, whose result is a probability, are called full contractions. It is also possible to define partial contraction of states and effects.
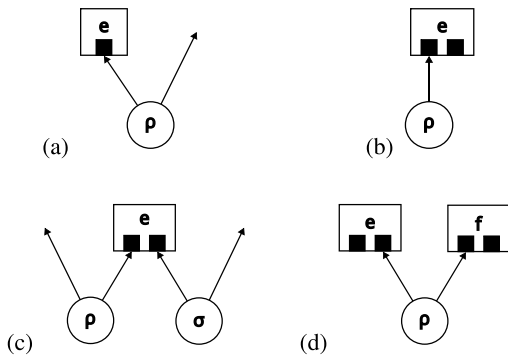

FIG. 2. Operational depiction of various partial contractions on the level of a bipartite theory. (a) depicts partial contraction of a unipartite effect and a bipartite state, (b) depicts partial contraction of a bipartite ieffect and a unipartite state, (c) depicts entanglement swapping, and (d) depicts dual entanglement swapping

Note for example the situation in Fig. 2(a), where a bipartite state ρ is measured only on subsystem 1 and the outcome corresponding to the effect e is obtained. We will interpret the resulting object as a (not necessarily normalized) conditional state, which acts on effects f on the second subsystem as

$$\langle \rho_{12}, e_1 \otimes \mathbb{1}_2 \rangle : f_2 \mapsto \langle \rho_{12}, e_1 \otimes f_2 \rangle. \tag{2}$$

Here $\mathbb{1}$ is the identity channel, which operationally corresponds to "do nothing."

Similarly, one can pair a bipartite effect with a unipartite state. This situation is depicted in Fig. 2(b) and mathematically corresponds to

$$\langle \rho_1 \otimes \mathbb{1}_2, e_{12} \rangle : \sigma_2 \mapsto \langle \rho_1 \otimes \sigma_2, e_{12} \rangle. \tag{3}$$

We can also extend these notions to the case of several bipartite objects. Consider the situation where two bipartite states are partially contracted with a bipartite effect, as depicted in Fig. 2(c). This is the generalization of the quantum-mechanical notion of entanglement swapping to GPTs. These operations will occur so frequently that we introduce a compact notation: For bipartite states ρ, σ and a bipartite effect e, we denote the conditional bipartite state resulting from an entanglement swapping procedure as

$$\llbracket \rho\,^e\sigma \rrbracket := \langle \rho_{12} \otimes \sigma_{34}, \mathbb{1}_1 \otimes e_{23} \otimes \mathbb{1}_4 \rangle. \tag{4}$$

Similarly, we can instead contract two bipartite effects, partially, with a bipartite state, as depicted in Fig. 2(d). This situation we call dual entanglement swapping, and as before, for bipartite effects e, f and a bipartite state ρ, we denote it by

$$\llbracket {}^e\rho\,^f \rrbracket := \langle \mathbb{1}_1 \otimes \rho_{23} \otimes \mathbb{1}_4, e_{12} \otimes f_{23} \rangle. \tag{5}$$

In the following, we generally drop subscripts labeling systems in contractions, if the system an object belongs to is clear from context.

C. Outline

This paper is structured as follows. In Sec. II we state our axioms for a theory under which entanglement swapping is well defined. We also present an algorithm to either show that a given bipartite theory is inconsistent or extend it to a multipartite theory in a consistent manner. In Sec. III we discuss the iterated CHSH game. In Sec. IV we discuss the notion of composite GPTs, which can swap PR-box correlations for a finite number of steps. Finally, in Sec. V we present a GPT, which sustains a maximal violation of Tsirelson's bound indefinitely under entanglement swapping, and provide an optimal strategy for the iterated CHSH game.

II. DEFINITION OF A THEORY

The main goal of this paper is to analyze the phenomenon of entanglement swapping in the general context of GPTs. To this end, it is necessary to specify the conditions on a theory under which entanglement swapping is well defined.

There are a few choices we make therein. First, we will choose the fundamental objects of our theory to be convex cones rather than starting with the state or effect spaces. This is because it is sufficient (and more convenient) to work with cones in the present context, i.e., describing notions such as

entanglement swapping and partial contractions. The state and effects spaces of the theory are obtained as derived objects, in the usual way, as specified below.

Second, we find it a more natural construction to reinterpret the effects as the primal objects and the states as linear functionals on the effects. The two formulations are clearly equivalent. For a detailed exposition of such a construction refer to the review [9].

Definition 1. A theory in which entanglement swapping is well defined is specified by the following:

(i) A finite-dimensional vector space V;

(ii) An element $\underline{1} \in V$, called the unit effect;

(iii) A collection of convex cones $\{\mathcal{P}^{(n)}\}_{n\in\mathbb{N}}$, $\mathcal{P}^{(n)} \subset V^{\otimes n}$, representing (unnormalized) effects, subject to:

    (a) $\underline{1} \in \mathcal{P}^{(1)}$.

    (b) Closure under tensor products, i.e., If $e \in \mathcal{P}^{(n)}$, $f \in \mathcal{P}^{(m)}$, then $e \otimes f \in \mathcal{P}^{(n+m)}$;

(iv) A collection of convex cones $\{\mathcal{D}^{(n)}\}_{n\in\mathbb{N}}$, $\mathcal{D}^{(n)} \subset (V^{\otimes n})^*$, representing (unnormalized) states, subject to:

    (a) Positivity, i.e., $\mathcal{D}^{(n)} \subset (\mathcal{P}^{(n)})'$, the polar dual of $\mathcal{P}^{(n)}$.

    (b) Closure under tensor products, i.e., If $\rho \in \mathcal{D}^{(n)}$, $\sigma \in \mathcal{D}^{(m)}$, then $\rho \otimes \sigma \in \mathcal{D}^{(n+m)}$;

(v) Closure under partial contractions, i.e., For $e \in \mathcal{P}^{(n)}$, $\rho \in \mathcal{D}^{(m)}$:

    (a) If $n > m$, then $\langle \rho, e \rangle \in \mathcal{P}^{(n-m)}$

    (b) If $n < m$, then $\langle \rho, e \rangle \in \mathcal{D}^{(m-n)}$

    (c) (for completeness) If $n = m$, then $\langle \rho, e \rangle \in \mathbb{R}^+$, is the canonical pairing;

(vi) Invariance under permutations of systems, i.e., for every $n \in \mathbb{N}$ and $\pi \in S_n$, we have $\pi(\mathcal{D}^{(n)}) = \mathcal{D}^{(n)}$ and $\pi(\mathcal{P}^{(n)}) = \mathcal{P}^{(n)}$. Where, for $\rho_{1\cdots n} \in \mathcal{D}^{(n)}$

$$\pi(\rho_{1\cdots n}) := \rho_{\pi(1)\cdots\pi(n)}.$$

From the above definition, we can derive the following: For an effect $e \in \mathcal{P}^{(n)}$, define the negation

$$\neg e := \underline{1}^{\otimes n} - e. \tag{6}$$

Define the n-partite effect space as

$$\mathcal{E}^{(n)} := \mathcal{P}^{(n)} \cap \neg(\mathcal{P}^{(n)}). \tag{7}$$

Define the n-partite state space as

$$\mathcal{S}^{(n)} := \{\rho \in \mathcal{D}^{(n)} \mid \langle \rho, \underline{1}^{\otimes n} \rangle = 1\}. \tag{8}$$

Define the set of unipartite correlators as

$$\mathcal{X}^{(1)} := \{e - \neg e \mid e \in \mathcal{E}^{(1)}\}. \tag{9}$$

For every choice $A_0, A_1, B_0, B_1 \in \mathcal{X}^{(1)}$, we define the CHSH observable as

$$\text{CHSH}(A_0, A_1; B_0, B_1)$$
$$:= A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1. \tag{10}$$

The CHSH value corresponding to the above choice of correlators and some choice of $\rho \in \mathcal{S}^{(2)}$ is the expectation value of the CHSH observable with respect to ρ, i.e.,

$$\langle \rho, \text{CHSH}(A_0, A_1; B_0, B_1) \rangle. \tag{11}$$

We can associate with every theory a CHSH value defined to be the supremum over all possible choices ρ, A_0, A_1, B_0, B_1.

$$\sup_{\substack{A_0, A_1, B_0, B_1 \in \mathcal{X}^{(1)} \\ \rho \in \mathcal{S}^{(2)}}} |\langle \rho, \text{CHSH}(A_0, A_1; B_0, B_1) \rangle|. \tag{12}$$

Another important notion in the discussion that follows, is that of closure under entanglement swapping. Given a theory with bipartite cones $\mathcal{P}^{(2)}$ and $\mathcal{D}^{(2)}$, we say $\mathcal{D}^{(2)}$ is closed under entanglement swapping if

$$\llbracket {}_{\mathcal{D}^{(2)}}\mathcal{P}^{(2)}{}_{\mathcal{D}^{(2)}} \rrbracket \subset \mathcal{D}^{(2)}. \tag{13}$$

Further, we say $\mathcal{D}^{(2)}$ is stable under entanglement swapping if

$$\text{conv}\left(\llbracket {}_{\mathcal{D}^{(2)}}\mathcal{P}^{(2)}{}_{\mathcal{D}^{(2)}} \rrbracket\right) = \mathcal{D}^{(2)}. \tag{14}$$

The same notions apply to dual entanglement swapping.

A. Consistent bipartite theory

In this section, we are concerned with the following question. Given:

(i) A finite-dimensional vector space V,

(ii) a nonzero element $\underline{1} \in V$,

(iii) a convex cone $P \subset V \otimes V$,

(iv) a convex cone $D \subset (V \otimes V)^*$,

is it possible to produce a multipartite theory that has the same unit effect, for which $\mathcal{P}^{(2)} = P$ and $\mathcal{D}^{(2)} = D$?

We answer this question by constructing an algorithm that, for an input $n \in \mathbb{N}$, either produces n-partite cones $\mathcal{P}^{(n)}$ and $\mathcal{D}^{(n)}$ by extending P and D, if consistently possible, or detects the inconsistency. The algorithm can be split into two parts. First, a consistency check on the objects $\underline{1}, P$, and D, followed by an explicit construction of $\mathcal{P}^{(n)}$ and $\mathcal{D}^{(n)}$ for every $n \in \mathbb{N}$. The algorithm is as follows:

In the presentation of Algorithm 1, $\dot{\otimes}$ stands for the minimal tensor product, defined as the conal hull over the tensor product of the respective sets:

$$\mathcal{P}^{(n)} \dot{\otimes} \mathcal{P}^{(m)} := \text{cone}(\mathcal{P}^{(n)} \otimes \mathcal{P}^{(m)}). \tag{15}$$

Additionally, we have not specified in which format inputs such as V or P are to be supplied, or how the checks should be performed. In this sense, it is a template for a concrete algorithm that depends on the mathematical properties of the input data. For example, if P, D are polyhedral cones in $\mathbb{R}^d$, then V can be represented by the integer d and the cones by the facet inequalities. In this case, all tests in Algorithm 1 are linear programs. But more general situations also make sense, e.g., cones defined in terms of semidefinite constraints.

Lemma 1. Given the input data $V, \underline{1}, P, D$, Algorithm 1 either detects inconsistency, or returns consistent n-partite cones $\mathcal{P}^{(n)}$ and $\mathcal{D}^{(n)}$ induced by the data, for any $n \in \mathbb{N}$, in a finite number of steps.

Proof. If the function "check consistency" of Algorithm 1 returns "inconsistent," then one of the axioms in Definition 1 is violated already at the level of bipartite objects and so there is no consistent extension of the input.

We will now verify that if "check consistency" returns "consistent," then there exists a theory as advertised. Indeed, the objects whose existence is posited by Axioms 1 and 2 are part of the problem data. Axiom 6 holds by construction given that the bipartite cones are invariant under

ALGORITHM 1. Induced theory.

1: **input:** $(V, \underline{1}, P, D, n \in \mathbb{N})$
2: **if** CHECK CONSISTENCY $(V, \underline{1}, P, D) = $ "Inconsistent"
3:   **return** "Inconsistent"
4: **else**
5:   $\mathcal{D}^{(1)} \leftarrow \text{cone}\langle D, \underline{1}\rangle$ ▷ partial contractions
6:   $\mathcal{P}^{(1)} \leftarrow \text{cone}\langle \mathcal{D}^{(1)}, P\rangle$
7:   $\mathcal{P}^{(2)} \leftarrow P$
8:   $\mathcal{D}^{(2)} \leftarrow D$
9:   **if** n is even **then**
10:      $\mathcal{P}^{(n)} \leftarrow S_n \cdot (\mathcal{P}^{(2)\dot{\otimes}n/2})$
11:      $\mathcal{D}^{(n)} \leftarrow S_n \cdot (\mathcal{D}^{(2)\dot{\otimes}n/2})$
12:   **else**
13:      $\mathcal{P}^{(n)} \leftarrow S_n \cdot (\mathcal{P}^{(1)}\dot{\otimes}\mathcal{P}^{(2)\dot{\otimes}\lfloor n/2\rfloor})$
14:      $\mathcal{D}^{(n)} \leftarrow S_n \cdot (\mathcal{D}^{(1)}\dot{\otimes}\mathcal{D}^{(2)\dot{\otimes}\lfloor n/2\rfloor})$
15:   **end if**
16:   **return** $\mathcal{P}^{(n)}, \mathcal{D}^{(n)}$
17: **end if**
1: **function** CHECK CONSISTENCY $V, \underline{1}, P, D$
2:   $\mathcal{D}^{(1)} \leftarrow \text{cone}\langle D, \underline{1}\rangle$ ▷ partial contractions
3:   $\mathcal{P}^{(1)} \leftarrow \text{cone}\langle \mathcal{D}^{(1)}, P\rangle$
4:   **if**
5:      $\underline{1} \notin \mathcal{P}^{(1)}$ ▷ Axiom 3(a)
6:      **or**
7:      $\langle \mathcal{D}^{(1)}, \mathcal{P}^{(1)}\rangle < 0$ ▷ positivity, unipartite
8:      **or**
9:      $\langle D, P\rangle < 0$ ▷ positivity, bipartite
10:     **or**
11:     $\forall \pi \in S_2, \; \pi(P) \neq P$ ▷ Axiom 6, bipartite
12:     **or**
13:     $\forall \pi \in S_2, \; \pi(D) \neq D$
14:     **or**
15:     $\mathcal{P}^{(1)} \otimes \mathcal{P}^{(1)} \not\subset P$ ▷ Axiom 3(b), unipartite
16:     **or**
17:     $\mathcal{D}^{(1)} \otimes \mathcal{D}^{(1)} \not\subset D$ ▷ Axiom 4(b), unipartite
18:     **or**
19:     $\left[\!\left[_D{}^P{}_D \right]\!\right] \not\subset D$ ▷ Axiom 5(b), bipartite
20:     **or**
21:     $\left[\!\left[^P{}_D{}^P \right]\!\right] \not\subset P$ ▷ Axiom 5(a), bipartite
22:   **then**
23:      **return** "inconsistent"
24:   **else**
25:      **return** "consistent"
26: **end function**

permutations. Axiom 3(a) is checked directly. The n-partite cones are constructed only using the unipartite and bipartite cones. Therefore, Axioms 5(a) and 5(b) follow from Axiom 6 and closure under partial contractions, entanglement swapping and dual entanglement swapping, at the bipartite level. Axiom 4(a) follows from Axioms 5(a) and 5(b) plus the positivity of the unipartite and bipartite cones, since tensor products preserve positivity. Axioms 3(b) and 4(b) follow by construction. ∎

III. ITERATED CHSH GAME

The iterated CHSH game, as previously alluded to, is an extension of the adaptive CHSH game from Ref. [24], to
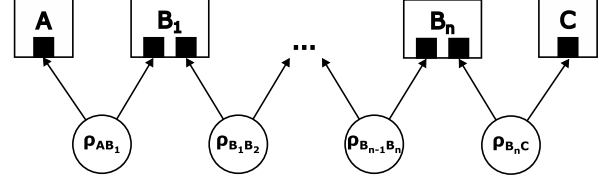

FIG. 3. Diagrammatic representation of the structure of the iterated CHSH game. A and C stand for Alice and Charlie, respectively. The first and the last Bob are shown as B_1 and B_n, respectively. The other Bobs are similarly iterated in order. Each of the nearest neighbors shares a bipartite resource. The local effect space of all the Bobs is bipartite, and that of Alice and Charlie is unipartite.

include multiple rounds of entanglement swapping. The game can be viewed as implementing a CHSH test after the use of a GPT repeater, the generalization of a quantum repeater to GPTs, in order to probe the capacity of the repeater to propagate entanglement.

The iterated CHSH game is parameterized by n, referring to the number of repeater units between the start and end nodes. The players of the game are Alice (A), a collection of n Bobs ($\{B_i\}_{i=1,\cdots,n}$), and Charlie (C). Alice corresponds to the start node, Charlie to the end node, and the Bobs to repeaters. The diagramatic representation of the structure thus produced is shown in Fig. 3.

Each round of the game proceeds as follows. First, each of the Bobs performs a bipartite measurement on the subsystems available to him, effectively leading to n rounds of entanglement swapping. They subsequently broadcast their outcomes. Alice and Charlie are then allowed to implement local corrections based on the outcomes of the Bobs. Finally, Alice and Charlie perform a CHSH test. In addition to shared randomness among all parties, each of the nearest neighbors depicted in Fig. 3 may share a bipartite resource, but no other multipartite resource. (The adaptive CHSH game of Ref. [24], depicted in Fig. 4, is the iterated CHSH game for $n = 1$.)

To calculate the resulting CHSH value between Alice and Charlie in the general case, we simply sum over the CHSH value corresponding to each outcome vector $\vec{b} = (b_1, \cdots, b_n) \in (\mathbb{Z}_k)^n$ (for a k-outcome measurement) produced by the Bobs, weighted by the respective probabilities. That is, if $\beta_{\vec{b}}$ and $p_{\vec{b}}$ are the CHSH value and probability corresponding to the outcome vector $\vec{b}$, respectively, the CHSH
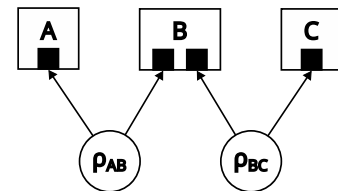

FIG. 4. Diagrammatic representation of the structure of the adaptive CHSH game. A, B, and C stand for Alice, Bob, and Charlie, respectively. Bob shares a bipartite resource ρ_{AB} with Alice and ρ_{BC} with Charlie.

value of the game is calculated as

$$\beta = \sum_{\vec{b} \in (\mathbb{Z}_k)^n} p_{\vec{b}} \beta_{\vec{b}}. \tag{16}$$

In the case of quantum mechanics, it is trivial to describe an optimal strategy for the iterated CHSH game: Choose all the shared states to be the Bell state $|\Phi^+\rangle$. The Bobs perform a Bell basis measurement each. Conditioned on the outcomes, the state shared by Alice and Charlie is then one of four elements of the Bell basis, all of which are equiprobable. A local Pauli operation, performed by either of them, can map it to $|\Phi^+\rangle$. Hence, it follows that quantum theory retains a CHSH value of $2\sqrt{2}$ for all n.

IV. COMPOSITE GPTs

Here we discuss a family of constructions that can each swap PR-box correlations for a fixed number of iterations m. That is, for all $n \leqslant m$ they have a CHSH value of 4 in the iterated CHSH game. But, as we will also show, for all $n > m$ they have a CHSH value of 2.

The main idea of this construction is to use particles with multiple internal degrees of freedom (DOFs). Roughly speaking, we assign to some DOFs the task of carrying the entanglement, and to others the task of supporting entangled measurements. As described below, this way, we can circumvent the tension between state and effect space sizes. We call these particles composites, as they can be viewed as a composite of particles from smaller GPTs.

A. An example

Consider the following example. Say we have particles with two DOFs, labeled 1 and 2. Each DOF supports the same local effects and local states as a unipartite boxworld theory ([9]). That is, if $\mathcal{P}$ and $\mathcal{D}$ are the unipartite boxworld state and effect cones, then the effects measurable on each composite particle are

$$\mathcal{P}_c = \mathcal{P}_1 \dot{\otimes} \mathcal{P}_2. \tag{17}$$

Physically, this says that we do not allow for entangled measurements between the DOFs within each composite particle. We define the joint state space within each particle in the same way:

$$\mathcal{D}_c = \mathcal{D}_1 \dot{\otimes} \mathcal{D}_2. \tag{18}$$

We now define the effects and states realisable between two composite particles. The n-partite theory then results from this input data via the general construction of Sec. II A.

Let

$$\widehat{\mathcal{D}} = \mathcal{D} \hat{\otimes} \mathcal{D} \tag{19}$$

be the cone of bipartite boxworld states. Here, $\hat{\otimes}$ is the maximal tensor product. Mathematically, it is the polar dual to the minimal tensor effects:

$$\mathcal{D} \hat{\otimes} \mathcal{D} := (\mathcal{P} \dot{\otimes} \mathcal{P})'. \tag{20}$$

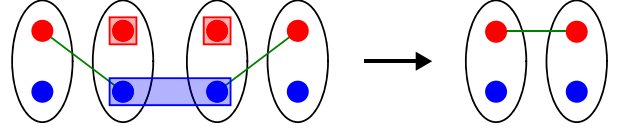Physically it gives rise to the maximally entangled boxworld states that achieve CHSH values of 4 [8,9,21].

FIG. 5. Caricature of PR-box entanglement swapping using composite particles. The red dot signifies DOF 1, the blue dot DOF 2. The green lines signify entangled states. The blue rectangle signifies an entangled measurement, the red squares signify product measurements.

In our theory, in addition to product states, we allow for maximally entangled boxworld states between the first DOFs of any two composite particles, and between the first DOF of one and the second DOF of the other. We thus arrive at the following cone of composite bipartite states

$$D_c := \mathrm{cone}\big(\widehat{\mathcal{D}}_{1_A,1_B} \otimes \mathcal{D}_{2_A} \otimes \mathcal{D}_{2_B} \cup \widehat{\mathcal{D}}_{1_A,2_B} \otimes \widehat{\mathcal{D}}_{2_A,1_B}\big), \tag{21}$$

where we have indicated the tensor factors that any object acts on in the superscripts: Letters A, B refer to the first and second composite particle, respectively; and numbers 1, 2 to the internal degrees of freedom.

Finally, the bipartite effects are just the minimal tensor product of the local ones, except that we allow for maximally entangled measurements between the second DOFs of any two composite particles (this is consistent because no entanglement exists between these particular DOFs). As in the case of states, denote the cone of maximally entangled bipartite effects as

$$\widehat{\mathcal{P}} = \mathcal{P} \hat{\otimes} \mathcal{P}. \tag{22}$$

Then, for the composite bipartite effects we get:

$$P_c := \mathrm{cone}\big(\mathcal{P}_{1_A} \otimes \mathcal{P}_{1_B} \otimes \widehat{\mathcal{P}}_{2_A,2_B}\big). \tag{23}$$

One can now check that P_c, D_c consistently define a theory. This theory allows for a CHSH value of 4 after one round of entanglement swapping, according to the scheme visualized in Fig. 5.

B. General construction

In the example just treated, by adding an additional DOF, we managed to move a PR-box past one round of entanglement swapping. It is reasonable to conjecture that with access to more DOFs one may be able to find a way to sustain entanglement longer, maybe even indefinitely.

To investigate this, we generalize the previous example to m degrees of freedom. Each DOF supports unipartite boxworld states and effects. Now, for each DOF k we have to specify the two disjoint subsets of DOFs that form entangled states and entangled effects with k. This amounts to specifying two symmetric bipartite graphs as follows:

Start with a collection of $2m$ vertices, two per DOF

$$V = \{v_1, \cdots, v_m, w_1, \cdots, w_m\}$$
$$= \{v_i\}_{i=1,\cdots,m} \cup \{w_i\}_{i=1,\cdots,m}. \tag{24}$$

On this vertex set, define two symmetric bipartite graphs, $G = (V, E)$ and $H = (V, F)$ (with the same bipartition over the above indicated subsets), such that they lie in each other's

complement. The elements of edge sets E and F therefore, are ordered pairs of vertices, one from each of the subsets $\{v_i\}_{i=1,\cdots,m}$ and $\{w_i\}_{i=1,\cdots,m}$. Indeed, the edges of graph G specifies those pairs of DOFs that can support entangled states, whereas the edges of graph H give those pairs that can support entangled measurements.

The entire construction can be summarized by the following rules imposed on the edge sets E and F:

(i) If $(v_i, w_j) \in E$ then $(v_j, w_i) \in E$, by symmetry. Same for F.

(ii) If $(v_i, w_j) \in E$ then $(v_i, w_j) \notin F$, because they lie in each other's complements.

(iii) If $(v_i, w_j), (v_k, w_l) \in E$ and $(v_j, w_k) \in F$ then, we can concatenate edges via entanglement swapping as follows:

$$(v_i, w_j) \circ (v_j, w_k) \circ (v_k, w_l) \equiv (v_i, w_l) \in E.$$

(iv) If $(v_i, w_j), (v_k, w_l) \in F$ and $(v_j, w_k) \in E$ then, we can concatenate edges via dual entanglement swapping as follows:

$$(v_i, w_j) \circ (v_j, w_k) \circ (v_k, w_l) \equiv (v_i, w_l) \in F.$$

Lemma 2. Given a composite particle with m boxworld DOFs, there exists $l \in \mathbb{N}$ such that $\forall\, n \geqslant l$ the CHSH value of the composite GPT in the iterated CHSH game parametrized by n, is 2.

Proof. In order to sustain entanglement indefinitely under entanglement swapping with a finite number of DOFs, we must have a closed cycle. That is, there is some chain of concatenations, alternating edges from E and F (starting and ending with E) that reproduces the first edge. In equations:

$$(v_1, w_2) \circ (v_2, w_3) \circ \cdots$$
$$\cdots \circ (v_{k-1}, w_k) \circ (v_k, w_2) \equiv (v_1, w_2) \in E, \quad (25)$$

with (v_i, w_{i+1}) belongs to E for odd i and F for even i. Since the second and penultimate edge belong to F, we can use Rule 4 to simplify the above chain to

$$(v_1, w_2) \circ (v_2, w_k) \circ (v_k, w_2) \equiv (v_1, w_2). \quad (26)$$

The above equation implies that $(v_1, w_2), (v_k, w_2) \in E$, and $(v_2, w_k) \in F$. But, by Rule 1 we get $(v_k, w_2) \in F$, which contradicts Rule 2. This then means, each time we do an additional round of entanglement swapping, we have to add a new degree of freedom (if not more). This implies, given access to m DOFs, the maximum number of entanglement swapping rounds such that the output is still entangled is less than or equal to $m - 1$. Therefore, for any $n \geqslant m$ we end up with only product states in the iterated CHSH game, meaning the CHSH value is no more than 2. ∎

V. OBLATE STABILIZER THEORY

In this section we present our main result. That is, we construct the oblate stabilizer theory, which not only achieves a CHSH value of 4 but is also stable under entanglement swapping. In other words, it can sustain this CHSH value indefinitely under entanglement swapping. We also show that, given the resources described by this theory, there is an optimal strategy by which we get a CHSH value of 4 in the iterated CHSH game.

A. Setup

The theory can be obtained by slightly deforming the set of quantum-mechanical stabilizer states. For this reason, we will use objects from the mathematical description of quantum mechanics to construct it.

Consider the one-qubit stabilizer polytope, i.e., the convex hull of the following states on the Bloch sphere

$$|x_{\pm}\rangle\langle x_{\pm}| := \tfrac{1}{2}(\mathbb{1} \pm \sigma_1),$$
$$|y_{\pm}\rangle\langle y_{\pm}| := \tfrac{1}{2}(\mathbb{1} \pm \sigma_2), \quad (27)$$
$$|z_{\pm}\rangle\langle z_{\pm}| := \tfrac{1}{2}(\mathbb{1} \pm \sigma_3),$$

where $\sigma_1, \sigma_2, \sigma_3$ are the Pauli matrices. Now perform a uniform stretch in the equatorial plane of the Bloch sphere, i.e., for some $r > 1$, set

$$|\tilde{x}_{\pm}\rangle\langle \tilde{x}_{\pm}| := \tfrac{1}{2}(\mathbb{1} \pm r\sigma_1),$$
$$|\tilde{y}_{\pm}\rangle\langle \tilde{y}_{\pm}| := \tfrac{1}{2}(\mathbb{1} \pm r\sigma_2), \quad (28)$$
$$|\tilde{z}_{\pm}\rangle\langle \tilde{z}_{\pm}| := |z_{\pm}\rangle\langle z_{\pm}| = \tfrac{1}{2}(\mathbb{1} \pm \sigma_3).$$

These will be the building blocks for the unipartite state space. The Bell state is defined as usual as:

$$|\Phi^+\rangle\langle\Phi^+| := \tfrac{1}{4}(\sigma_0 \otimes \sigma_0 + \sigma_1 \otimes \sigma_1 - \sigma_2 \otimes \sigma_2 + \sigma_3 \otimes \sigma_3).$$
$$(29)$$

It satisfies the standard identities

$$(A \otimes B)|\Phi^+\rangle = (\mathbb{1} \otimes BA^t)|\Phi^+\rangle \quad (30)$$

and

$$\mathrm{tr}_1(|\Phi^+\rangle\langle\Phi^+|) = \mathrm{tr}_2(|\Phi^+\rangle\langle\Phi^+|) = \tfrac{1}{2}\mathbb{1}. \quad (31)$$

Where A^t stands for the matrix transpose of A, and

$$|\Phi^+\rangle = \tfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \quad (32)$$

From the Bell state, we can generate the Bell basis as:

$$|\Phi_\mu^+\rangle\langle\Phi_\mu^+| := (\sigma_\mu \otimes \mathbb{1})|\Phi^+\rangle\langle\Phi^+|(\sigma_\mu \otimes \mathbb{1}), \quad \mu \in \mathbb{Z}_4.$$
$$(33)$$

In keeping with the quantum formalism, effects will also be represented by 2×2 matrices and the pairing between states and effects by the trace inner product. Additionally, partial contractions between states and effects are realized as partial traces. For example, for a unipartite effect e_1 and bipartite state ρ_{12}, the partial contraction is

$$\langle \rho_{12}, e_1 \rangle := \mathrm{tr}_1[\rho_{12}(e_1 \otimes \mathbb{1})]. \quad (34)$$

Similarly, for a bipartite effect e and bipartite states ρ and σ, the entanglement swapping map is

$$[\![_\rho{}^e{}_\sigma]\!] := \mathrm{tr}_{23}[\rho_{12} \otimes \mathbb{1}^{\otimes 2}(\mathbb{1} \otimes e_{23} \otimes \mathbb{1})\mathbb{1}^{\otimes 2} \otimes \sigma_{34}]. \quad (35)$$

B. Unipartite theory

Let $R = e^{-i\frac{\pi}{8}\sigma_3}$ be the unitary, which, by conjugation $[R(\cdot)R^\dagger]$, implements a $\pi/4$-rotation of the Bloch sphere about the z axis. Let Ω be the stretched stabilizer states, for some choice of $r > 1$:

$$\Omega := \{|\tilde{x}_{\pm}\rangle\langle \tilde{x}_{\pm}|, |\tilde{y}_{\pm}\rangle\langle \tilde{y}_{\pm}|, |\tilde{z}_{\pm}\rangle\langle \tilde{z}_{\pm}|\}. \quad (36)$$
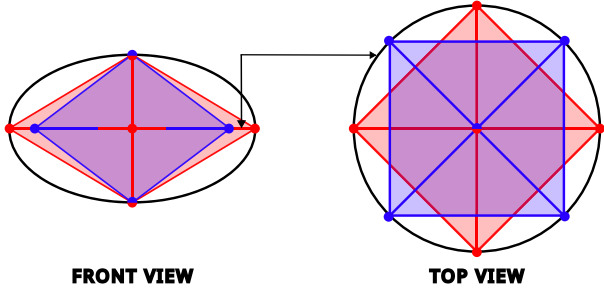
FIG. 6. Caricature of the geometry of the unipartite oblate stabilizer theory. The black outlines represent a Bloch sphere with a scaled up equatorial plane. The rays formed by vertices of the red polytope (what used to be the stabilizer states) represent the extremal rays of the state cone. The rays formed by vertices of the rotated (about the z axis, by $\pi/4$) blue polytope represent the extremal rays of the effect cone.

Definition 2 (Unipartite oblate stabilizer theory). Choose $r = (\cos(\frac{\pi}{4}))^{-\frac{1}{2}} = \sqrt[4]{2}$, and define the following:

(i) The unit effect, $\underline{1} := \mathbb{1}$, the 2×2 identity matrix;

(ii) The effect cone, $\mathcal{P}^{(1)} := \text{cone}(R\Omega R^\dagger)$;

(iii) The state cone, $\mathcal{D}^{(1)} := \text{cone}(\Omega)$.

Indeed, we get the following derived objects: the negation of an effect is

$$\neg e = \mathbb{1} - e. \tag{37}$$

The state space is simply

$$\mathcal{S}^{(1)} = \text{conv}(\Omega). \tag{38}$$

Similarly, the effect space is

$$\mathcal{E}^{(1)} = \text{conv}(\{0, \mathbb{1}\} \cup R\Omega R^\dagger). \tag{39}$$

It is easily verified that we have a well-defined unipartite theory as: (i) The effect space is closed under negations

$$\neg \mathcal{E}^{(1)} = \mathcal{E}^{(1)}. \tag{40}$$

(ii) The states are normalized

$$\rho \in \mathcal{S}^{(1)} \quad \Rightarrow \quad \text{tr}(\rho \mathbb{1}) = 1. \tag{41}$$

(iii) Pairing a state and an effect gives a probability, i.e.,

$$\rho \in \mathcal{S}^{(1)}, \; e \in \mathcal{E}^{(1)} \quad \Rightarrow \quad \text{tr}(\rho e) \in [0, 1]. \tag{42}$$

It suffices to check (42) for the extremal vertices. Additionally, due to (40) only the upper bound needs to be checked. From Fig. 6, the inner product on the z axis is unchanged from quantum theory and hence bounded between 0 and 1. On the equatorial plane the largest inner product is between two nearest neighbors. By construction, the angle enclosed by the corresponding Bloch vectors is $\pi/4$, meaning

$$\sup_{\rho, e} \text{tr}(\rho e) = \frac{1}{2}[1 + r^2 \cos(\pi/4)] = 1. \tag{43}$$

C. Bipartite theory

To construct the bipartite theory we first interpret the bipartite states as maps from the unipartite effect cone to the unipartite state cone, via partial contraction. Similarly for bipartite effects [refer to Figs. 2(a) and 2(b)]. Then, we use the following properties of the unipartite theory:

(i) Both the state and effect cones are invariant under conjugation by Pauli matrices. In the Bloch picture, these correspond to reflections about the x, y, and z axes.

(ii) Both state and effect cones are invariant under matrix transpose. In the Bloch picture, this corresponds to reflections about the xz plane.

(iii) Both the state and effect cones are invariant under conjugation by R^m for $m \in \{0, 2, 4, 6\}$. In the Bloch picture this corresponds to a $m\pi/4$ rotation about the z axis.

(iv) On the other hand, conjugation by R^m for $m \in \{1, 3, 5, 7\}$ maps the state cone to the effect cone and vice versa.

Therefore, for $\mu \in \mathbb{Z}_4$, $m \in \mathbb{Z}_8$ define the projections

$$|\Phi^+_{\mu,m}\rangle\langle\Phi^+_{\mu,m}| := (\sigma_\mu R^m)^\dagger \otimes \mathbb{1}|\Phi^+\rangle\langle\Phi^+|(\sigma_\mu R^m) \otimes \mathbb{1}. \tag{44}$$

Further, define the set

$$\Phi := \{|\Phi^+_{\mu,m}\rangle\langle\Phi^+_{\mu,m}| \mid \mu \in \mathbb{Z}_4, \; m \in \mathbb{Z}_8, \; m \text{ odd}\}. \tag{45}$$

Definition 3 (Oblate stabilizer theory). In the sense of Sec. II A, oblate stabilizer theory is the theory specified by the following data:

(i) V, the set of 2×2 Hermitian matrices,

(ii) $\underline{1} = \mathbb{1}$, the 2×2 identity matrix,

(iii) $P = \text{cone}(R\Omega R^\dagger \otimes R\Omega R^\dagger \cup \Phi)$,

(iv) $D = \text{cone}(\Omega \otimes \Omega \cup \Phi)$.

Lemma 3. The data specified in Definition 3 can be consistently extended to a theory using Algorithm 1.

Proof. We verify that Algorithm 1 accepts the data, which is thus consistent by Lemma 1.

The partial trace of all entangled states introduced is $\frac{1}{2}\mathbb{1}$. Partial contractions of entangled effects with unipartite states can be viewed as conjugation by a odd rotation, and then by a Pauli matrix. This, as discussed before, maps the unipartite state cone to the unipartite effect cone. Therefore, by construction, the $\mathcal{D}^{(1)}$ and $\mathcal{P}^{(1)}$ assigned by the algorithm are the same $\mathcal{D}^{(1)}$ and $\mathcal{P}^{(1)}$ as in Definition 2. Hence, we already have $\underline{1} \in \mathcal{P}^{(1)}$ and $\langle\mathcal{D}^{(1)}, \mathcal{P}^{(1)}\rangle \geqslant 0$. Since the entangled states and effects are projectors taken from quantum theory without modification, the trace inner product between them is nonnegative.

We have verified above that pairing unipartite states and effects leads to positive outcomes, and this property is preserved under tensor products. Using identities (30) and (31), it can be easily shown that

$$\text{tr}(|\Phi^+_{\mu,m}\rangle\langle\Phi^+_{\mu,m}|e \otimes f) = \frac{1}{2}\text{tr}[f\sigma_\mu R^m e^t (R^m)^\dagger \sigma_\mu], \tag{46}$$

which is just $\frac{1}{2}$ times the pairing between a unipartite state and effect, and hence, is positive. The same argument extends to entangled effects and product states. Therefore, we have $\langle D, P\rangle \geqslant 0$.

For invariance under permutation of systems, we need only check S_2 invariance for the set of entangled states and effects, since everything else is permutation invariant by construction. This is readily verified by using identity (30) since $|\Phi^+\rangle\langle\Phi^+|$ is already S_2 invariant. The tensor products of the unipartite cones, $\mathcal{P}^{(1)} \otimes \mathcal{P}^{(1)}$ and $\mathcal{D}^{(1)} \otimes \mathcal{D}^{(1)}$ are subsets of, respectively, P and D by construction.

It remains to be shown that the theory is closed under entanglement swapping and dual entanglement swapping. We only treat the first case explicitly. The dual version follows in complete analogy. We separate the entanglement swapping contractions into four types.

(i) The effect factorizes, i.e., $[\![_{\bullet}{}^{e \otimes f}{}_{\bullet}]\!]$. In this case the result is an element of $\mathcal{D}^{(1)} \otimes \mathcal{D}^{(1)}$, because the contraction splits as follows:
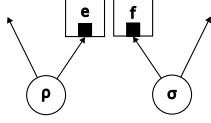
(ii) Both states are product states, i.e., $[\![_{\rho_1 \otimes \rho_2}{}^{\bullet}{}_{\rho_3 \otimes \rho_4}]\!]$. In this case the result is again an element of $\mathcal{D}^{(1)} \otimes \mathcal{D}^{(1)}$, because the central objects can be grouped as:
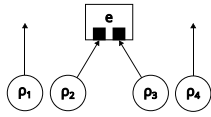
(iii) There is one entangled state and one entangled effect, i.e., $[\![_{\Phi}{}^{\Phi}{}_{\rho \otimes \sigma}]\!]$ for example. In this case the result is once again an element of $\mathcal{D}^{(1)} \otimes \mathcal{D}^{(1)}$, because we can split up the contraction as:
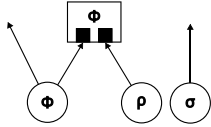
(iv) All three objects are entangled. This leads to a simple but lengthy calculation, which we have deferred to Appendix A 2. The result is

$$[\![_{\Phi}{}^{\Phi}{}_{\Phi}]\!] \subset \tfrac{1}{4}\Phi. \tag{47}$$

Therefore, none of the conditions required to detect inconsistency in Algorithm 1 are met and hence, it does not return inconsistent. ∎

D. Iterated CHSH game

The theory exhibits a CHSH value of 4 for the observable (10) and the choices

$$\begin{aligned}
\rho &= |\Phi_{0,1}^+\rangle\langle\Phi_{0,1}^+|, \\
A_0 &= R|\tilde{x}_+\rangle\langle\tilde{x}_+|R^\dagger - R|\tilde{x}_-\rangle\langle\tilde{x}_-|R^\dagger = rR\sigma_1 R^\dagger, \\
A_1 &= R|\tilde{y}_+\rangle\langle\tilde{y}_+|R^\dagger - R|\tilde{y}_-\rangle\langle\tilde{y}_-|R^\dagger = rR\sigma_2 R^\dagger, \\
B_0 &= R|\tilde{x}_+\rangle\langle\tilde{x}_+|R^\dagger - R|\tilde{x}_-\rangle\langle\tilde{x}_-|R^\dagger = rR\sigma_1 R^\dagger, \\
B_1 &= R|\tilde{y}_+\rangle\langle\tilde{y}_+|R^\dagger - R|\tilde{y}_-\rangle\langle\tilde{y}_-|R^\dagger = rR\sigma_2 R^\dagger.
\end{aligned} \tag{48}$$

Since this is the maximum possible value, the CHSH value associated with this theory, as defined in Eq. (12), is also 4.

Also, $\forall \rho \in \mathcal{D}^{(2)}$ we have (refer to Appendix A 1)

$$[\![_{|\Phi_{0,1}^+\rangle\langle\Phi_{0,1}^+|}{}^{|\Phi_{0,1}^+\rangle\langle\Phi_{0,1}^+|}{}_{\rho}]\!] \propto \rho \tag{49}$$

Therefore, we not only have closure but also stability under both entanglement swapping and dual entanglement swapping.

Remark 1. It is worth pointing out here that the choice of correlators being only on the xy plane is no coincidence. For any situation with a z measurement, oblate stabilizer theory no longer has such a strong CHSH violation.

The general strategy for the iterated CHSH game for oblate stabilizer theory is the same as the optimal strategy for quantum theory discussed in Sec. III. Alice and Charlie have access to two-setting two-outcome measurement machines. The Bobs have access to four-outcome bipartite measurement machines. In each run of the experiment, nearest neighbors share a bipartite oblate stabilizer state. Each Bob performs a bipartite measurement on the subsystems available to him and broadcasts his outcome. Based on this, either Alice or Charlie apply a correction locally and perform a CHSH test.

Theorem 1. Oblate stabilizer theory reaches a value of $\beta = 4$ in the iterated CHSH game.

Proof. Build the two-setting two-outcome measurement machines of Alice and Charlie out of the correlators

$$\begin{aligned}
A_0 &= rR\sigma_1 R^\dagger, \quad A_1 = rR\sigma_2 R^\dagger, \\
C_0 &= rR\sigma_1 R^\dagger, \quad C_1 = rR\sigma_2 R^\dagger.
\end{aligned} \tag{50}$$

As we have noted in (48), the above are valid correlators of the theory. For the four-outcome measurements, choose

$$\mathcal{M}_{B_i} = \{|\Phi_{\mu,1}^+\rangle\langle\Phi_{\mu,1}^+| \mid \mu \in \mathbb{Z}_4\}, \; \forall i. \tag{51}$$

From (33) it follows that $\mathcal{M}_{B_i}$ is a measurement. For the shared bipartite states, choose

$$\rho_{AB_1} = \rho_{B_1 B_2} = \cdots = \rho_{B_n C} = |\Phi_{0,1}^+\rangle\langle\Phi_{0,1}^+|. \tag{52}$$

Given these choices one can verify (refer to Appendix A 2) that the four output states obtained after each consecutive entanglement swap are the same. Each state is also equiprobable. The output state after n rounds depends solely on the multiplicity of each of the four outcomes of $\mathcal{M}_{B_i}$, in the vector $\vec{b} = (b_1, \cdots, b_n) \in (\mathbb{Z}_4)^n$. Therefore to obtain the proper correction one can convert the outcomes to binary and perform a bitwise XOR (equivalent to finding the resultant element of the Klein four-group). The output state and correction corresponding to each $\mu \in \mathbb{Z}_4$ is tabulated in Table I.

The correction maps each output state back to $|\Phi_{0,1}^+\rangle\langle\Phi_{0,1}^+|$, which gives a CHSH value of 4 with the above choice of correlators as was stated earlier. And therefore, we have

$$\beta = \sum_{\vec{b} \in (\mathbb{Z}_4)^n} p_b \beta_b = \sum_{\vec{b} \in (\mathbb{Z}_4)^n} \left(\tfrac{1}{4}\right)^n 4 = 4. \tag{53}$$

∎

TABLE I. The output state and correction corresponding to each outcome of Bob's measurement.

$\mu \in \mathbb{Z}_4$	State	Correction
0	$|\Phi_{0,1}^+\rangle\langle\Phi_{0,1}^+|$	$\sigma_0 \otimes \mathbb{1}$
1	$|\Phi_{1,-1}^+\rangle\langle\Phi_{1,-1}^+|$	$\sigma_1 \otimes \mathbb{1}$
2	$|\Phi_{2,-1}^+\rangle\langle\Phi_{2,-1}^+|$	$\sigma_2 \otimes \mathbb{1}$
3	$|\Phi_{3,1}^+\rangle\langle\Phi_{3,1}^+|$	$\sigma_3 \otimes \mathbb{1}$

VI. CONCLUSION AND OUTLOOK

We have constructed a GPT in which a CHSH violation of 4 can be sustained indefinitely under entanglement swapping. As a consequence, the iterated CHSH game is insufficient to single out QM among GPTs. In the process of obtaining this result, we have also set up a framework to turn bipartite theories into multipartite theories in which entanglement swapping is consistently defined.

As an outlook to future work, Ref. [20] suggests that theories should satisfy stronger symmetry conditions, i.e., "for any state and set of local measurements, if the local outcome probabilities are permuted, then there is a state that achieves these permuted correlations under the same measurements." If this is interpreted as an invariance under permutation of subsystems, then oblate stabilizer theory satisfies this requirement. If instead we interpret this as a symmetry under permutation of local extremal effects, then our construction fails to satisfy this requirement. In particular, our construction breaks the

symmetry between z observables and those on the equatorial plane. This raises two complementary questions for further work: (i) Are there natural, stronger conditions on multipartite correlations for which QM is indeed optimal? (ii) Can one find a theory that beats QM in the iterated CHSH game and that is isotropic in the sense of having a transitive symmetry group action on all extremal effects?

APPENDIX

The Bell state can be written concisely as

$$|\Phi^+\rangle\langle\Phi^+| = \frac{1}{4}\sum_{\mu=0}^{3}(-1)^{\delta_{2\mu}}\sigma_\mu \otimes \sigma_\mu. \tag{A1}$$

1. Stability of OST

To verify that oblate stabilizer theory is stable under entanglement swapping we show that

$$\left[\!\left[_{|\Phi_{0,1}^+\rangle\langle\Phi_{0,1}^+|}{}^{|\Phi_{0,1}^+\rangle\langle\Phi_{0,1}^+|}\bullet\right]\!\right] : \sigma_\mu \otimes \sigma_\nu \longmapsto \tfrac{1}{4}\sigma_\mu \otimes \sigma_\nu. \tag{A2}$$

To this end, note that

$$\left[\!\left[_{|\Phi_{0,1}^+\rangle\langle\Phi_{0,1}^+|}{}^{|\Phi_{0,1}^+\rangle\langle\Phi_{0,1}^+|}{}_{\sigma_\mu\otimes\sigma_\nu}\right]\!\right]$$

$$= \mathrm{tr}_{23}((R^\dagger\otimes\mathbb{1})_{12}|\Phi^+\rangle\langle\Phi^+|_{12}(R\otimes\mathbb{1})_{12}(R^\dagger\otimes\mathbb{1})_{23}|\Phi^+\rangle\langle\Phi^+|_{23}(R\otimes\mathbb{1})_{23}(\sigma_\mu\otimes\sigma_\nu)_{34})$$

$$= \mathrm{tr}_{23}((\mathbb{1}\otimes R^{-1})_{12}|\Phi^+\rangle\langle\Phi^+|_{12}(R^{-1}R\otimes\mathbb{1}^{\otimes2})_{123}|\Phi^+\rangle\langle\Phi^+|_{23}(R\otimes\mathbb{1})_{23}(\sigma_\mu\otimes\sigma_\nu)_{34})$$

$$= \mathrm{tr}_3\left(\frac{1}{16}\sum_{\alpha,\beta=0}^{3}(-1)^{\delta_{2\alpha}+\delta_{2\beta}}\mathrm{tr}(R^{-1}\sigma_\alpha\sigma_\beta R)(\sigma_\alpha\otimes\sigma_\beta)_{13}(\sigma_\mu\otimes\sigma_\nu)_{34}\right)$$

$$= \frac{1}{8}\sum_{\alpha=0}^{3}\mathrm{tr}(\sigma_\alpha\sigma_\mu)(\sigma_\alpha\otimes\sigma_\nu) = \frac{1}{4}\sigma_\mu\otimes\sigma_\nu. \tag{A3}$$

Every $\rho \in \mathcal{D}^{(2)}$ is a linear combination of $\sigma_\mu \otimes \sigma_\nu$ for $\mu, \nu \in \mathbb{Z}_4$, hence the claim follows.

2. Iterated entanglement swapping

In order to find the result of $\left[\!\left[_\Phi{}^\Phi{}_\Phi\right]\!\right]$, let us first calculate the following identity: for some operators A, B, C, D

$$\mathrm{tr}_{23}[(A\otimes\mathbb{1})_{12}|\Phi^+\rangle\langle\Phi^+|_{12}(B\otimes\mathbb{1}^{\otimes2})_{123}|\Phi^+\rangle\langle\Phi^+|_{23}(\mathbb{1}^{\otimes2}\otimes C)_{234}|\Phi^+\rangle\langle\Phi^+|_{34}(\mathbb{1}\otimes D)_{34}]$$

$$= \frac{1}{64}\sum_{\mu,\nu,\lambda=0}^{3}((-1)^{\delta_{2\mu}+\delta_{2\nu}+\delta_{2\lambda}}\mathrm{tr}(\sigma_\mu\sigma_\nu)\mathrm{tr}(\sigma_\nu\sigma_\lambda)A\sigma_\mu B\otimes C\sigma_\lambda D)$$

$$= \frac{1}{16}(A\otimes C)\left(\sum_{\mu=0}^{3}(-1)^{\delta_{2\mu}}\sigma_\mu\otimes\sigma_\mu\right)(B\otimes D)$$

$$= \frac{1}{4}(A\otimes C)|\Phi^+\rangle\langle\Phi^+|(B\otimes D) = \frac{1}{4}(AC^t\otimes\mathbb{1})|\Phi^+\rangle\langle\Phi^+|(D^tB\otimes\mathbb{1}). \tag{A4}$$

Using this we can now compute

$$\text{tr}_{23}((A^\dagger \otimes \mathbb{1})_{12}|\Phi^+\rangle\langle\Phi^+|_{12}(A \otimes \mathbb{1})_{12}(B^\dagger \otimes \mathbb{1})_{23}|\Phi^+\rangle\langle\Phi^+|_{23}(B \otimes \mathbb{1})_{23}(C^\dagger \otimes \mathbb{1})_{34}|\Phi^+\rangle\langle\Phi^+|_{34}(C \otimes \mathbb{1})_{34})$$

$$= \text{tr}_{23}((A^\dagger \otimes \mathbb{1})_{12}|\Phi^+\rangle\langle\Phi^+|_{12}(\overline{B}A \otimes \mathbb{1}^{\otimes 2})_{123}|\Phi^+\rangle\langle\Phi^+|_{23}(\mathbb{1}^{\otimes 2} \otimes \overline{C}B)_{234}|\Phi^+\rangle\langle\Phi^+|_{34}(\mathbb{1} \otimes C^t)_{34})$$

$$= \tfrac{1}{4}(A^\dagger(\overline{C}B)^t \otimes \mathbb{1})|\Phi^+\rangle\langle\Phi^+|((C^t)^t\overline{B}A \otimes \mathbb{1})$$

$$= \tfrac{1}{4}(C\overline{B}A)^\dagger \otimes \mathbb{1}|\Phi^+\rangle\langle\Phi^+|(C\overline{B}A) \otimes \mathbb{1}. \tag{A5}$$

Therefore, the entanglement swapping map

$$\left[\!\!\left[{}_{|\Phi_{\mu,m}^+\rangle\langle\Phi_{\mu,m}^+|}^{|\Phi_{\nu,m'}^+\rangle\langle\Phi_{\nu,m'}^+|}_{|\Phi_{\lambda,m''}^+\rangle\langle\Phi_{\lambda,m''}^+|} \right]\!\!\right] = \text{tr}_{23}([(\sigma_\mu R^m)^\dagger \otimes \mathbb{1}]_{12}|\Phi^+\rangle\langle\Phi^+|_{12}[(\sigma_\mu R^m) \otimes \mathbb{1}]_{12}[(\sigma_\nu R^{m'})^\dagger \otimes \mathbb{1}]_{23}|\Phi^+\rangle\langle\Phi^+|_{23}[(\sigma_\nu R^{m'})$$

$$\otimes \mathbb{1}]_{23}[(\sigma_\lambda R^{m''})^\dagger \otimes \mathbb{1}]_{34}|\Phi^+\rangle\langle\Phi^+|_{34}[(\sigma_\lambda R^{m''}) \otimes \mathbb{1}]_{34}) \tag{A6}$$

reduces to (A5) with

$$A = \sigma_\mu R^m, \quad B = \sigma_\nu R^{m'}, \quad C = \sigma_\lambda R^{m''}.$$

which gives

$$C\overline{B}A = \sigma_\lambda R^{m''}\overline{\sigma_\nu R^{m'}}\sigma_\mu R^m = (\pm i)\sigma_\xi R^{m^*}, \tag{A7}$$

with $\xi \in \mathbb{Z}_4$ and $m^* \in \mathbb{Z}_8$ and odd. Therefore, it follows that

$$\left[\!\!\left[{}_\Phi{}^\Phi{}_\Phi \right]\!\!\right] = \tfrac{1}{4}\Phi. \tag{A8}$$

Finally, we can specialize to the case of the optimal strategy for the iterated CHSH game. For $n = 1$:

$$A = \sigma_0 R, \quad B = \sigma_\mu R, \quad C = \sigma_0 R \Rightarrow C\overline{B}A = R\sigma_\mu$$

We can calculate the output of $n = 2$ by entering the output of the first round into the second round, i.e.,

$$A = R\sigma_\mu, \quad B = \sigma_\nu R, \quad C = \sigma_0 R \Rightarrow C\overline{B}A = \sigma_0 R\overline{\sigma_\nu R}R\sigma_\mu = R\sigma_\nu\sigma_\mu.$$

These are the same four states as $n = 1$ up to scaling and factors of $\pm i$, which are eliminated since they come in complex conjugate pairs. This means that the set of output states is closed. Moreover, each time we have conjugation by an additional Pauli matrix. The final result therefore depends only on the number of time each Pauli matrix occurs. That is, it depends on the multiplicity of each member of $\mathbb{Z}_4$ in the outcome vector $\vec{b} = (b_1, \cdots, b_n)$.

[1] I. E. Segal, Postulates for general quantum mechanics, Ann. Math. **48**, 930 (1947).

[2] G. Ludwig, Attempt of an axiomatic foundation of quantum mechanics and more general theories, II, Commun. Math. Phys. **4**, 331 (1967).

[3] G. Ludwig, Attempt of an axiomatic foundation of quantum mechanics and more general theories, III, Commun. Math. Phys. **9**, 1 (1968).

[4] G. Dähn, Attempt of an axiomatic foundation of quantum mechanics and more general theories. IV, Commun. Math. Phys. **9**, 192 (1968).

[5] P. Stolz, Attempt of an axiomatic foundation of quantum mechanics and more general theories V, Commun. Math. Phys. **11**, 303 (1969).

[6] P. Stolz, Attempt of an axiomatic foundation of quantum mechanics and more general theories VI, Commun. Math. Phys. **23**, 117 (1971).

[7] E. B. Davies and J. T. Lewis, An operational approach to quantum probability, Commun. Math. Phys. **17**, 239 (1970).

[8] J. Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007).

[9] M. Plávala, General probabilistic theories: An introduction, Phys. Rep. **1033**, 1 (2023).

[10] L. Hardy, Quantum theory from five reasonable axioms, arXiv:quant-ph/0101012.

[11] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Information causality as a physical principle, Nature **461**, 1101 (2009).

[12] N. Miklin and M. Pawłowski, Information causality without concatenation, Phys. Rev. Lett. **126**, 220403 (2021).

[13] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, Limit on nonlocality in any world in which communication complexity is not trivial, Phys. Rev. Lett. **96**, 250401 (2006).

[14] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum nonlocality and beyond: limits from nonlocal computation, Phys. Rev. Lett. **99**, 180502 (2007).

[15] M. Navascués and H. Wunderlich, A glance beyond the quantum model, Proc. R. Soc. London A **466**, 881 (2010).

[16] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín, Local orthogonality as a multipartite principle for quantum correlations, Nat. Commun. **4**, 2263 (2013).

[17] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed experiment to test local hidden-variable theories, Phys. Rev. Lett. **23**, 880 (1969).

[18] J. S. Bell, On the Einstein Podolsky Rosen paradox, Phys. Phys. Fiz. **1**, 195 (1964).

[19] B. S. Cirel'son, Quantum generalizations of Bell's inequality, Lett. Math. Phys. **4**, 93 (1980).

[20] M. Weilenmann and R. Colbeck, Toward correlation self-testing of quantum theory in the adaptive Clauser-Horne-Shimony-Holt game, Phys. Rev. A **102**, 022203 (2020).

[21] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, Found. Phys. **24**, 379 (1994).

[22] A. J. Short, S. Popescu, and N. Gisin, Entanglement swapping for generalized nonlocal correlations, Phys. Rev. A **73**, 012101 (2006).

[23] P. Skrzypczyk, N. Brunner, and S. Popescu, Emergence of quantum correlations from nonlocality swapping, Phys. Rev. Lett. **102**, 110402 (2009).

[24] M. Weilenmann and R. Colbeck, Self-testing of physical theories, or, is quantum theory optimal with respect to some information-processing task? Phys. Rev. Lett. **125**, 060406 (2020).

[25] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner, Limits on nonlocal correlations from the structure of the local state space, New J. Phys. **13**, 063024 (2011).