

## MERMIN POLYTOPES IN QUANTUM COMPUTATION AND FOUNDATIONS

CIHAN OKAY<sup>a</sup> HO YIU CHUNG<sup>b</sup> SELMAN IPEK<sup>c</sup>  
*Department of Mathematics, Bilkent University, Ankara, Turkey*

Received November 19, 2022

Revised June 27, 2023

Mermin square scenario provides a simple proof for state-independent contextuality. In this paper, we study polytopes  $MP_\beta$  obtained from the Mermin scenario, parametrized by a function  $\beta$  on the set of contexts. Up to combinatorial isomorphism, there are two types of polytopes  $MP_0$  and  $MP_1$  depending on the parity of  $\beta$ . Our main result is the classification of the vertices of these two polytopes. In addition, we describe the graph associated with the polytopes. All the vertices of  $MP_0$  turn out to be deterministic. This result provides a new topological proof of a celebrated result of Fine characterizing noncontextual distributions on the CHSH scenario.  $MP_1$  can be seen as a nonlocal toy version of  $\Lambda$ -polytopes, a class of polytopes introduced for the simulation of universal quantum computation. In the 2-qubit case, we provide a decomposition of the  $\Lambda$ -polytope using  $MP_1$ , whose vertices are classified, and the nonsignaling polytope of the (2, 3, 2) Bell scenario, whose vertices are well-known.

*Keywords:* quantum computation, quantum contextuality, algebraic topology, polytope theory

### 1 Introduction

Central to many of the paradoxes arising in quantum theory is that the act of measurement cannot be understood as merely revealing the pre-existing values of some hidden variables.<sup>d</sup> Instead, as shown by the ‘no-go’ theorems of Bell [4], and Kochen-Specker (KS) [5], the outcomes of quantum measurements depend crucially on what else they are being measured with, a phenomenon known as contextuality. (For a recent review, see e.g., [6].) A particularly accessible illustration of this quantum mechanical feature using just two spin-1/2 particles was given some years ago by Mermin [7], an example which is now commonly called Mermin’s square. This scenario, as illustrated in Fig. (2a), consists of 9 measurements  $M$  and 6 contexts  $\mathcal{C}$  given by the rows and the columns of the square grid. Together with the function  $\beta$  which assigns a value in  $\mathbb{Z}_2 = \{0, 1\}$  to each context this scenario specifies a binary linear system [8]. It is known that this binary system  $(M, \mathcal{C}, \beta)$  has a classical solution if and only if

$$[\beta] = \sum_{C \in \mathcal{C}} \beta(C) = 0 \pmod{2}. \quad (1)$$

---

<sup>a</sup>cihan.okay@bilkent.edu.tr

<sup>b</sup>hoyiu.chung@bilkent.edu.tr

<sup>c</sup>selman.ipek@bilkent.edu.tr

<sup>d</sup>A classic counterexample to this viewpoint is the well-known de Broglie Bohm pilot wave theory [1]. For more modern approaches seeking to bypass these claims, see e.g., [2, 3].

However, even in the case of  $[\beta] = 1$  there is a quantum solution, e.g., over 2-qubits as

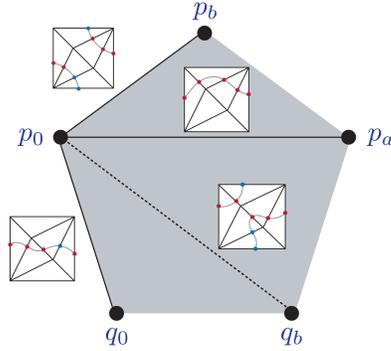


Fig. 1: Local structure  $MP_1$  at the type 1 vertex  $q_0$  and type 2 vertex  $p_0$ . The former has only type 2 neighbors, a single orbit under the action of the stabilizer of the vertex. The latter has both type 1 (single orbit) and type 2 neighbors (breaks into two orbits with representatives  $p_b$  and  $p_a$ ). Edges in the polytope are represented by loops on the Mermin torus.  $q_b$  can be connected to  $p_0$  by a path corresponding to a loop but is not a neighbor.

given in Fig. (2b). The quantity  $[\beta]$  is, in fact, cohomological, as first observed in [9]. The cohomological perspective is based on reorganizing the scenario into a space. Then the Mermin scenario is represented as a torus; see Fig. (3b). In this representation, measurements label the edges of the triangles, and  $\beta$  assigns a value in  $\mathbb{Z}_2$  to each triangle. Choosing a quantum state induces a nonsignaling distribution on the Mermin scenario with support on each context  $C$  consisting of the set  $O_\beta(C)$  of outcome assignments  $s : C \rightarrow \mathbb{Z}_2$  that satisfy  $\sum_{m \in C} s(m) = \beta(C)$ . Let  $NS_C$  denote the nonsignaling polytope for the Mermin scenario. We introduce a



Fig. 2: Mermin scenario with  $\beta = 1$  for the context indicated in red color.

subpolytope<sup>e</sup> called the *Mermin polytope*,

$$MP_\beta \subset NS_C \tag{2}$$

<sup>e</sup>Throughout the paper by a subpolytope we mean a polytope contained in the larger polytope as a subset, not necessarily the convex hull of a subset of vertices.

that consists of nonsignaling distributions, that is, tuples of  $p = (p_C)_{C \in \mathcal{C}}$  probability distributions compatible under marginalization, such that the support of each  $p_C$  is contained in  $O_\beta(C)$ . We show that the combinatorial isomorphism type of the polytope  $MP_\beta$  is determined by  $[\beta]$ . As canonical representatives for  $[\beta] = 0$  and 1 we take the choices of  $\beta$ 's given in Fig. (3a) and Fig. (3c); respectively. The resulting Mermin polytopes will be denoted by  $MP_0$  and  $MP_1$ . One of our main technical contributions is the classification of the vertices of these two polytopes.

**Theorem 1** *Let  $MP_\beta$  denote the Mermin polytope.*

1. *All the vertices of  $MP_0$  are deterministic distributions corresponding to the functions*

$$s : \{m_{00}, m_{01}, m_{10}, m_{11}\} \rightarrow \mathbb{Z}_2.$$

*There are 16 vertices.*

2. *For  $MP_1$  the vertices are given by pairs  $(\Omega, s)$  where  $\Omega \subset M$  is a maximal closed non-contextual (cnc) set and  $s : \Omega \rightarrow \mathbb{Z}_2$  is an outcome assignment. There are two types of vertices:*

- *Type 1: When  $\Omega$  is of type 1. There are 48 vertices of this type.*
- *Type 2: When  $\Omega$  is of type 2. There are 72 vertices of this type.*

Although the Mermin polytopes are numerically tractable (e.g., using Polymake [10]), our goal in this paper is to identify structures in the combinatorial description of these polytopes. Our vertex classification result relies on the symmetries of the Mermin polytopes. We identify a subgroup  $G_\beta$  of the combinatorial automorphisms of  $MP_\beta$ . We show that  $G_0$  acts transitively on the vertices of  $MP_0$ . This means that for any pair of vertices, there is a symmetry of the polytope that moves one to the other. For  $MP_1$  the symmetry group  $G_1$  acts transitively within each type of vertices. We also study the stabilizer group of the vertices, that is, symmetry elements that fix a given vertex, and the action of this group on the neighbor vertices to obtain a description of the graph associated to the polytopes. In the graph of  $MP_1$  the main structural elements are the loops on the Mermin torus that give the edges of the graph connecting a pair of neighbor vertices; see Fig. (1). We hope that these new structural elements will help us understand more complicated polytopes, such as the close relatives of  $MP_1$  known as  $\Lambda$ -polytopes [11].

**Theorem 2** *Let  $MP_\beta$  denote the Mermin polytope.*

1. *The graph of  $MP_0$  is the complete graph  $K_{16}$ .*
2. *The graph of  $MP_1$  consists of 120 vertices and the local structure at the type 1 and 2 vertices is depicted in Fig. (1).*

Let us put our result into context: The polytope  $MP_0$  can be best studied within the framework of simplicial distributions introduced in [12]. In this framework, nonsignaling distributions can be interpreted as distributions on spaces, as in Fig. (5b). The nonsignaling conditions are encoded at the faces of the triangles. The Mermin scenario can be regarded as an extension of the well-known Clauser, Horne, Shimony, Holt (CHSH) [13] scenario, a Bell

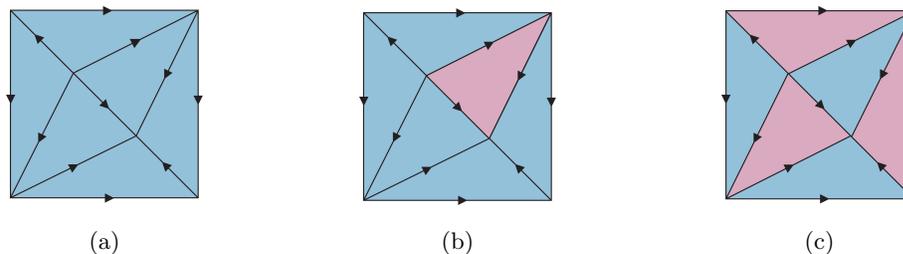


Fig. 3: Mermin scenario represented as a torus: Top and bottom (left and right) edges are identified. In this representation  $\beta$  assigns  $\{0, 1\}$  to each triangle. Red color indicates that  $\beta = 1$ ; otherwise  $\beta = 0$ .

scenario consisting of two parties and two measurements  $x_i, y_j$ , where  $i, j \in \mathbb{Z}_2$ , for each party with binary outcomes; see Fig. (5a). A fundamental result for the CHSH scenario is Fine's theorem [14]. This theorem says that a distribution on the CHSH scenario is noncontextual if and only if the CHSH inequalities are satisfied. Our vertex classification for  $MP_0$  can be turned into a new topological proof of Fine's theorem. This proof diverges from Fine's original argument; see [12, Thm. 4.12] for an alternative topological proof closer to Fine's original argument. We present our topological proof of Fine's theorem in Section 5.1.

The other polytope,  $MP_1$ , can be seen as a toy model of a more complicated polytope introduced in [11] for classical simulation of universal quantum computation. For  $n$ -qubits, the polytope  $\Lambda_n$  used in this classical simulation is defined as the polar dual of the  $n$ -qubit stabilizer polytope. These polytopes are only fully understood in the case of a single qubit:  $\Lambda_1$  is a 3-dimensional cube containing the Bloch sphere. The combinatorial structure of  $\Lambda_n$  for  $n \geq 2$  is yet to be understood. This mathematical problem is the main obstacle to quantifying the complexity of the  $\Lambda$ -simulation algorithm, a fundamental question in the study of quantum computational advantage. The next case,  $\Lambda_2$ , is only understood numerically (e.g., using Polymake [10]). A geometric understanding of  $\Lambda_2$  will bring insight into the structure of  $\Lambda$ -polytopes with higher number of qubits. Tensoring a vertex of  $\Lambda_2$  with an  $(n-2)$ -qubit stabilizer state produces a vertex in  $\Lambda_n$  [15, Theorem 2]. Some of the vertices of  $\Lambda_2$  are similar to the vertices of  $MP_1$ . These vertices are also described by cnc sets [16]. In fact, the Mermin polytope  $MP_1$  can be seen as a nonlocal version of  $\Lambda_2$ . The local part is captured by the nonsignaling polytope  $NS_{232}$  of the two party Bell scenario, consisting of two measurements with binary outcomes per parties. Our decomposition result provides a description of  $\Lambda_2$  in terms of two well-understood polytopes:  $NS_{232}$  whose vertices are described in [17] and  $MP_1$  described in Theorem 3. Vertices of  $MP_1$  are also described in [18] in the dual form by solving the facet enumeration problem for the polytope given by the convex hull of the non-local stabilizer states; see also [19] for applications to quantum computational universality.

Our main contributions in this paper can be summarized as follows:

- We define families of Mermin polytopes parametrized by a function  $\beta$  and classify the corresponding polytopes by the cohomology class  $[\beta]$  (Proposition 1).
- The symmetry groups  $G_\beta$  of each equivalence class of Mermin polytopes are described and we demonstrate that they are isomorphic (Proposition 3).

- A complete characterization of the vertices for both classes of Mermin polytopes is given (Theorem 3).
- $G_0$  acts transitively on the vertices of  $MP_0$  (Lemma 13) and  $G_1$  acts transitively on the vertices of  $MP_1$  of a fixed type (Lemma 15). The latter result also describes the stabilizers of each type of vertices.
- Graphs of both Mermin polytopes are described (Theorem 4 and Theorem 5).
- We exploit the relationship between the Mermin and CHSH scenarios to provide a new topological proof of Fine's theorem [14, 20] (Theorem 6). An important step is the vertex classification for  $MP_0$ , which implies that any distribution on the Mermin torus is noncontextual (Corollary 3).
- The  $\Lambda_2$  polytope is decomposed into local and nonlocal polytopes. The former is a well-known nonsignaling polytope  $NS_{232}$  [17], while the nonlocal part is precisely the Mermin polytope (Theorem 7).

The rest of the paper is organized as follows. In Section 2 we formalize the Mermin scenario and the notion of Mermin polytopes. In Section 3 we characterize the vertices of the Mermin polytopes. In Section 4 we describe the graphs of the polytopes. In Section 5 we apply the vertex characterization to problems in quantum foundations and quantum computation. More involved proofs for Propositions 1 and 3 can be found in Appendices A and B, respectively. Appendix C contains the description of the stabilizer groups of the vertices of  $MP_1$ .

## 2 Mermin polytopes

Mermin polytopes mentioned in this paper are certain subpolytopes of nonsignaling polytopes associated to the Mermin scenario. In this section we introduce these polytopes formally and show that up to combinatorial isomorphism of polytopes there are two types denoted by  $MP_0$  and  $MP_1$ . Our main result is a classification theorem for the vertices of these polytopes.

### 2.1 Definition

A *measurement scenario*, or more briefly a *scenario*, consists of the following data:

- a set  $M$  of measurements,
- a collection  $\mathcal{C}$  of subsets  $C \subset M$ , called contexts, that cover the whole set of measurements, i.e.

$$M = \cup_{C \in \mathcal{C}} C,$$

- a set of outcomes, which through the paper is fixed as  $\mathbb{Z}_2 = \{0, 1\}$ .

Since the outcome set is fixed we will write  $(M, \mathcal{C})$  to denote a scenario. For a set  $U$  we will write  $\mathbb{Z}_2^U$  for the set  $\{s : U \rightarrow \mathbb{Z}_2\}$  of functions on a context  $C \in \mathcal{C}$ . The *nonsignaling polytope* on this scenario, denoted by  $NS_{\mathcal{C}}$ , consists of collections  $(p_C)_{C \in \mathcal{C}}$  of probability distributions, each given by a function  $p_C : \mathbb{Z}_2^C \rightarrow \mathbb{R}_{\geq 0}$  where  $\sum_{m \in \mathbb{Z}_2^C} p(m) = 1$ , satisfying the nonsignaling condition given by

$$p_C|_{C \cap C'} = p_{C'}|_{C' \cap C} \quad \forall C, C' \in \mathcal{C}.$$

The restriction  $p_C|_{C \cap C'}$  corresponds to marginalization of the distribution to the intersection. A distribution  $p$  is called *noncontextual* if there exists a distribution  $d : \mathbb{Z}_2^M \rightarrow \mathbb{R}_{\geq 0}$  such that  $p_C = d|_C$  for all  $C \in \mathcal{C}$ . Otherwise,  $p$  is called *contextual*. For more details see [21]. We will write  $\text{NS}_{\mathcal{C}}$  for the polytope of nonsignaling distributions on the scenario  $(M, \mathcal{C})$ .

We are interested in polytopes associated to binary linear systems [8]. A *binary linear system* consists of a scenario  $(M, \mathcal{C})$  together with a function  $\beta : \mathcal{C} \rightarrow \mathbb{Z}_2$ . For each  $C$  we will write

$$O_{\beta}(C) = \{s : C \rightarrow \mathbb{Z}_2 : \sum_{m \in C} s(m) = \beta(C)\} \subset \mathbb{Z}_2^C.$$

A function in this set will be referred to as an *outcome assignment* on the context  $C$ . We introduce a subpolytope

$$\text{NS}_{\mathcal{C}, \beta} \subset \text{NS}_{\mathcal{C}} \quad (3)$$

that consists of nonsignaling distributions  $p = (p_C)_{C \in \mathcal{C}}$  such that

$$\text{supp}(p_C) \subset O_{\beta}(C) \quad \forall C \in \mathcal{C}$$

where  $\text{supp}(p_C)$  stands for the support of  $p_C$ , i.e., the set of functions  $s : C \rightarrow \mathbb{Z}_2$  such that  $p_C(s) > 0$ .

**Definition 1** The *Mermin scenario* consists of

- the measurement set  $M = \{m_{ij} : i, j \in \mathbb{Z}_3\}$ , and
- the cover  $\mathcal{C}$  given by two types of contexts:
  - Horizontal:  $C^h = \{C_i^h : i \in \mathbb{Z}_3\}$  where  $C_i^h = \{m_{ij} : j \in \mathbb{Z}_3\}$ ,
  - Vertical:  $C^v = \{C_j^v : j \in \mathbb{Z}_3\}$  where  $C_j^v = \{m_{ij} : i \in \mathbb{Z}_3\}$ .

The *Mermin polytope* for a function  $\beta : \mathcal{C} \rightarrow \mathbb{Z}_2$  is defined to be  $\text{MP}_{\beta} = \text{NS}_{\mathcal{C}, \beta}$ . Analogously we can consider quasiprobability distributions on the Mermin scenario with restricted support. We will write  $\text{MP}_{\beta}^{\mathbb{R}}$  for this polytope.

In this paper we will study the Mermin polytope associated to the Mermin scenario  $(M, \mathcal{C})$ .

## 2.2 Topological representation

In [7] it was shown that the Mermin scenario can be represented by a torus with a certain triangulation. In this representation contexts are represented by triangles. We will follow the more recent approach developed in [12] to represent nonsignaling distributions in a topological way. Given a context  $C = \{x, y, z\}$  in  $\mathcal{C}$  we represent the distribution  $p_C$  as in Fig. (4). For a measurement  $x$  we write  $p_x^0$  for the probability of measuring outcome 0. Similarly given a pair  $x, y$  of measurements  $p_{xy}^{ab}$  denotes the probability for the outcome assignment  $(x, y) \mapsto (a, b)$ . Given a triangle with a probability distribution as in Fig. (4) the probabilities at the  $x, y$  edges are given by

$$\begin{aligned} p_x^0 &= p^{01} + p^{00} \\ p_y^0 &= p^{10} + p^{00}. \end{aligned}$$

Fig. (4a) represents the case where  $\beta = 0$ . In this case

$$p_z^0 = p^{00} + p^{11}, \quad (4)$$

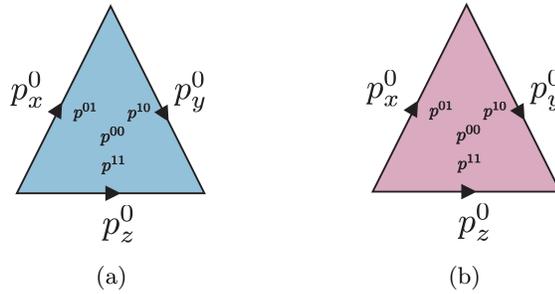


Fig. 4: (a) Triangle with  $\beta = 0$ . (b) Triangle with  $\beta = 1$ . The marginal at  $z$  is given by Eq. (4), or Eq. (5); respectively.

whereas if  $\beta = 1$  as in Fig. (4b) then

$$p_z^0 = p^{01} + p^{10}. \tag{5}$$

Therefore, in effect  $z$  is the XOR measurement  $x \oplus y$  in the first case, and the NOT of the XOR measurement  $\overline{x \oplus y}$  in the second.

In this description the triangle with edges  $(x, y)$  is our measurement space and the outcome space is kept implicit for convenience. In general a simplicial distribution [12] is defined on a pair  $(X, Y)$  of spaces where  $X$  and  $Y$  represent the space of measurements and outcomes, respectively. The model used to represent these spaces is called simplicial sets. A simplicial set  $X$  consists of a sequences of sets  $X_n$  indexed over the natural numbers  $n \geq 0$ . These sets represent the set of  $n$ -simplices and they are related by the face maps<sup>f</sup>encoding the gluing relations. Given an outcome space  $Y$  there is a way to obtain another simplicial set denoted by  $D(Y)$  whose  $n$ -simplices are distributions on  $Y_n$ . Then a simplicial distribution is given by a simplicial set map  $p : X \rightarrow D(Y)$ , i.e., a sequence of functions  $X_n \rightarrow D(Y_n)$  assigning to each  $n$ -simplex  $\sigma$  of  $X$  a distribution  $p_\sigma$  on  $Y_n$ . These assignments are compatible with the gluing relations of both spaces. In this paper we restrict to measurement spaces obtained by gluing triangles and our outcome space is fixed to the nerve space of  $\mathbb{Z}_2$ . Our pictorial description in the rest of the paper will suffice to describe the relevant polytopes. See [22] for more examples of simplicial distributions on 2-dimensional measurement spaces.

In Fig. (3) Mermin scenario with various choices of  $\beta$ 's are represented on a torus. In this framework,  $\beta$  assigns 0 or 1 to each triangle, hence can be interpreted as a cochain from algebraic topology. The value given by the sum in Eq. (1) has a special meaning in this context known as the cohomology class of  $\beta$ . In this paper we don't assume familiarity with cochains, or with other topological notions such as cohomology; see [9] for more on the cohomological perspective.

**Proposition 1** *Given two functions  $\beta, \beta' : \mathcal{C} \rightarrow \mathbb{Z}_2$  the Mermin polytope  $MP_\beta$  is combinatorially isomorphic to  $MP_{\beta'}$  if and only if  $[\beta] = [\beta']$ .*

<sup>f</sup>There is also a second type of maps called degeneracy maps that encode the “collapsing relations”. We will ignore them for simplicity of the presentation.

Proof of this result is given in Appendix A. As a consequence there are two types of Mermin polytopes, up to combinatorial isomorphism, corresponding to the cases  $[\beta] = 0$  and 1.

**2.3 The even case:**  $MP_0$

Let  $\beta_0 : \mathcal{C} \rightarrow \mathbb{Z}_2$  denote the function defined by

$$\beta_0(C) = 0, \quad \forall C \in \mathcal{C}. \tag{6}$$

We will simply write  $MP_0$  to denote the Mermin polytope  $MP_{\beta_0}$ . Note that this notation is justified by the observation that the isomorphism type of  $MP_{\beta}$  only depends on  $[\beta]$  as proved in Proposition 1. Our goal in this section is to relate this polytope to a famous bipartite Bell scenario, usually referred to as the CHSH scenario.

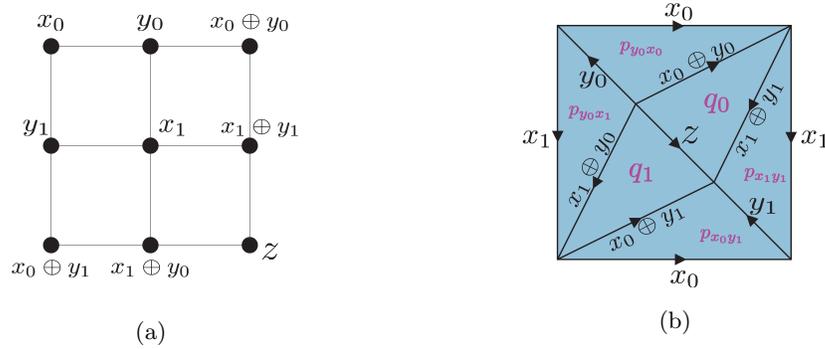


Fig. 5: (a) Mermin scenario in the conventional representation. Vertices correspond to measurement labels. (b) Mermin scenario in the topological representation. Measurements label the edges.

The CHSH scenario is a particular type of Bell scenario for 2 parties, 2 measurements per party and 2 outcomes per measurement. More precisely, this scenario consists of

- the measurement set  $\{x_i, y_j : i, j \in \mathbb{Z}_2\}$  where  $x_i$ 's are for Alice and  $y_j$ 's are for Bob, and
- the contexts  $\{x_i, y_j\}$  where  $i, j \in \mathbb{Z}_2$ .

Mermin scenario can be obtained from the CHSH scenario by adding two additional contexts  $\{x_0 \oplus y_0, x_1 \oplus y_1, z\}$  and  $\{x_0 \oplus y_1, x_1 \oplus y_0, z\}$ , where  $z = x_0 \oplus y_0 \oplus x_1 \oplus y_1$ , consisting of the XOR's of the measurements of Alice and Bob; see Fig. (5a). See Fig. (5b) for a topological

representation. For the convenience of the reader we list the nonsignaling conditions

$$\begin{aligned}
 p_{x_0}^0 &= p_{y_0x_0}^{10} + p_{y_0x_0}^{00} = p_{x_0y_1}^{01} + p_{x_0y_1}^{00} \\
 p_{y_0}^0 &= p_{y_0x_0}^{01} + p_{y_0x_0}^{00} = p_{y_0x_1}^{01} + p_{y_0x_1}^{00} \\
 p_{x_1}^0 &= p_{y_0x_1}^{10} + p_{y_0x_1}^{00} = p_{x_1y_1}^{01} + p_{x_1y_1}^{00} \\
 p_{y_1}^0 &= p_{x_0y_1}^{10} + p_{x_0y_1}^{00} = p_{x_1y_1}^{10} + p_{x_1y_1}^{00} \\
 p_{x_0 \oplus y_0}^0 &= p_{y_0x_0}^{11} + p_{y_0x_0}^{00} = q_0^{01} + q_0^{00} \\
 p_{x_1 \oplus y_0}^0 &= p_{y_0x_1}^{11} + p_{y_0x_1}^{00} = q_1^{01} + q_1^{00} \\
 p_{x_0 \oplus y_1}^0 &= p_{x_0y_1}^{11} + p_{x_0y_1}^{00} = q_1^{10} + q_1^{00} \\
 p_{x_1 \oplus y_1}^0 &= p_{x_1y_1}^{11} + p_{x_1y_1}^{00} = q_0^{10} + q_0^{00} \\
 p_z^0 &= q_0^{11} + q_0^{00} = q_1^{11} + q_1^{00}.
 \end{aligned} \tag{7}$$

**Proposition 2** *A distribution  $p$  on the CHSH scenario is noncontextual if and only if it extends to a distribution on the Mermin scenario.*

**Remark 1** This result first appeared in [12]. Its proof relies on Fine’s theorem characterizing noncontextual distributions using the CHSH inequalities. We will provide a proof of this result independent of Fine’s theorem (see Proposition 5) by describing all the vertices of  $MP_0$ . Then this observation will be used to provide a new topological proof of Fine’s theorem.

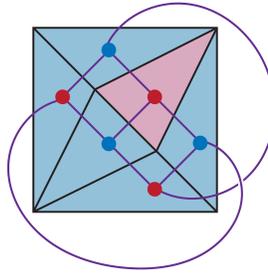


Fig. 6: Graph of  $K_{3,3}$ .

Next we discuss the symmetries of  $MP_0$ . For a polytope  $P$  let  $\text{Aut}(P)$  denote the *group of combinatorial automorphisms* of the polytope. We begin by describing certain elements of this symmetry group. First we consider a graph obtained from the Mermin scenario. The vertices of this graph are given by the contexts, i.e.,  $\mathcal{C} = \mathcal{C}^h \sqcup \mathcal{C}^v$ , and the edges are given by the set  $M$  of measurements. The resulting graph is the *bipartite complete graph*  $K_{3,3}$ ; see Fig (6). The automorphism group  $\text{Aut}(K_{3,3})$  of this graph is generated by the following operations [23]:

- (1) Permutation of the vertices in  $\mathcal{C}^h$  while keeping  $\mathcal{C}^v$  fixed.
- (2) Permutation of the vertices in  $\mathcal{C}^v$  while keeping  $\mathcal{C}^h$  fixed.

(3) The permutation exchanging

$$\begin{aligned} C_1^v &\leftrightarrow C_0^h \\ C_2^v &\leftrightarrow C_2^h \\ C_0^v &\leftrightarrow C_1^h \end{aligned}$$

Denoting the symmetric group on  $n$  letters by  $\Sigma_n$  the symmetry group can be expressed as a semidirect product

$$\text{Aut}(K_{3,3}) = (\Sigma_3 \times \Sigma_3) \rtimes \mathbb{Z}_2.$$

Each factor represents a type of symmetry given in (1), (2) and (3); respectively. Geometrically the symmetry operation (3) corresponds to a reflection about the diagonal in the torus; see Fig. (5b).

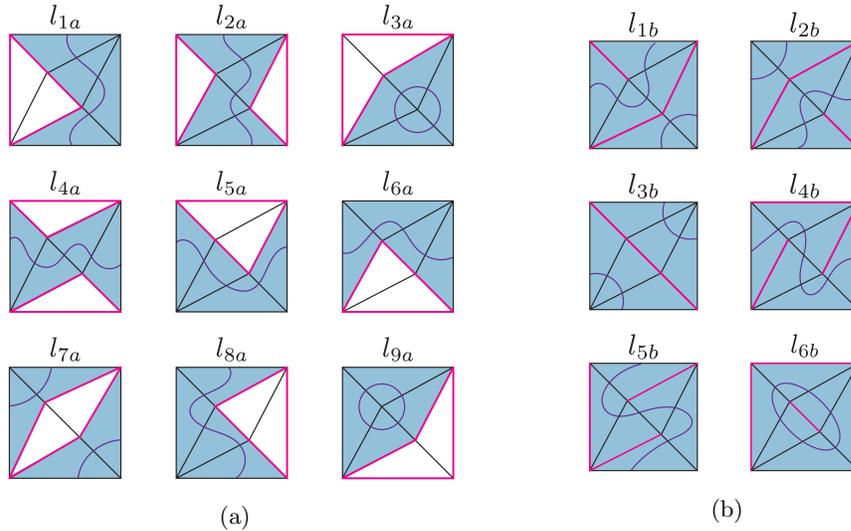


Fig. 7: There are two types of loops denoted by  $l_{1a}, l_{2b}, \dots, l_{9a}$  and  $l_{1b}, l_{2b}, \dots, l_{6b}$ .

Another kind of symmetry of  $\text{MP}_0$  comes from flipping the outcomes of the measurements in  $M$ . Each such symmetry operation can be represented by a loop in the graph  $K_{3,3}$ . Let  $\ell(K_{3,3})$  denote the set of loops on the graph; see Fig. (7) for the complete list of loops. For each such loop  $l$  there is a group element  $g_l$  that acts on  $\text{MP}_0$  by flipping the outcomes of the measurements that live on the loop. Let  $G_\ell$  denote the subgroup of  $\text{Aut}(\text{MP}_0)$  generated by the elements  $g_l$  for  $l \in \ell(K_{3,3})$ .

**Lemma 1**  $G_\ell$  is isomorphic to  $\mathbb{Z}_2^4$  with the canonical generators given by the loops  $l_{x_0} = l_{2a}$  (flipping  $x_0$ ),  $l_{x_1} = l_{4a}$  (flipping  $x_1$ ),  $l_{y_0} = l_{9a}$  (flipping  $y_0$ ) and  $l_{y_1} = l_{3a}$  (flipping  $y_1$ ).

*Proof.* Proof of this result follows from directly verifying that  $g_l$  for each loop  $l$  in Fig. (7) can be decomposed into a product of these canonical generators. For example,  $g_{l_{1a}} = g_{l_{x_0}} g_{l_{y_1}}$  and  $g_{l_{1b}} = g_{l_{x_0}} g_{l_{y_1}} g_{l_{x_1}}$ . Similarly for the remaining loops.  $\square$

We will write  $G_0$  for the subgroup of  $\text{Aut}(\text{MP}_0)$  generated by the two subgroups  $\text{Aut}(K_{3,3})$  and  $G_\ell$ . This group can be expressed as an extension

$$0 \rightarrow G_\ell \rightarrow G_0 \rightarrow \text{Aut}(K_{3,3}) \rightarrow 1, \tag{8}$$

that is  $G_\ell$  is a normal subgroup of  $G_0$  and the quotient group  $G_0/G_\ell$  is given by  $\text{Aut}(K_{3,3})$ .

**2.4 The odd case:  $\text{MP}_1$**

Let  $\beta_1 : \mathcal{C} \rightarrow \mathbb{Z}_2$  denote the function defined by

$$\beta_1(C) = 0, \quad \forall C \in \mathcal{C}^h \quad \text{and} \quad \beta_1(C) = 1, \quad \forall C \in \mathcal{C}^v. \tag{9}$$

We will write  $\text{MP}_1$  for the the Mermin polytope  $\text{MP}_{\beta_1}$ . (This notation is justified by Proposition 1.) Our goal in this section is to provide a quantum mechanical description of  $\text{MP}_1$ . Using this description we will study the symmetries of the polytope.

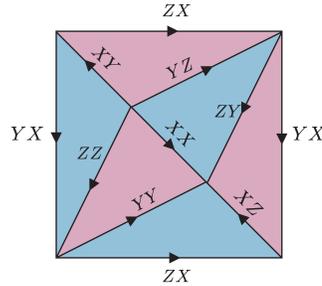


Fig. 8: Mermin square whose edges are labeled by nonlocal Pauli operators. The tensor product is omitted from the notation.

The connection to quantum theory is via the notion of binary linear systems. A *quantum solution* to the Mermin square binary linear system  $(M, \mathcal{C}, \beta)$  consists of unitary operators  $A_{ij} \in U(\mathcal{C}^d)$  where  $i, j \in \mathbb{Z}_3$  such that

- $A_{ij}^2 = \mathbb{1}$  for all  $i, j \in \mathbb{Z}_3$ ,
- $\{A_{ij} : i \in \mathbb{Z}_3\}$  and  $\{A_{ij} : j \in \mathbb{Z}_3\}$  consist of pairwise commuting unitaries,
- $A_{i0}A_{i1}A_{i2} = (-1)^{\beta(C_i^h)}$  and  $A_{0j}A_{1j}A_{2j} = (-1)^{\beta(C_j^v)}$  for all  $i, j \in \mathbb{Z}_3$ .

A quantum solution over  $U(\mathcal{C})$  is called a *classical solution*. It is known that a classical solution exists if and only if  $[\beta] = 0$ . This can either be proved directly by an argument similar to Mermin’s proof of contextuality [7], or by cohomological arguments [9]. Nonexistence of a classical solution is an indication of quantum contextuality in the sense of Kochen–Specker. As in the case of Mermin’s proof, quantum solutions can come from Pauli operators. The *n-qubit Pauli operators*<sup>9</sup> are given by

$$A = A_1 \otimes A_2 \otimes \cdots \otimes A_n, \quad A_i \in \{\mathbb{1}, X, Y, Z\},$$

<sup>9</sup>We only consider the ones whose eigenvalues are  $\pm 1$ .

where  $\mathbb{1}, X, Y, Z$  are the  $2 \times 2$  Pauli matrices. The *Pauli group*, denoted by  $P_n$ , consists of operators of the form  $i^\alpha A$  where  $\alpha \in \mathbb{Z}_4$ .

We partition the set of 2-qubit Pauli operators into local and nonlocal parts:

- Local 2-qubit Pauli operators

$$\{X \otimes \mathbb{1}, Y \otimes \mathbb{1}, Z \otimes \mathbb{1}, \mathbb{1} \otimes X, \mathbb{1} \otimes Y, \mathbb{1} \otimes Z\}.$$

- Nonlocal 2-qubit Pauli operators

$$\{X \otimes X, X \otimes Y, X \otimes Z, Y \otimes X, Y \otimes Y, Y \otimes Z, Z \otimes X, Z \otimes Y, Z \otimes Z\}.$$

For  $\beta_1$  defined as in Eq. (9) nonlocal Pauli operators constitute a quantum solution; see Fig. (8). For a pair  $A, B$  of distinct and commuting Pauli operators let  $\Pi_{AB}^{ab}$  denote the projector onto the simultaneous eigenspace corresponding to the eigenvalues  $(-1)^a$  and  $(-1)^b$  of  $A$  and  $B$ ; respectively. More concretely, we have  $\Pi_{AB}^{ab} = (\mathbb{1} + (-1)^a A + (-1)^b B + (-1)^{a+b} AB)/4$ . These projectors constitute the set  $\mathcal{S}_2$  of 2-qubit stabilizer states. There is a corresponding local vs nonlocal decomposition:

$$\mathcal{S}_2 = \mathcal{S}_2^{(l)} \sqcup \mathcal{S}_2^{(nl)} \tag{10}$$

where

- $\mathcal{S}_2^{(l)}$  consists of projectors  $\Pi_{AB}^{ab}$  where  $A, B$  are local Pauli operators.
- $\mathcal{S}_2^{(nl)}$  consists of projectors  $\Pi_{AB}^{ab}$  where  $A, B$  are nonlocal Pauli operators.

**Lemma 2** *MP<sub>1</sub> can be identified with the set of Hermitian operators  $\rho \in \text{Herm}((\mathbb{C}^2)^{\otimes n})$  of trace 1 such that  $\text{Tr}(\rho A) = 0$  for all local Pauli's  $A$  and  $\text{Tr}(\rho \Pi_{AB}^{ab}) \geq 0$  for all pairwise commuting nonlocal Pauli operators  $A, B$  and  $a, b \in \mathbb{Z}_2$ .*

*Proof.* Let  $Q$  denote the set of operators  $\rho$  described in the statement. The Born rule gives a map  $p : Q \rightarrow \text{MP}_1$  sending  $\rho \mapsto p_\rho$ . If we know  $p_\rho^{ab}$  for all  $a, b \in \mathbb{Z}_2$  then we can compute the expectation  $\langle A \rangle_\rho$  for any nonlocal Pauli. Since by assumption  $\langle B \rangle_\rho = 0$  for every local Pauli  $B$  this way we can determine  $\rho$ . In other words, we can define a map  $e : \text{MP}_1 \rightarrow Q$  by sending a distribution  $d$  to the operator

$$\rho_d = \frac{1}{4} \left( \mathbb{1} + \sum_A \alpha_A A \right)$$

where  $A$  runs over nonlocal Pauli's and  $\alpha_A$  is the expectation obtained from  $d$ . Then  $e$  is the inverse of  $p$ . Therefore  $p$  is a bijection.  $\square$

Lemma 2 provides a quantum mechanical description of  $\text{MP}_1$ . In particular, some of the symmetries of  $\text{MP}_1$  come from quantum mechanics, that is, by conjugation with a Clifford unitary. The *n-qubit Clifford group*  $\text{Cl}_n$  is the quotient of the normalizer of  $P_n$ , the group of unitaries  $U \in U((\mathbb{C}^2)^{\otimes n})$  such that  $UAU^\dagger \in P_n$  for all  $A \in P_n$ , by the central subgroup

$\{e^{i\theta}\mathbb{1} : 0 \leq \theta < 2\pi\}$ . Acting by the elements of  $Cl_1$  on each qubit preserves the set of nonlocal Pauli's. By Lemma 2 this group acts on the polytope  $MP_1$ . Additionally, the SWAP gate that permutes the parties is also a symmetry of the polytope. Let us define the following subgroup of  $Cl_2$ :

$$G_1 = \langle Cl_1 \times Cl_1, SWAP \rangle. \tag{11}$$

As we observed this is also a subgroup of  $Aut(MP_1)$ . Next we will express  $G_1$  as an extension similar to the one for  $G_0$  given in Eq. (8). First recall that  $Cl_1$  has two parts: the Pauli part isomorphic to  $\mathbb{Z}_2^2$  generated by conjugation with  $X$  and  $Z$ , and the symplectic part  $Sp_2(\mathbb{Z}_2)$ . The latter group is isomorphic to  $\Sigma_3$  since in the single qubit case the symplectic action is determined by the permutation of the subgroups  $\langle -\mathbb{1}, X \rangle$ ,  $\langle -\mathbb{1}, Y \rangle$ ,  $\langle -\mathbb{1}, Z \rangle$ . We can express this decomposition as an exact sequence

$$0 \rightarrow \mathbb{Z}_2^4 \rightarrow G_1 \rightarrow Aut(K_{3,3}) \rightarrow 1. \tag{12}$$

The quotient is given by  $Aut(K_{3,3})$  since  $G_1$  (up to signs) permutes the set of contexts which in return induces as an action on the graph  $K_{3,3}$ . By comparing sizes we conclude that the quotient group is the whole automorphism group of the graph.

**Proposition 3** *There is an isomorphism of groups  $\phi : G_1 \rightarrow G_0$ .*

The proof can be found in Appendix B.

### 3 Vertices of the Mermin polytopes

The description of Mermin polytopes is most naturally given in terms of the intersection of a finite number of half-spaces, or  $H$ -representation. However, by the Minkowski-Weyl theorem [24] there is an equivalent representation of a polytope in terms of the convex hull of a finite number of vertices, called the  $V$ -representation. The problem of switching from the  $H$  to the  $V$ -description is called the vertex enumeration problem; see e.g., [25]. Here we do precisely this and enumerate the vertices of  $MP_\beta$ , using the rich structure of these polytopes to aid in this task.

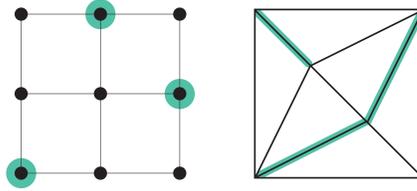
#### 3.1 Closed noncontextual subsets

We recall some definitions from [16].

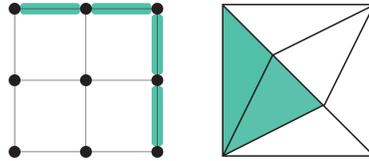
**Definition 2** *A subset  $\Omega \subset M$  is called closed if  $M \cap C \neq \emptyset$  implies  $C \subset M$ . An outcome assignment on a closed subset  $\Omega$  is a function  $s : \Omega \rightarrow \mathbb{Z}_2$  such that*

$$s(c) = s(a) + s(b) + \beta(a, b)$$

*for all  $C = \{a, b, c\} \subset \Omega$ . A closed subset  $\Omega$  is called noncontextual if it admits an outcome assignment. We call subsets  $\Omega$  obeying both of these properties closed noncontextual, or cnc sets. And as noted in [16], it suffices to consider just the maximal cnc sets, which we do from here forth.*



(a) Type 1 cnc set: All edges correspond to anti-commuting operators, thus never in the same triangle.



(b) Type 2 cnc set: Two adjacent triangles.

Fig. 9

As we observed in the description of  $MP_1$  (see Section 2.4) contexts of the Mermin scenario can be realized by the commutation relation of nonlocal Pauli operators. Next we make this connection more precise. Recall that the measurements  $m_{ij} \in M$  are labeled by pairs  $(i, j) \in \mathbb{Z}_3^2$ . We consider a map  $\iota : \mathbb{Z}_3 \rightarrow \mathbb{Z}_2^2 - \{0\}$  defined as follows:

$$\iota(0) = (0, 1), \quad \iota(1) = (1, 1), \quad \iota(2) = (1, 0).$$

The corresponding Pauli operators are  $T_{(0,0)} = \mathbb{1}$ ,  $T_{(0,1)} = X$ ,  $T_{(1,1)} = Y$  and  $T_{(1,0)} = Z$ . On the other hand, 2-qubit Pauli operators can be labeled by  $(v, w) \in E = \mathbb{Z}_2^2 \times \mathbb{Z}_2^2$ , which corresponds to  $T_v \otimes T_w$ . This way we obtain an embedding

$$M \rightarrow E, \quad m_{ij} \mapsto (\iota(i), \iota(j)). \tag{13}$$

Throughout we will use this identification. Given  $(v, w)$  and  $(v', w')$  there is a symplectic form

$$[(v, w), (v', w')] = v \cdot w' + v' \cdot w \pmod{2}.$$

We say that such a pair *commutes* if  $[(v, w), (v', w')] = 0$ ; otherwise we say that they *anticommute*. A subspace  $I \subset E$  is called *isotropic* if each pair of elements in this subspace commute. Observe that contexts in  $\mathcal{C}$  are precisely the maximal isotropic subspaces of  $M$ .

**Lemma 3** *The structure of maximal closed noncontextual subsets of  $M$  with respect to  $\beta_1$  is given as follows:*

- *Type 1: the subset  $\Omega$  consists of three distinct pairwise anticommuting elements; i.e., none lie within the same context. We have 6 such sets  $\Omega$  and an outcome assignment  $s : \Omega \rightarrow \mathbb{Z}_2$  is a function; see Fig. (9a).*
- *Type 2: the subset  $\Omega$  is a union of two distinct contexts with a single measurement  $m \in M$  lying on their (nonempty) intersection, and hence consists of 5 elements total.*

There are 9 such subsets  $\Omega$ , one for each  $m \in M$ . Additionally, there are 3 elements that generate the set and an outcome assignment  $s : \Omega \rightarrow \mathbb{Z}_2$  is determined by a function on these 3 generators; see Fig. (9b).

*Proof.* This result follows from Raussendorf, et al. [16], specifically Lemma 3. The case of the Mermin square is treated in more detail in Example 2. Quoting their results, there are  $48 = 2^3 \times 6$  Type 1 cnc sets and  $72 = 2^3 \times 9$  Type 2 cnc sets.  $\square$

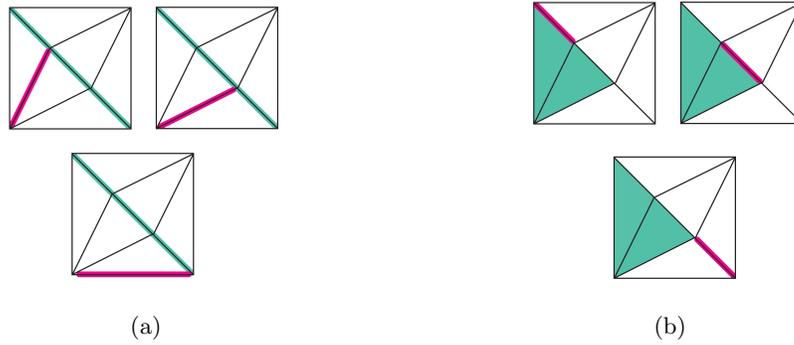


Fig. 10: Type 1 and 2 cnc sets indicated by green color and the nonlocal Pauli's that flip only one outcome indicated by red color.

**Lemma 4**  $G_1$  acts transitively on the set

$$\{(\Omega, s) : \Omega \text{ is a maximal cnc set of type } k, \text{ and } s : \Omega \rightarrow \mathbb{Z}_2 \text{ is an outcome assignment}\}$$

where  $k$  is either 1, or 2.

*Proof.* We begin with type 1 cnc sets. First let us ignore the outcome assignment. A type 1 cnc set is specified by three pairwise anticommuting nonlocal Pauli operators. There are 6 of these sets, three are of the form  $\{A \otimes X, A \otimes Y, A \otimes Z\}$  and the remaining three are of the form  $\{X \otimes A, Y \otimes A, Z \otimes A\}$ , where  $A = X, Y, Z$ . To move from one such cnc set to another one one can use a local Clifford unitary and the SWAP gate if needed. Now including the outcome assignments, for a fixed cnc set  $\Omega$  we can move from one outcome assignment to the other by flipping the signs of the outcomes by conjugating with a nonlocal Pauli operator that commutes with two of them, but anticommutes with the remaining one; see Fig. (10a).

In the case of type 2 cnc sets we can move between the  $\Omega$  sets since the  $\text{Aut}(K_{3,3})$  quotient of  $G_1$  acts transitively on the edges of the torus (or the dual  $K_{3,3}$  graph) [26, Sec. 3.2]. To move between the outcome assignments on a fixed type 2 cnc set  $\Omega$  we can conjugate with a nonlocal Pauli; see Fig. (10b).  $\square$

**Corollary 1**  $\text{Aut}(K_{3,3})$  acts transitively on the set of maximal cnc sets of a fixed type.

*Proof.* As we observed in the proof of Lemma 4 the  $G_1$  action factors through the action of the quotient group  $\text{Aut}(K_{3,3})$  when the outcome assignments are ignored.  $\square$

### 3.2 Vertex classification

In the rest of this section we will prove the following result. Recall the embedding  $M \subset E$  given in Eq. (13).

**Theorem 3** *There is a bijection between the set of vertices of  $MP_\beta$  and the set of functions  $s : \Omega \rightarrow \mathbb{Z}_2$  satisfying the following properties:*

(i) *For  $MP_0$  the subset  $\Omega = M$  and the functions are group homomorphisms  $s : M \rightarrow \mathbb{Z}_2$ . Each such function is given by specifying  $s(v_i, w_j) \in \mathbb{Z}_2$  for  $v_i, w_j \in \{(0, 1), (1, 1)\}$ . In particular, there are 16 vertices.*

(ii) *For  $MP_1$  the subset  $\Omega$  is a maximal closed noncontextual subset (one of the two types in Lemma 3) and  $s : \Omega \rightarrow \mathbb{Z}_2$  is an outcome assignment. There are two types of vertices:*

- *Type 1: When  $\Omega$  is of type 1. In particular, there are 48 vertices of this type.*
- *Type 2: When  $\Omega$  is of type 2. In particular, there are 72 vertices of this type.*

Given a function  $s : \Omega \rightarrow \mathbb{Z}_2$  as in (i) or (ii) the corresponding vertex  $p \in MP_\beta$  is uniquely determined by

$$p_m^0 = \begin{cases} \frac{1+(-1)^{s(m)}}{2} & \forall m \in \Omega \\ 0 & \text{otherwise} \end{cases} . \quad (14)$$

We begin with some recollections from polytope theory; see [24, 25]. Let  $P(A, b) = \{x : Ax \geq b\}$  denote a polytope where  $A \in \mathbb{R}^{n \times m}$  and  $b \in \mathbb{R}^n$ . Assume that  $P(A, b) \subset \mathbb{R}^m$  is full dimensional. Let us establish some terminology. If an inequality is satisfied with equality then we call that inequality *tight*. For a point  $p \in P(A, b)$  we refer to the *active set* at  $p$  as a subset  $\mathcal{Z}_p \subset \{1, \dots, n\}$  which indexes the set of tight inequalities at  $p$ . A point  $p$  is a vertex of  $P(A, b)$  if and only if there exists a subset of tight inequalities  $Z \subseteq \mathcal{Z}_p$  with  $|Z| = m$  such that

$$v = A[Z]^{-1}b,$$

where  $A[Z]$  is the matrix obtained from  $A$  by removing all the rows whose index is not in  $Z$ . Note that  $|Z| \leq |\mathcal{Z}_p|$ .

Let us apply these observations to Mermin polytopes  $MP_\beta$ . We can express  $MP_\beta$  in the form  $P(A, b)$ . Let us write  $x \circ y$  to mean the XOR measurement  $x \oplus y$  if  $\beta(x, y) = 0$ , or the NOT of the XOR measurement  $\overline{x \oplus y}$  if  $\beta(x, y) = 1$ .

**Proposition 4** *The Mermin polytope  $MP_\beta$  has a description in the form  $P(A, b)$  where  $b = -\mathbb{1}_{9 \times 1}$ ,  $A \in \mathbb{R}^{24 \times 9}$  is a matrix whose rows are labeled by the set*

$$S = \{(C, ab) : C \in \mathcal{C}, a, b \in \mathbb{Z}_2\},$$

columns labeled by  $M$  (once both sets are ordered) and for  $C = \{x, y, x \circ y\}$  its entries are given by

$$A_{(C,ab),m} = \begin{cases} (-1)^a & m = x \\ (-1)^b & m = y \\ (-1)^{a+b+\beta(C)} & m = x \circ y \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Before we proceed to proving Proposition 4, let us prove the following useful lemma:

**Lemma 5** *The distribution  $p_C^{ab}$  in a single triangle (see Fig. (4)) with edges labeled  $\{x, y, z\}$  and outcomes  $a, b, c \in \mathbb{Z}_2$ , respectively, with  $c = a + b + \beta_C$ , is uniquely determined by the marginals along the edges  $\{p_{\{x\}}^a, p_{\{y\}}^b, p_{\{z\}}^c\}$  according to*

$$p_C^{ab} = \frac{1}{2} \left( p_{\{x\}}^a + p_{\{y\}}^b - p_{\{z\}}^{c+1} \right). \tag{16}$$

*Proof.* First note that we have:

$$p_{\{x\}}^0 = p_C^{00} + p_C^{01} \tag{17}$$

$$p_{\{y\}}^0 = p_C^{00} + p_C^{10} \tag{18}$$

$$p_{\{z\}}^{\beta_C} = p_C^{00} + p_C^{11} \tag{19}$$

from which the normalization condition  $\sum_{ab} p_C^{ab} = 1$  becomes:

$$p_C^{00} + p_C^{01} + p_C^{10} + p_C^{11} = p_{\{x\}}^0 + p_{\{y\}}^0 + p_{\{z\}}^{\beta_C} - 2p_C^{00} = 1.$$

From this we can obtain:

$$\begin{aligned} p_C^{00} &= \frac{1}{2} \left( p_{\{x\}}^0 + p_{\{y\}}^0 + p_{\{z\}}^{\beta_C} - 1 \right) \\ &= \frac{1}{2} \left( p_{\{x\}}^0 + p_{\{y\}}^0 - p_{\{z\}}^{\beta_C+1} \right), \end{aligned} \tag{20}$$

where in the second line we used  $p_{\{z\}}^{c+1} = 1 - p_{\{z\}}^c$ . Equation (16) then follows by inserting Eq. (20) into Eqns. (17)-(19) and solving for the remaining  $p_C^{ab}$ .  $\square$

**Remark 2** Notice that if two contexts  $C$  and  $C'$  intersect on an edge  $m = C \cap C'$  then distributions  $p_C^{ab}, p_{C'}^{ab}$ , represented as in Eq. (16) will automatically satisfy the nonsignaling conditions if one and the same marginal  $p_{\{m\}}^a$  is used in both.

The 24 probabilities  $p_C^{ab}$  can therefore be uniquely expressed by the marginal probabilities  $p_{\{m\}}^a$ , where  $m \in M$ . In particular, any  $p_C^{ab}$  can be expressed by just the 0-outcome marginals  $p_{\{m\}}^0$  since their complement is given by  $p_{\{m\}}^1 = 1 - p_{\{m\}}^0$ . These nine marginal probabilities

therefore serve as a system of coordinates for  $MP_\beta$ , which can be embedded in  $\mathbb{R}^9$ .

We now introduce a new set of coordinates in terms of the *expectation values* of the measurement outcomes, denoted  $\bar{m}$ . The two are related by an affine transformation. The expectation value of a measurement  $m \in M$  is then given by

$$\bar{m} := \sum_a (-1)^a p_{\{m\}}^a = p_{\{m\}}^0 - p_{\{m\}}^1.$$

Using that  $p_{\{m\}}^1 = 1 - p_{\{m\}}^0$  and solving for  $p_{\{m\}}^0$  we obtain the desired relationship

$$p_{\{m\}}^0 = \frac{1}{2}(1 + \bar{m}). \tag{21}$$

noindent [Proof of Proposition 4] Note that  $MP_\beta$  is defined as the intersection of the half-space inequalities  $p_C^{ab} \geq 0$  intersected by the affine subspace generated by the nonsignaling conditions and normalization. By Lemma 5 this is equivalent to requiring the nonnegativity of Eq. (16), for every  $C \in \mathcal{C}$  and  $a, b \in \mathbb{Z}_2$ . Plugging in Eq. (21) for  $p_{\{m\}}^0$  in terms of  $\bar{m}$  gives us the expression

$$p_C^{ab} = \frac{1}{4} (1 + (-1)^a \bar{x} + (-1)^b \bar{y} + (-1)^{a+b+\beta} \bar{z}). \tag{22}$$

Requiring nonnegativity and rearranging yields

$$A_{(C,ab)}^T x \geq -1 \tag{23}$$

where  $A_{(C,ab)}$  is given as in Eq. (15) and  $x \in \mathbb{R}^9$  has components  $\bar{m}$ , where  $m \in M$ . The 24 inequalities defining the polytope can now be compactly expressed as  $Ax \geq b$ , which concludes the proof.

Let  $Z \subset S$  be a subset of indices such that  $|Z| = 9$ . For each  $C \in \mathcal{C}$  let us write  $n(C) = |Z \cap \{(C, ab) : a, b \in \mathbb{Z}_2\}|$ . The numbers  $n(C)$  satisfy the following properties:

- $\sum_{C \in \mathcal{C}} n(C) = 9$  since  $|Z| = 9$ .
- $0 \leq n(C) \leq 3$  since  $\sum_{a,b \in \mathbb{Z}_2} p_C^{ab} = 1$  for each context.

Our case classification will be in terms of the following numbers:

$$n_k = |\{C \in \mathcal{C} : n(C) = k\}|$$

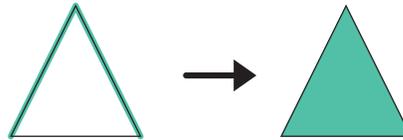
Table (1) displays all the cases that can occur. These cases will be denoted by  $(n_3, n_2)$ .

A triangle representing a context  $C$  is called a *deterministic triangle* if  $p_C$  is a deterministic distribution. An edge labeled by a measurement  $x \in C$  is called a *deterministic edge* if  $p_C|_{\{x\}}$  is a deterministic distribution.

$n_3$	$n_2$	$n_1$
3	0	0
2	1	1
2	0	3
1	3	0
1	2	2
1	1	4
0	3	3

Table 1: Each row displays the number of contexts with the indicated number of zeros. These are the triples  $(n_3, n_2, n_1)$  satisfying  $0 \leq n_i \leq 3$  and  $3n_3 + 2n_2 + n_1 = 9$ .

**Lemma 6** *A triangle  $C$  with two deterministic edges is deterministic.*



*Proof.* We can assume  $\beta = 0$  on the triangle, the case  $\beta = 1$  is treated similarly. Let  $p_C = \{p_C^{ab}\}_{a,b}$  be a distribution on the triangle where  $C = \{x, y, x \oplus y\}$ . Assume that  $p|x = \delta_x^a$  and  $p|_y = \delta^b$  for some  $a, b \in \mathbb{Z}_2$ . This implies

$$p^{\bar{a}0} + p^{\bar{a}1} = p^{0\bar{b}} + p^{1\bar{b}} = 0,$$

where  $\bar{a} = a + 1$ . In every case three of the four probabilities are zero giving us a deterministic distribution. Other cases where  $x, x \oplus y$  and  $y, x \oplus y$  are deterministic are also treated similarly  $\square$

**Lemma 7** *Let  $p$  be a distribution on a single triangle  $C$ . Let  $Z_C = \{A_{(C,ab)} : a, b \in \mathbb{Z}_2\}$  be the set of tight inequalities. Then  $\text{rank}(A[Z_C]) = |Z_C|$ .*

*Proof.* We can assume  $\beta = 0$ , the case  $\beta = 1$  is similar. Assume that  $Z_C$  is nonempty, otherwise the rank is zero. Let us write  $C = \{x, y, x \oplus y\}$ . We will use the symmetry group  $G_C$  generated by flipping the outcomes of  $x, y$ , which is isomorphic to  $\mathbb{Z}_2^2$ . By Lemma 6 there are two cases (up to  $G_C$  action):

- Single deterministic edge:  $p|_{\{x\}} = \delta^0$ . In this case

$$\text{rank}(A) = \text{rank} \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} = 2, \tag{24}$$

where a row corresponds to outcomes of  $(x, y, x \oplus y)$  in this order.

- Deterministic triangle:  $p = \delta^{00}$ . In this case

$$\text{rank}(A) = \text{rank} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix} = 3. \tag{25}$$

□

Next, we consider distributions on the *diamond*  $D$ , which is obtained by gluing two triangles,  $C$  and  $C'$ , along a common edge. We will denote the common edge by  $z$ . See Fig. (11a).

**Lemma 8** *Let  $p$  be a distribution on the diamond  $D$ . Let  $Z_C, Z_{C'}$  be the set of tight inequalities in triangles  $C, C'$ ; respectively. Define  $Z_{C,C'} = Z_C \cup Z_{C'}$ . Then*

$$\text{rank}(A[Z_{C,C'}]) = \begin{cases} |Z_{C,C'}| - 1 & z \text{ is deterministic,} \\ |Z_{C,C'}| & \text{otherwise.} \end{cases}$$

*Proof.* We assume  $\beta = 0$  on both triangles, the case  $\beta = 1$  on one of the triangles is treated similarly. Assume that  $Z_{C,C'}$  is nonempty, otherwise the rank is zero. We will also use symmetries to reduce the number of cases. Let us write  $C = \{x, z, x \oplus z\}$  and  $C' = \{x', z, x' \oplus z\}$  for the contexts. Let  $G_D$  denote the symmetry group generated by flipping the outcomes of  $x, z, x'$ , which is isomorphic to  $\mathbb{Z}_2^3$ . First, let us assume that  $z$  is not deterministic. The cases where either  $Z_C$  or  $Z_{C'}$  are empty can be deduced from Lemma 7. By Lemma 6 the remaining cases are as follows (up to  $G_D$  action):

- $|Z_C| = |Z_{C'}| = 1$  with  $p_C^{00} = p_{C'}^{00} = 0$ . In this case

$$\text{rank}(A[Z_{C,C'}]) = \text{rank} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} = 2, \tag{26}$$

where a row corresponds to outcomes of  $(x, x \oplus z, z, x', x' \oplus z)$  in this order.

- $|Z_C| = 2, |Z_{C'}| = 1$  with  $p_C|_{\{x\}} = \delta^0$  and  $p_{C'}^{00} = 0$ . In this case

$$\text{rank}(A[Z_{C,C'}]) = \text{rank} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} = 3. \tag{27}$$

The case  $|Z_C| = 1, |Z_{C'}| = 2$  is similar.

- $|Z_C| = 2, |Z_{C'}| = 2$  with  $p_C|_{\{x\}} = p_{C'}|_{\{x'\}} = \delta^0$ . In this case

$$\text{rank}(A[Z_{C,C'}]) = \text{rank} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & 1 & -1 \end{bmatrix} = 4. \tag{28}$$

Next, we consider the case where  $z$  is deterministic. Again up to the action of  $G_D$  we have the following cases:

- $|Z_C| = 2, |Z_{C'}| = 2$  with  $p_C|_{\{z\}} = p_{C'}|_{\{z\}} = \delta^0$ . In this case

$$\text{rank}(A[Z_C, C']) = \text{rank} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & -1 \end{bmatrix} = 3. \tag{29}$$

- $|Z_C| = 3, |Z_{C'}| = 2$  with  $p_C = \delta^{00}$  and  $p_{C'}|_{\{z\}} = \delta^0$ . In this case

$$\text{rank}(A[Z_C, C']) = \text{rank} \begin{bmatrix} 1 & -1 & -1 & 0 & 0 \\ -1 & 1 & -1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & -1 \end{bmatrix} = 4. \tag{30}$$

The case  $|Z_C| = 2, |Z_{C'}| = 3$  is similar.

- $|Z_C| = 3, |Z_{C'}| = 3$  with  $p_C = p_{C'} = \delta^{00}$ . In this case

$$\text{rank}(A[Z_C, C']) = \text{rank} \begin{bmatrix} 1 & -1 & -1 & 0 & 0 \\ -1 & 1 & -1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & -1 & -1 & 1 \end{bmatrix} = 5. \tag{31}$$

□

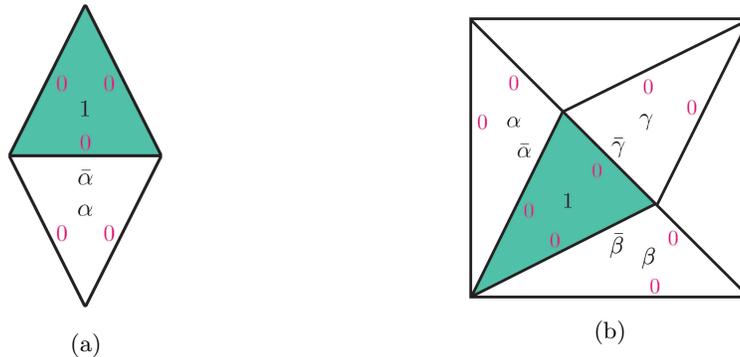


Fig. 11: (a) Two triangles glued along a common edge. (b) A distribution on the torus with one deterministic triangle.

**Lemma 9** *Let  $p$  be a distribution in  $MP_\beta$  and  $Z$  denote the set of tight inequalities. Assume that there exists a deterministic triangle  $C$ . Then  $\text{rank}(A[Z]) \geq 6$ .*

*Proof.* Considering the action of  $G_\beta$  will simplify the discussion. Our argument does not depend on  $\beta$ , so we assume  $\beta = 0$ . Up to the action of the symmetry group we can assume

that  $p_C$  has the form given in Fig. (11b). Let us write  $C = \{x, y, z\}$  and  $C_t = \{t, t', t \oplus t'\}$  where  $t = x, y, z$  for the adjacent triangles. Then

$$\text{rank}(A[Z]) = \text{rank} \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} = 6, \quad (32)$$

where a row corresponds to outcomes of  $(x, y, z, x', x \oplus x', y', y \oplus y', z', z \oplus z')$ . Lemma 7 and Lemma 8 can be used to compute the rank. We are looking at a region obtained by gluing three diamonds along a triangle. The rank of the matrix above is the sum of the ranks of each diamond minus two times the rank of the deterministic triangle.  $\square$

**Remark 3** Lemma 9 implies that to obtain a vertex we must fix three additional (linearly independent) zeros. This forces at least one additional edge to be deterministic, which by Lemma 6 implies that we must have at least two adjacent deterministic triangles.

We will use the number of deterministic triangles and the number of deterministic edges (that lie out side the boundary of the triangles) to organize the cases. In Fig. (12) we see a diagram that illustrates all the possibilities. The base cases consist of three deterministic edges. Successive application of Lemma 6 together with Lemma 9 and Remark 3 reduces the diagram to three main cases:

- (C1) All triangles are deterministic.
- (C2) Two adjacent deterministic triangles.
- (C3) Three anticommuting deterministic edges.

Note that if there are more than three deterministic edges again Lemma 6 can be used to reduce to (C1). The case where there is only one deterministic triangle and no additional deterministic edges does not appear since Lemma 8 implies that such a configuration can have rank at most 8.

**Remark 4** Note that up to  $\text{Aut}(K_{3,3})$  there are only three representative cases. For (C1) this is obvious since all triangles are deterministic. For cases (C2) and (C3) we observe that these correspond to type 2 and type 1 cnc sets, respectively. Thus by Corollary 1 it suffices to consider a single representative for each case. The representatives are given in Fig. (12).

**Lemma 10** Assume  $p \in MP_\beta$  is a vertex that satisfies (C1). Then  $p$  belongs to  $MP_0$ . No distribution in  $MP_1$  is deterministic.

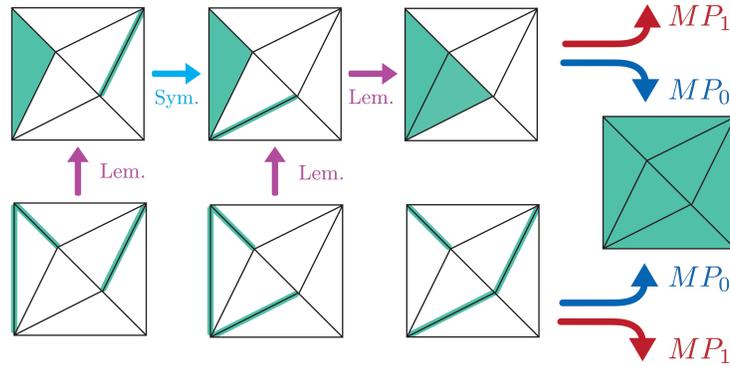


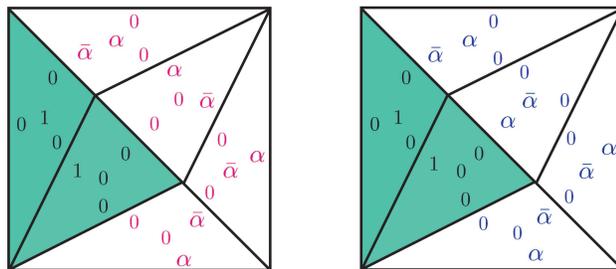
Fig. 12: Starting from the base case of three deterministic edges, we can obtain the vertices of  $MP_\beta$  by repeated application of Lemma 6. Up to symmetry we have three representative cases: (C1) All deterministic triangles (C2), two adjacent deterministic triangles, and (C3) three anti-commuting deterministic edges. For  $\beta = 0$  all of these cases lead to a deterministic distribution. For  $\beta = 1$  deterministic distributions are not allowed and we have type 1 and type 2 cnc distributions corresponding to cases (C2) and (C3), respectively.

*Proof.* First let us note that (C1) implies that  $A[Z]$  has full rank. To see this, take three mutually nonadjacent (i.e., the set of edges  $C_i \cap C_j$  is empty) triangles as deterministic, which implies that all edges are deterministic. By applying Lemma 7 for each triangle (after an appropriate permutation of columns) we have that  $A[Z]$  has full rank. Next observe that a set of deterministic edges implies a classical solution to the binary linear system  $(M, \mathcal{C}, \beta)$ . Since this is possible only for  $[\beta] = 0$ , we have (C1) defines a vertex of  $MP_0$ , but not of  $MP_1$ .  $\square$

**Lemma 11** *Let  $p$  be a vertex of  $MP_\beta$  that satisfies (C2). Then either*

- (i)  $p$  satisfies (C1), or
- (ii)  $p$  is a type 2 vertex of  $MP_1$ .

*Proof.* Consider a configuration for (C2), which specifies a cnc set  $\Omega$  of type 2. We can fix a deterministic distribution on  $\Omega$ . Any other choice can be dealt with similarly by the help of Lemma 4. Then we use the compatibility conditions, as shown below:



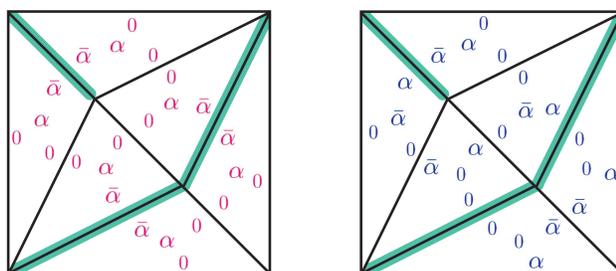
Here we have  $\beta_0$  ( $\beta_1$ ) in red (blue). For  $\beta = 0$  there is a one parameter family of distributions. A vertex is specified by choosing  $\alpha = 0, 1$ , which implies a deterministic distribution

and thus reduces to (C1). For  $\beta = 1$  the compatibility conditions imply that  $\alpha = 1 - \alpha = 1/2$ .

**Lemma 12** *Let  $p$  be a vertex of  $MP_\beta$  that satisfies (C3). Then either*

- (i)  $p$  satisfies (C1), or
- (ii)  $p$  is a type 1 vertex of  $MP_1$ .

*Proof.* Similar to the case (C2) let us consider a configuration, choose a convenient distribution consistent with the case (C3) (other choices can be handled using symmetry, i.e., Lemma 4), and solve for the probabilities using the compatibility conditions:



As with (C2), here for  $\beta = 0$  we have a one-parameter family of distributions (red) where vertices are specified by  $\alpha = 0, 1$ , reducing to the deterministic case (C3). For  $\beta = 1$  we have that  $\alpha = 1 - \alpha = 1/2$  (blue).  $\square$

[**Proof of Theorem 3**] To begin, the diagram in Fig. (12) implies that we need only consider cases (C1)-(C3). Focusing first on  $\beta = 0$ , Lemmas 10-12 imply that all vertices of  $MP_0$  are deterministic. These vertices are determined by the marginals on the measurements  $(v_i, w_j)$  where  $v_i, w_j \in \{(0, 1), (1, 1)\}$ . Hence there are 16 such vertices.

Turning now to  $MP_1$ , note that by Lemma 10 that no deterministic distribution is a point of  $MP_1$ , thus by Lemmas 11 and 12, the only vertices of  $MP_1$  are those of the form of (C2) and (C3). Observe that for (C2) and (C3) that  $p_{\{m\}}^0 \in \{0, 1\}$  (i.e., the edge is deterministic) if and only if  $m \in \Omega$ , where  $\Omega$  are the maximal cnc sets described in Lemma 3, and  $p_{\{m\}}^0 = 1/2$  (or  $\bar{m} = 0$ ) for all other observables  $m \notin \Omega$ . For example, the deterministic edges in (C3) are described by a type 1 cnc set since they correspond to a maximal set of anti-commuting observables. Using Lemma 3, we know that there are  $48 = 2^3 \times 6$  type 1 and  $72 = 2^3 \times 9$  type 2 cnc sets, which then correspond to 48 type 1 and 72 type 2 vertices of  $MP_1$ , respectively.  $\square$

#### 4 Graph of the Mermin polytopes

In this section, we determine the graph of  $MP_\beta$  consisting of the vertices of the polytope together with the edges connecting two neighbor vertices in the polytope.

##### 4.1 Graph of $MP_0$

**Lemma 13**  $G_0$  acts transitively on deterministic vertices of  $MP_0$ .

*Proof.* Take an arbitrary deterministic vertex  $p \in MP_0$  and act on it by  $G_\ell \subset G_0$ . There are 15 elements of  $G_\ell$  listed in Fig. (7a) and the action of each permutes the outcomes of a different subset of measurements and thus generates 15 distribution distinct from  $p$ . Since there are  $16 = 2^4$  outcome assignments in total, we obtain all possible deterministic distributions by the action of  $G_0$   $\square$

Let  $q$  denote the deterministic distribution in  $MP_0$  given by  $q_C^{00} = 1$  for all triangles  $C$ ; see Fig. (13). We will take this as the canonical vertex of this polytope. The other vertices can be obtained by using the action of the loops as a consequence of Lemma 13. We will write

$$q_l = g_l \cdot q \quad \text{where } g_l \in G_\ell, l \in \ell(K_{3,3})$$

for the remaining vertices obtained via the action of  $G_\ell$  (see Eq. (8)).

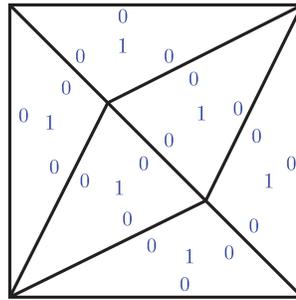


Fig. 13

**Corollary 2** *Let  $p$  be a vertex of  $MP_0$ . Then  $Stab_{G_0}(p)$  is isomorphic to  $Aut(K_{3,3})$ . Moreover, the stabilizer acts transitively on the set of remaining vertices.*

*Proof.* By Lemma 13 we know that the action of  $G_0$  on the set of vertices is transitive. Therefore the stabilizers of each vertex are isomorphic. It suffices to compute the stabilizer of the canonical vertex  $q$ . By definition of  $q$ , permutation of the contexts does not change it. That is,  $Aut(K_{3,3}) \subset Stab_{G_0}(p)$ . Since there are 16 vertices, this implies  $|G_0/Stab_{G_0}(p)| = 16$  and we have  $Aut(K_{3,3}) = Stab_{G_0}(p)$ .

For the second part of the statement observe that the set of edges in a loop is precisely the complement  $\Omega^c$  of a maximal cnc set (Definition 2). Therefore there is a one-to-one correspondence between the set of loops and the set of maximal cnc sets (both types combined). Since  $Aut(K_{3,3})$  acts transitively on the set of cnc sets (Corollary 1), it also acts transitively on the set  $\ell(K_{3,3})$  of loops. This implies that the action of the stabilizer, that is  $Aut(K_{3,3})$ , on the vertices  $\{g_l \cdot q : l \in \ell(K_{3,3})\}$  is transitive since  $\sigma \cdot q_l = g_{\sigma \cdot l} \cdot q$  for  $\sigma \in Aut(K_{3,3})$ , where  $\sigma \cdot l$  is the loop obtained by the permutation action of  $\sigma$ .  $\square$

**Theorem 4** *The graph of  $MP_0$  is the complete graph  $K_{16}$ .*

*Proof.* Let us consider  $q$  and another vertex  $q_l = g_l \cdot q$ . By Corollary 2 we can assume  $l = l_{x_0}$  corresponding to flipping the outcome of  $x_0$ ; see Lemma 1. It suffices to show that  $q$  and  $q_l$  are neighbors. The distribution  $p(\alpha) = \alpha q + \bar{\alpha} q_l$ , where  $0 \leq \alpha \leq 1/2$ , is given as follows:

The diagram shows three square diagrams representing distributions. The first square on the left is labeled with  $\alpha$  and  $\bar{\alpha}$  at its vertices and contains various 0 and 1 values along its edges and diagonals. This square is equal to the sum of two other squares. The second square is labeled with  $\alpha$  and contains 0 and 1 values. The third square is labeled with  $\bar{\alpha}$  and contains 0 and 1 values.

Note that  $p(\alpha)$  for  $\alpha \in (0, 1/2)$  specifies an edge in  $MP_0$  from  $q$  to  $q_l$  since the rank of  $A[Z]$ , where  $Z$  is the set of tight inequalities, is equal to 8. This is because, the zeros in  $Z$  together with the nonsignaling conditions leaves a single parameter, that is  $\alpha$ .  $\square$

### 4.2 Graph of $MP_1$

Our goal is to describe the graph of  $MP_1$ . We will follow a similar approach to the vertex classification. This time we consider 8 linearly independent inequalities instead of 9. Considering the number of deterministic edges on the torus representation is a good way to organize the cases. Our main technical result describes an edge between two neighboring vertices of  $MP_1$  in terms of the loops on the torus given in Fig. (7). We begin by introducing some notation: We have seen that the complement of a loop  $l \in \ell(K_{3,3})$  corresponds to a cnc set (Definition 2). Denoting a maximal cnc set that corresponds to loop  $l$  by  $\Omega_l$  we will write  $\Omega_l^c$  for its complement, consisting of the edges that belong to the loop  $l$ . A *signed loop* consists of a loop together with a function

$$\varphi : \Omega_l^c \rightarrow \{\pm 1\}.$$

Corresponding to this function we will define a collection  $p^\varphi = (p^\varphi)_{C \in \mathcal{C}}$  of functions  $p^\varphi : \mathbb{Z}_2^{\mathcal{C}} \rightarrow \mathbb{R}$  such that  $\sum_s p^\varphi(s) = 0$ . Note that this is similar to a distribution but the values sum to zero instead of one, and can be negative. For  $C = \{x, y, z\}$  our definition of  $p_C^\varphi$  uses a version of Eq. (22):

$$(p_C^\varphi)^{ab} = \frac{1}{4}((-1)^a \varphi(x) + (-1)^b \varphi(y) + (-1)^{a+b+\beta(C)} \varphi(z)).$$

**Lemma 14** *Let  $p$  be a distribution in  $MP_1$  and  $Z$  denote a subset of tight inequalities such that  $|Z| = 8$ . If  $\text{rank}(A[Z]) = 8$  then there exists precisely two deterministic edges. Moreover, an edge  $p \in MP_1$  between two vertices  $q_1$  and  $q_2$  is given by*

$$p(\alpha) = q_1 + \alpha p^\varphi, \quad \alpha \in [0, 1/2], \tag{33}$$

for some signed loop  $\varphi : \Omega_l^c \rightarrow \{\pm 1\}$ , where  $l \in \ell(K_{3,3})$ , such that  $p(1/2) = q_2$ .

*Proof.* Let us start with the case of no deterministic edges. In this case  $A[Z] \leq 6$ , hence we don't have enough zeros to obtain an edge in the polytope. When there is one deterministic edge, say denoted by  $z$ , we consider the diamond  $D$  consisting of two adjacent triangles  $C, C'$  at  $z$ . By Lemma 8 we have  $\text{rank}(A[Z_{C,C'}]) = |Z_{C,C'}| - 1$  which implies  $\text{rank}(A[Z]) \leq 7$ . The

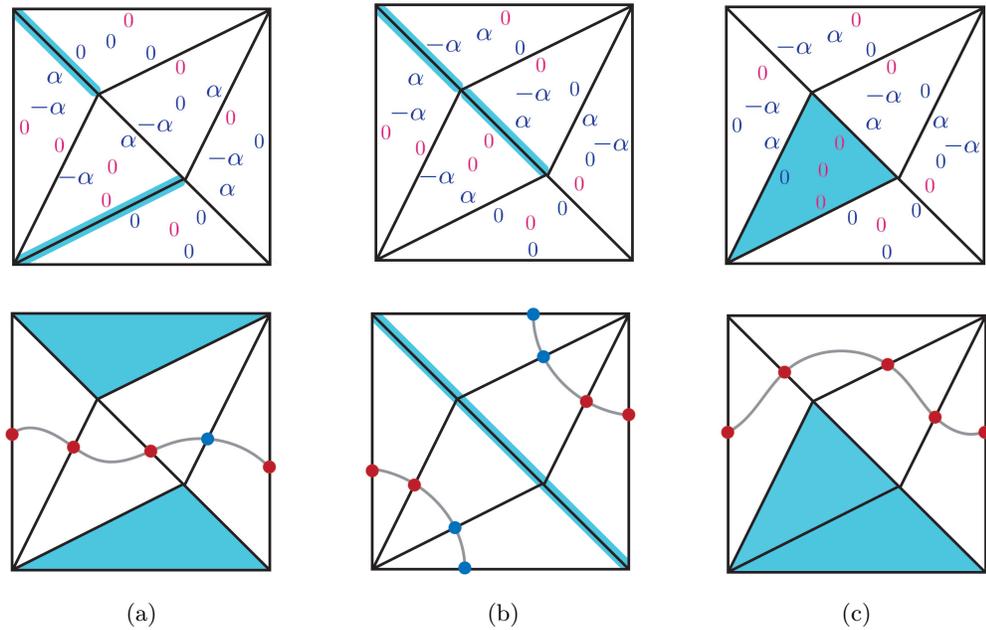


Fig. 14: An edge (top) in  $MP_1$  between (a) canonical type 2 vertex  $p_0$  and canonical type 1 vertex  $q_0$  (b)  $p_0$  and  $p_b$  (c)  $p_0$  and  $p_a$ , as well as the corresponding loops (bottom). The edges are parameterized by  $\alpha \in [0, 1/2]$  such that  $\alpha = 0$  corresponds to  $p_0$  and  $\alpha = 1/2$ .

case of three deterministic edges, or more, with at least two of them anticommuting is studied in Section 3. According to Lemmas 10-12 we obtain either a vertex of  $MP_1$  or a distribution that lies outside of this polytope. Remaining cases are two deterministic edges which either commute or anticommute. Note that by Lemma 6 the commuting case also covers the three pairwise commuting deterministic edges. To establish Eq. (33) we note that two distributions  $q_1$  and  $q_2$  are connected by an edge if and only if they have in common 8 linearly independent tight inequalities preserved along the edge. Given such a set of tight inequalities, we proceed to construct  $p^\varphi$  by placing the corresponding zeros on the torus and then use the compatibility conditions together with the fact that  $\sum_s p^\varphi(s) = 0$ .

To see how this works, let us consider a representative  $\Omega = \{x, y\} \subset M$  for the case of two (a) anti-commuting and (b) commuting deterministic edges (see Fig. (15)) and notice that these are both cnc sets, although not maximal. Moreover, let us choose a value assignment  $s : \Omega \rightarrow \mathbb{Z}_2$ , which by Eq. (14) determines the marginals  $p_{\{x\}}$  and  $p_{\{y\}}$ . By Lemma 4 the action of  $G_1$  on the set of pairs  $(\Omega, s)$  is transitive. Even though  $\Omega$  is not maximal we can always embed it into a maximal one, extend  $s$  and apply the transitivity of the action of  $G_1$ .

In both cases, as depicted in Fig. (15a) and (15b), there are 6 linearly independent tight inequalities, thus we must choose two additional probabilities to set to zero. The possible choices are as follows: (1) Set two (or one) of the given parameters  $\alpha, \beta, \gamma = 0, 1$ . (2) Place both remaining zeros in a single shaded triangle. (3) Place one zero in each of the shaded triangles. It is straightforward to see that both options (1) and (2) will fix the distribution to be a specific vertex, and thus will not be an edge. For (3) we let  $p$  and  $q$  denote the

distribution on the shaded triangles. In Fig. (15a) suppose  $p$  corresponds to the triangle whose boundary has marginals for the outcome 0 given by  $(-\alpha, -\beta, -\gamma)$  and  $q$  corresponds to  $(-\alpha, \beta, -\gamma)$ . Then from these marginals one can compute

$$\begin{aligned} p^{01} &= -\alpha - p^{00} \\ p^{10} &= -\beta - p^{00} \\ p^{11} &= \gamma - p^{00} \end{aligned}$$

and similarly for  $q$ . Using  $\sum_{a,b} p^{ab} = \sum_{c,d} q^{cd} = 0$  and solving for  $p^{ab}$  and  $q^{cd}$  we obtain

$$\begin{aligned} p^{00} &= (-\alpha - \beta + \gamma)/2 & q^{00} &= (-\alpha + \beta + \gamma)/2 = -p^{10} \\ p^{01} &= (-\alpha + \beta - \gamma)/2 & q^{01} &= (-\alpha - \beta - \gamma)/2 = -p^{11} \\ p^{10} &= (\alpha - \beta - \gamma)/2 & q^{10} &= (\alpha + \beta - \gamma)/2 = -p^{00} \\ p^{11} &= (\alpha + \beta + \gamma)/2 & q^{11} &= (\alpha - \beta + \gamma)/2 = -p^{01}. \end{aligned}$$

If  $p^{ab}$  is set to zero then we can set one of  $q^{cd}$ , where  $(c, d) \in \mathbb{Z}_2^2 - \{a+1, b\}$ , equal to zero. In this way we obtain a type 2 loop. For example, setting  $p^{01} = q^{10} = 0$  gives the signed loop Fig. (14c) For this choice  $p(1/2)$  is the vertex  $p_a$ .

Fig. (15b) is handled similarly. Suppose  $p$  corresponds to the triangle whose boundary has marginals given by  $(-\alpha, -\beta, \gamma)$  and  $q$  corresponds to  $(\alpha, \beta, \gamma)$ . Then we have

$$\begin{aligned} p^{00} &= (-\alpha + \beta + \gamma)/2 & q^{00} &= (\alpha + \beta + \gamma)/2 = -p^{01} \\ p^{01} &= (-\alpha - \beta - \gamma)/2 & q^{01} &= (-\alpha + \beta - \gamma)/2 = -p^{10} \\ p^{10} &= (\alpha - \beta + \gamma)/2 & q^{10} &= (-\alpha - \beta + \gamma)/2 = -p^{11} \\ p^{11} &= (\alpha + \beta - \gamma)/2 & q^{11} &= (\alpha - \beta - \gamma)/2 = -p^{00}. \end{aligned}$$

This case gives either a type 1 or a type 2 loop. For example, setting  $p^{10} = q^{00} = 0$  gives the signed loop in Fig. (14b), and  $p(1/2)$  is the vertex  $p_b$ .

□

Next, we will describe the graph of  $MP_1$ . By Lemma 15 the action of  $G_1$  on the type 1 and 2 vertices is transitive. Therefore to understand the local structure, i.e., the neighbors, at a given vertex we can fix one type 1 vertex and one type 2 vertex. Our canonical representative for a type 1 vertex is  $q_0$  given in Fig. (16b), which as an operator given as follows:

$$q_0 = \frac{1}{4}(\mathbb{1} + X \otimes Y - Y \otimes Y + Z \otimes Y). \quad (34)$$

Here we are using Lemma 2 to identify points of  $MP_1$  as operators and don't distinguish them notationally from the probability distributions. For a type 2 vertex our canonical choice is  $p_0$  given in Fig. (16a):

$$p_0 = \frac{1}{4}(\mathbb{1} + X \otimes X + X \otimes Y + Y \otimes X - Y \otimes Y + Z \otimes Z). \quad (35)$$

**Lemma 15** *Let  $p$  be a vertex of  $MP_1$ .*



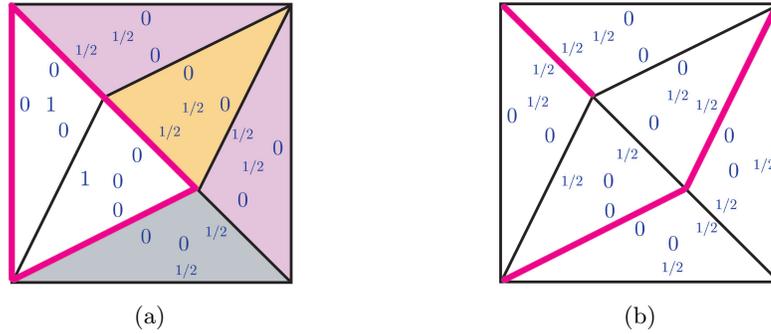


Fig. 16

- $N(q_0)$  consists of 12 type 2 vertices given in Fig. (19a).  $Stab_{G_1}(q_0)$  acts transitively on these neighbors.
- $N(p_0)$  consists of 8 type 1 vertices and 16 type 2 vertices given in Fig. (18a) and (18b); respectively.  $Stab_{G_1}(p_0)$  acts transitively on the type 1 neighbors, whereas the type 2 neighbors break into two orbits.

*Proof.* Vertices of  $MP_1$  are classified in part (ii) of Theorem 3. Lemma 14 shows that edges of  $MP_1$  are described by signed loops. To describe the local structure of the graph at a vertex we consider the canonical vertices in Eq. (34) and (35), since by Lemma 15  $G_1$  acts transitively on each type of vertex.

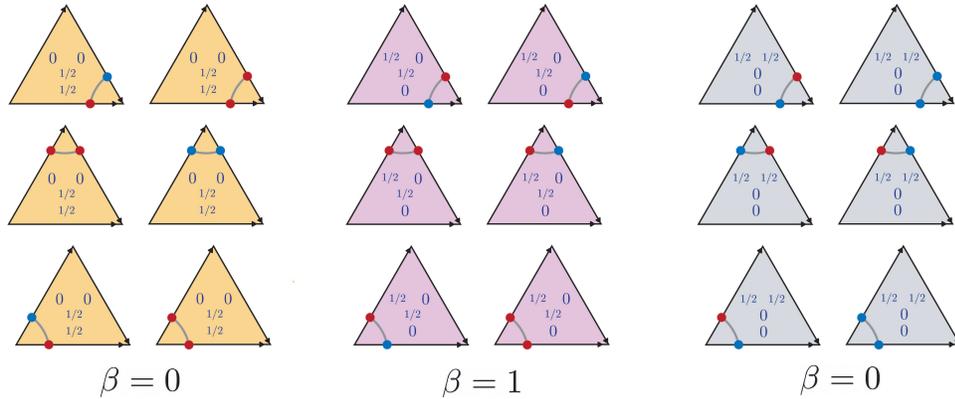
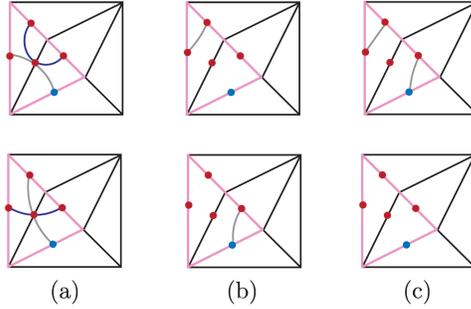


Fig. 17

Our strategy is to find the signed loops such that  $p(\alpha)$  in Eq. (33) gives  $p(0) = q_0$  or  $p_0$ . Let us start with  $p_0$ , the corresponding distribution is given in Fig(16a). Let  $\Omega$  denote the maximal cnc set corresponding to  $p_0$  and  $\Omega^c$  denote its complement. We will partition a loop  $l$  into two parts  $\Omega_l \cap \Omega$  and  $\Omega_l \cap \Omega^c$ . The restriction of  $\varphi$  to  $\Omega_l \cap \Omega$  is determined by the outcome assignment corresponding to  $p_0$ . We begin by considering the restriction of  $\varphi$  to  $\Omega_l \cap \Omega^c$ . The region  $\Omega^c$  consists of four triangles. Each of these triangles has exactly one

deterministic edge. Let  $C$  be one of those four triangles. The intersection  $\Omega_l \cap C$  is either empty or consists of two edges. There are two choices for the restriction of the sign  $\varphi$  to this intersection, which is dictated by the distribution on  $C$ . All the possibilities are given in Fig. (17). Observe that  $\Omega_l \cap \Omega$  can be given by one of the following possibilities:



We analyze each case.

- (a) There are two ways to complete the paths to a loop. The sign on  $\Omega_l \cap \Omega^c$  is determined by two adjacent triangles  $C, C'$ . There are two possibilities for the sign on  $(\Omega_l \cap \Omega^c) \cap (C \cup C')$ . If  $l$  is type 2 then we obtain the type 2 neighbors given in the first two columns of Fig. (18b). The action of  $\text{Stab}_{G_1}(p_0)$  is transitive by Lemma C.6 on these neighbors; see Table (C.2c). A representative vertex in this orbit is

$$p_b = \frac{1}{4}(\mathbb{1} + X \otimes X + X \otimes Y - Y \otimes Z + Z \otimes X - Z \otimes Y). \tag{36}$$

If  $l$  is type 1 we obtain the type 1 neighbors in Fig. (18a). By Lemma C.6  $\text{Stab}_{G_1}(p_0)$  acts transitively; see Table (C.2a). A representative vertex in this orbit is  $q_0$  given in Eq. (34).

- (b) This is similar to (a): Two ways to complete to a loop and two choices for the sign on the complement. We obtain the type 2 neighbors in the last two columns of Fig. (18b). The action of  $\text{Stab}_{G_1}(p_0)$  is transitive by Lemma C.6 on these neighbors; see Table (C.2c). A representative vertex in this orbit is

$$p_a = \frac{1}{4}(\mathbb{1} + X \otimes X - Y \otimes Y - Y \otimes Z - Z \otimes Y + Z \otimes Z). \tag{37}$$

- (c) Top figure: There are two ways to complete to a loop. The sign on the complement is determined by two nonadjacent triangles. Hence there are four possibilities for the sign on the complement. We obtain the signed loops in Fig. (19b). By Lemma C.7 the vertices at  $p(1/2)$  are not neighbors of  $p_0$ . Also in the proof of this lemma we see that  $\text{Stab}_{G_1}(p_0)$  acts transitively; see Table (C.2b). Our representative vertex is

$$q_b = \frac{1}{4}(\mathbb{1} + Z \otimes X - Z \otimes Y + Z \otimes Z). \tag{38}$$

Bottom figure: There is a unique loop on  $\Omega^c$ . However, no sign is compatible with the restrictions onto the triangles given in Fig. (17). This loop does not produce an edge in the graph that initiates from  $p_0$ ; see Lemma C.7.

The distributions connecting  $p_0$  to  $q_0$ ,  $p_b$  and  $p_a$  are given in Fig. (14a), (14b) and (14c); respectively.

For  $q_0$  given in Fig (16b) the argument is similar. Let  $p(\alpha)$  be a path obtained from a signed loop such that  $p(0) = q_0$ . The distribution  $p(1/2)$  will consist of triangles with a single deterministic edge on the boundary. Hence it is a vertex of type 1. However, we need to determine whether  $p(\alpha)$  is an edge in  $MP_1$ . There are three cases to consider.

- (a)  $l$  is of type 2: Then  $p(\alpha)$  will be obtained from  $q_0$  by swapping a 1/2 with 0 in each triangle. This means that the common set of zeros between  $q_0$  and  $p(1/2)$  is 6. Therefore  $p(\alpha)$  cannot be an edge.
- (b)  $l$  is of type 1 and intersects two of the edges in  $\{X \otimes Y, Z \otimes Y, Y \otimes Y\}$ : Similarly  $p(1/2)$  is a type 1 vertex. Looking at the common zeros we see that there are 8. However, as in the proof of Lemma C.7 we can argue that Lemma 8 to a pair of adjacent triangles to reduce the rank by 1. This implies that the path  $p(\alpha)$  is not an edge in  $MP_1$ .
- (b)  $l$  is of type 1 and intersects one of the edges in  $\{X \otimes Y, Z \otimes Y, Y \otimes Y\}$ : Then  $p(1/2)$  is a type 2 vertex as listed in Fig. (19a). By Lemma C.5  $\text{Stab}_{G_1}(q_0)$  acts transitively on them. The distribution connecting  $q_0$  to  $p_0$  is given in Fig. (14a).

□

## 5 Applications

Mermin polytopes  $MP_\beta$ , besides having an interesting structure in their own right, also have utility in understanding aspects of quantum foundations ( $MP_0$ ) as well as quantum computation ( $MP_1$ ). We explore these topics here.

### 5.1 A new topological proof of Fine's theorem

Here we combine the current results on Mermin polytopes together with the topological framework of [12] to provide a novel proof of Fine's theorem [14]. Before proceeding to the precise statement of Fine's theorem, however, we recall from Section 2.3 that the CHSH scenario consists of four measurements  $x_i, y_j$  and four measurement contexts consisting of pairs  $\{x_i, y_j\}$ , where  $i, j \in \mathbb{Z}_2$ . Our first goal will be to represent this scenario topologically.

In the simplicial approach to contextuality, first introduced in [12], and discussed briefly in Section 2.2, measurement contexts are represented by simplices (triangles) and to each simplex we associate a probability distribution. The collection of distributions on each simplex constitutes a *simplicial distribution*, which generalizes the notion of nonsignaling distributions. In particular, a well-studied class of measurement scenarios are the bipartite scenarios which in the simplicial framework are given by collections of triangles (i.e., 2-simplices) where edges (i.e., 1-simplices) represent measurements; not necessarily local. Nonsignaling (or compatibility) constraints are then formalized as the gluing of triangles along edges. For instance, the (2, 1, 2) Bell scenario is just a single triangle as in Fig. (20), while the so-called diamond scenario consists of two triangles glued along a single edge; see Fig. (23b). The diamond scenario will prove useful for our proof of Fine's theorem.

A topological representation of the CHSH scenario is given by four triangles glued along their  $x_i, y_j$  edges. We assemble these four triangles into a punctured torus as in Fig. (21).

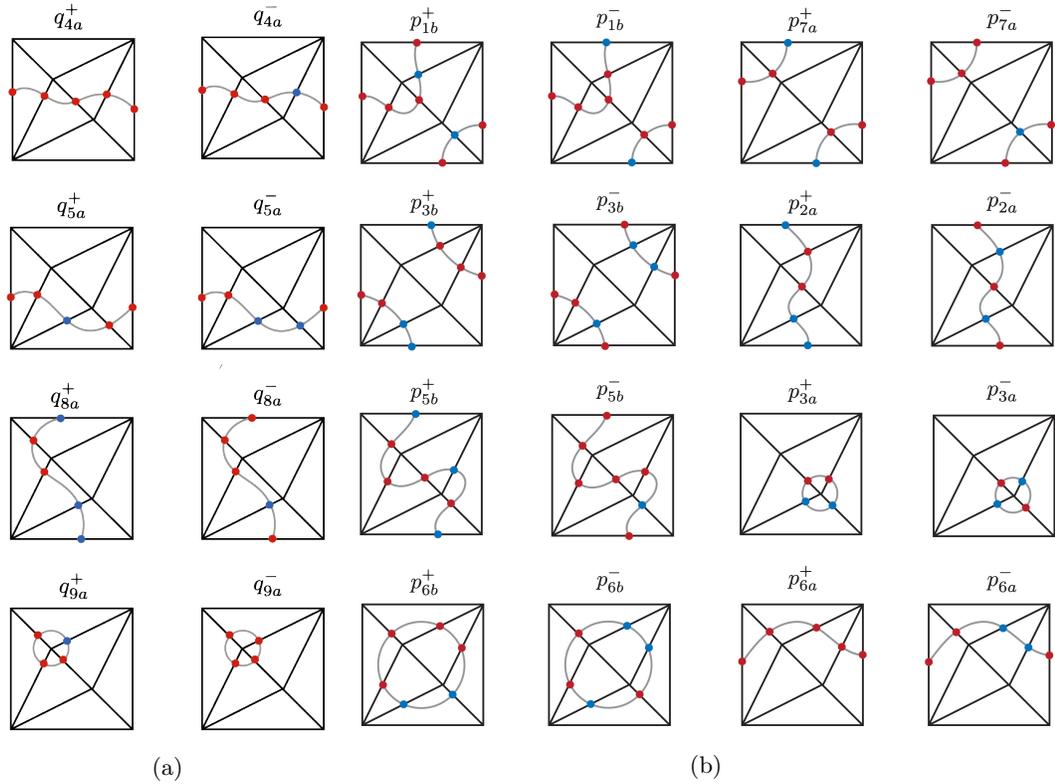


Fig. 18: (a) Type 1 neighbors of  $p_0$ . (b) Type 2 neighbors of  $p_0$ .

That is, as a Mermin scenario with the  $\{x_0 \oplus y_0, x_1 \oplus y_1, z\}$  and  $\{x_0 \oplus y_1, x_1 \oplus y_0, z\}$  contexts removed. For convenience we denote the CHSH scenario as  $T_0$  and the Mermin scenario as  $T$ .

Before we analyze this scenario, let us establish some terminology.

**Definition 3** [12, Def. 3.10] *A simplicial distribution  $p$  is called noncontextual if it can be written as a convex combination of deterministic distributions. Otherwise we call it contextual.*

This notion of contextuality specializes to the usual notion for the CHSH scenario. As is well-known, the CHSH scenario is contextual since there are distributions, the so-called Popescu-Rohrlich boxes [27], which cannot be written as a probabilistic mixture of deterministic distributions. It was established by CHSH [13] that necessary for a distribution on the CHSH

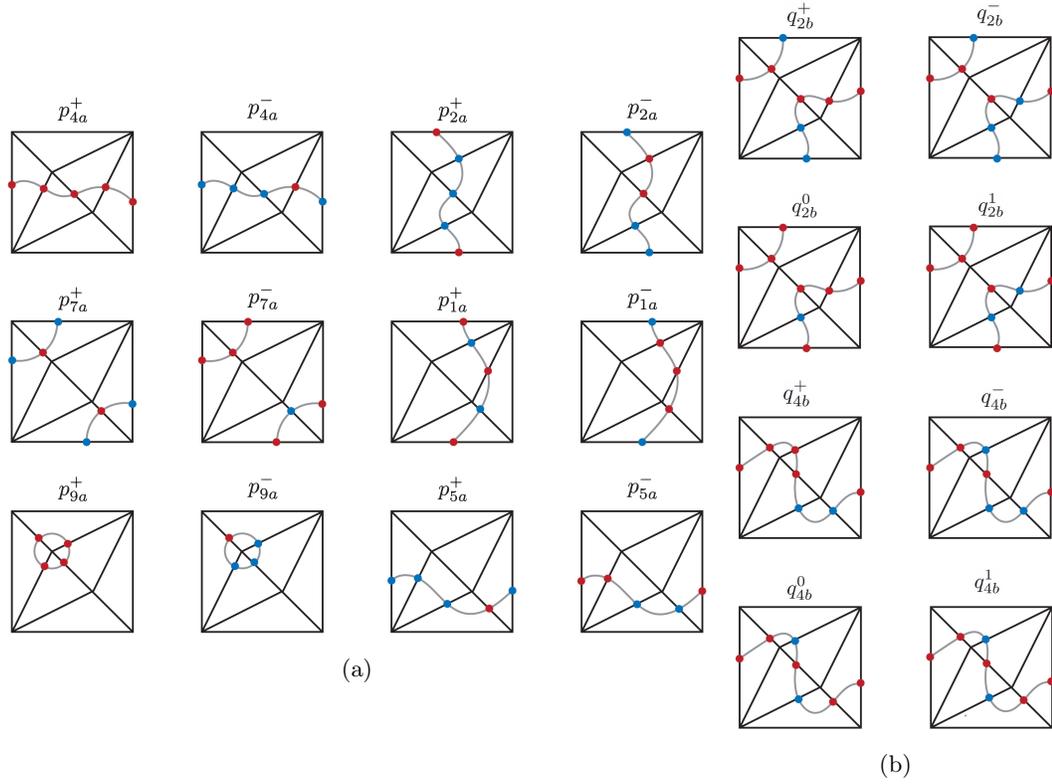


Fig. 19: (a) Type 2 neighbors of  $q_0$ . The neighbors  $p_{1a}^+, p_{1a}^-, p_{2a}^+, p_{7a}^+, p_{5a}^+, p_{9a}^-$  coincide with the type 2 neighbors  $p_{1b}^+, p_{1b}^-, p_{3b}^+, p_{3a}^-, p_0$  of the canonical type 2 vertex  $p_0$ ; respectively. (b) Signed loops that connect  $p_0$  to a type 1 vertex that is not a neighbor.

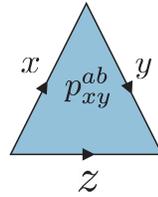


Fig. 20: The (2, 1, 2) Bell scenario in the simplicial setting.

scenario to be noncontextual is that the following CHSH inequalities be satisfied:

$$\begin{aligned}
 0 &\leq p_{x_0 \oplus y_0}^0 + p_{x_0 \oplus y_1}^0 + p_{x_1 \oplus y_0}^0 - p_{x_1 \oplus y_1}^0 \leq 2 \\
 0 &\leq p_{x_0 \oplus y_0}^0 + p_{x_0 \oplus y_1}^0 - p_{x_1 \oplus y_0}^0 + p_{x_1 \oplus y_1}^0 \leq 2 \\
 0 &\leq p_{x_0 \oplus y_0}^0 - p_{x_0 \oplus y_1}^0 + p_{x_1 \oplus y_0}^0 + p_{x_1 \oplus y_1}^0 \leq 2 \\
 0 &\leq -p_{x_0 \oplus y_0}^0 + p_{x_0 \oplus y_1}^0 + p_{x_1 \oplus y_0}^0 + p_{x_1 \oplus y_1}^0 \leq 2.
 \end{aligned}
 \tag{39}$$

Fine [14, 20] then established the sufficiency of these inequalities:

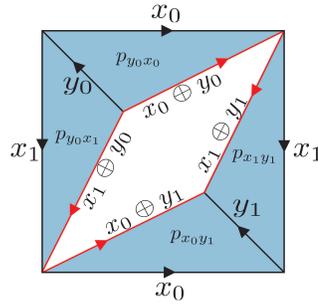


Fig. 21: CHSH scenario represented topologically as a punctured torus.

**Theorem 6** [Fine] *A distribution on the CHSH scenario is noncontextual if and only if the CHSH inequalities are satisfied.*

To provide a new proof of Fine’s theorem we will rely on a couple of key observations. One is that  $T_0$  can be embedded into  $T$  by inclusion, which allows us to study the CHSH scenario via the Mermin scenario. The other is the following immediate consequence of the vertex classification of  $MP_0$ .

**Corollary 3** *Any distribution on the Mermin torus, whose topological realization is given in Fig. (2a), is noncontextual.*

*Proof.* The distributions on the Mermin torus satisfying the nonsignaling conditions given in Eq. (7) constitute the polytope  $MP_0$ . In Theorem 3 part (1) we have seen that all the vertices of this polytope are deterministic. Therefore any distribution on the Mermin torus can be written as a probabilistic mixture of deterministic distributions.  $\square$

Next, we prove Proposition 2, which is stated in a more topological form below.

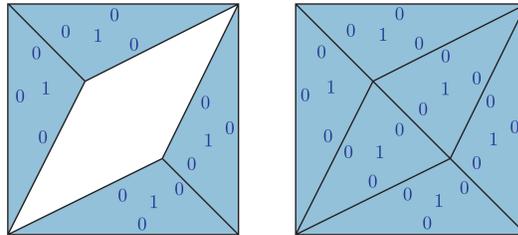


Fig. 22: (Left) The deterministic distribution  $\delta^s$  on the punctured torus corresponding to the outcome assignment  $s : x_0 \mapsto 1, x_1, y_j \mapsto 0$ . (Right) The extension  $\tilde{\delta}^s$  of the distribution to the torus.

**Proposition 5** *A distribution  $p$  on the punctured torus  $T_0$  extends to a distribution on the*

torus  $T$  if and only if  $p$  is noncontextual.

*Proof.* This result is a special case of the extension result proved in [12, Pro. 4.7]. The key observation is that an outcome assignment  $s : \{x_i, y_j : i, j, \in \mathbb{Z}_2\} \rightarrow \mathbb{Z}_2$  specifies both a deterministic distribution  $\delta^s$  on  $T_0$  and a deterministic distribution on  $T$ , which we denote by  $\tilde{\delta}^s$ . See Fig. (22). Assume that  $p$  is noncontextual. We can express  $p$  as a probabilistic mixture  $\sum_s \lambda(s) \delta^s$  of deterministic distributions. Then  $\tilde{p}$  defined as the probabilistic mixture  $\sum_s \lambda(s) \tilde{\delta}^s$  is the desired extension. Conversely, assume that  $p$  extends to a distribution  $\tilde{p}$  on the torus. By Corollary 3 every distribution in  $\text{MP}_0$  is noncontextual, i.e., can be expressed as a probabilistic mixture of deterministic distributions  $\tilde{\delta}^s$ . Then restricting onto  $T_0$  we can write  $p$  as a probabilistic mixture of  $\delta^s$ . Thus  $p$  is noncontextual.  $\square$

Since an extension from  $T_0$  to  $T$  amounts to filling in the diamond whose boundary is given by the measurements  $x_i \oplus y_j$ ,  $i, j \in \mathbb{Z}_2$ , it is useful to establish the following fact:

**Lemma 16** *A distribution  $p$  on the boundary of the diamond scenario extends to the diamond if and only if  $p$  satisfies the CHSH inequalities in Eq. (39).*

*Proof.* This is proved in [12, Pro 4.9], we include the proof here for the convenience of the reader. For our purposes we will assume that the diamond  $Z$  is such that the triangles are glued along their XOR edge; see Fig. (23).



Fig. 23: (a) The boundary of the diamond. (b) Topological representation of the diamond scenario.

The argument for the other choices is similar. The distribution  $p_{\partial Z}$  on the boundary of the diamond is specified by  $(p_{x_0}^0, p_{y_0}^0, q_{x_1}^0, q_{y_1}^0) \in [0, 1]^4$ . On the other hand, a distribution  $p_Z$  on the diamond, requires compatible distributions  $p_{xy}^{ab}$  and  $q_{vw}^{rs}$ , which by using Eq. (16) can be specified by  $(p_{x_0}^0, p_{y_0}^0, q_{x_1}^0, q_{y_1}^0, q_z^0)$ , where  $q_z^0$  is the marginal along the common edge. It is possible to extend from  $\partial Z$  to  $Z$  if and only if there exists a  $q_z^0$  such that all  $p_{x_0 y_0}^{ab}, q_{x_1 y_1}^{rs} \geq 0$ . This occurs precisely when

$$\max\{|p_{x_0}^0 + p_{y_0}^0 - 1|, |q_{x_1}^0 + q_{y_1}^0 - 1|\} \leq q_z^0 \leq \min\{1 - |p_{x_0}^0 - p_{y_0}^0|, 1 - |q_{x_1}^0 - q_{y_1}^0|\}.$$

By Fourier-Motzkin elimination this single inequality is equivalent to the following four

$$\begin{aligned} |p_{x_0}^0 + p_{y_0}^0 - 1| &\leq 1 - |p_{x_0}^0 - p_{y_0}^0| \\ |p_{x_0}^0 + p_{y_0}^0 - 1| &\leq 1 - |q_{x_1}^0 - q_{y_1}^0| \\ |q_{x_1}^0 + q_{y_1}^0 - 1| &\leq 1 - |p_{x_0}^0 - p_{y_0}^0| \\ |q_{x_1}^0 + q_{y_1}^0 - 1| &\leq 1 - |q_{x_1}^0 - q_{y_1}^0|, \end{aligned}$$

in addition to the trivial inequalities corresponding to  $0 \leq p_{i_0}^0, q_{i_1}^0 \leq 1$ , where  $i = x, y$ . Expanding the absolute values gives the inequalities

$$\begin{aligned} 0 &\leq p_{x_0}^0 + p_{y_0}^0 + q_{x_1}^0 - q_{y_1}^0 \leq 2 \\ 0 &\leq p_{x_0}^0 + p_{y_0}^0 - q_{x_1}^0 + q_{y_1}^0 \leq 2 \\ 0 &\leq p_{x_0}^0 - p_{y_0}^0 + q_{x_1}^0 + q_{y_1}^0 \leq 2 \\ 0 &\leq -p_{x_0}^0 + p_{y_0}^0 + q_{x_1}^0 + q_{y_1}^0 \leq 2. \end{aligned} \tag{40}$$

These equations are formally identical to the CHSH inequalities appearing in Eq.(39).  $\square$

**[Proof of Theorem 6]** Let  $p$  be a distribution on  $T_0$  and  $p_\partial$  denote the restriction (marginalization) of  $p$  to the boundary of  $T_0$ . Observe that the torus is obtained from the punctured torus by filling in the diamond in the middle. Therefore  $p$  extends to  $T$  if and only if  $p_\partial$  extends to the diamond. Combining this observation with Proposition 5 and Lemma 16 gives the desired result.  $\square$

### 5.2 Decomposing the 2-qubit $\Lambda$ -polytope

In this section, we provide a decomposition for  $\Lambda_2$ , the  $\Lambda$ -polytope for 2-qubits, using the Mermin polytope  $MP_1$ . This decomposition will provide valuable insight into the vertex enumeration problem for  $\Lambda$ -polytopes. This problem is a fundamental mathematical obstacle in the complexity analysis of the  $\Lambda$ -simulation algorithm introduced in [11].

Recall the set  $\mathcal{S}_2$  of 2-qubit stabilizer states and the (non)local version from Eq. (10). The 2-qubit  $\Lambda$ -polytope is defined as follows:

$$\Lambda_2 = \{X \in \text{Herm}((\mathcal{C}^2)^{\otimes 2}) : \text{Tr}(X) = 1, \text{Tr}(X\Pi) \geq 0, \forall \Pi \in \mathcal{S}_2\} \tag{41}$$

Our decomposition will be derived from the local vs. nonlocal decomposition of Pauli operators introduced in Section 2.4. Let us write  $E^{(l)}$  and  $E^{(nl)}$  for the subsets of  $E$  corresponding to the local and nonlocal Pauli operators; respectively. This gives us the following decomposition:

$$E = \{0\} \sqcup E^{(l)} \sqcup E^{(nl)}.$$

Let  $\mathcal{I}$  denote the set of maximal isotropic subspaces in  $E$ . This set also decomposes into a local  $\mathcal{I}^{(l)}$  and a nonlocal part  $\mathcal{I}^{(nl)}$ ; see Fig. (24). Recall that the Mermin scenario  $(M, \mathcal{C})$  can be identified with  $(E^{(nl)}, \mathcal{I}^{(nl)})$  via the map in Eq. (13). The function  $\beta_1 : \mathcal{C} \rightarrow \mathbb{Z}_2$  extends to a function  $\tilde{\beta} : \mathcal{I} \rightarrow \mathbb{Z}_2$  where  $\tilde{\beta}(C) = 0$  for  $C \in \mathcal{I}^{(l)}$ . We begin with a result that is a local version of Lemma 2. We define the local 2-qubit  $\Lambda$ -polytope:

$$\Lambda_2^{(l)} = \{X \in \text{Herm}((\mathcal{C}^2)^{\otimes 2}) : \text{Tr}(X) = 1, \text{Tr}(X\Pi) \geq 0, \forall \Pi \in \mathcal{S}_2^{(l)}\}.$$

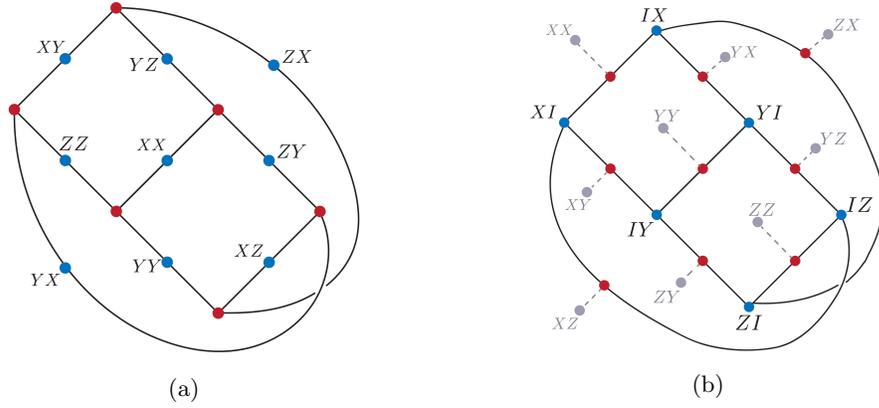


Fig. 24: (a) Subspaces in  $\mathcal{I}_2^{(nl)}$  and their intersections (zero subspace is omitted). (b) Subspaces in  $\mathcal{I}_2^{(l)}$  and their intersections (nonlocal operators do not belong to this set, they are only indicated to reveal the connection to the nonlocal part).

Note that by definition  $\Lambda_2 \subset \Lambda_2^{(l)}$ . This local polytope is, in fact, a well-known nonsignaling polytope. The  $(2, 3, 2)$  Bell scenario consists of

- the measurement set  $M_{232} = \{x_i, y_j : i, j \in \mathbb{Z}_3\}$ ,
- the collection  $\mathcal{C}_{232}$  of contexts  $C_{ij}$ , where  $i, j \in \mathbb{Z}_3$ , given by

$$C_{ij} = \{x_i, y_j\}.$$

**Lemma 17** *The local polytope  $\Lambda_2^{(l)}$  can be identified with the nonsignaling polytope  $NS_{232}$  of the  $(2, 3, 2)$  Bell scenario.*

*Proof.* The argument is similar to the proof of Lemma 2. An operator  $X \in \Lambda_2^{(l)}$  specifies a distribution  $p_X$  in  $NS_{\mathcal{I}^{(l)},0}$  (see Eq. (3) for the definition of the polytope) via the Born rule. For the bijection we need an inverse map, which comes from by first marginalizing to a single measurement and then computing the expectation  $\langle A \rangle_X$  of the corresponding Pauli operator. The identification of  $NS_{\mathcal{I}^{(l)},0}$  with the nonsignaling polytope  $NS_{232}$  follows from realizing the measurements in the Bell scenario as quantum mechanical measurements

$$\begin{aligned} x_0 &\mapsto X \otimes \mathbb{1}, & x_1 &\mapsto Y \otimes \mathbb{1}, & x_2 &\mapsto Z \otimes \mathbb{1} \\ y_0 &\mapsto \mathbb{1} \otimes X, & y_1 &\mapsto \mathbb{1} \otimes Y, & y_2 &\mapsto \mathbb{1} \otimes Z. \end{aligned} \tag{42}$$

□

Let  $MP_1^{\mathbb{R}}$  denote the Mermin polytope for quasiprobability distributions; see Definition 1. We introduce an important map

$$\text{ext} : NS_{232} \rightarrow MP_1^{\mathbb{R}} \tag{43}$$

using the identification of Lemma 17. For the explicit description of the ext map we need to extend the  $(2, 3, 2)$  Bell scenario  $(M_{232}, \mathcal{C}_{232})$  by including the nonlocal measurements. Formally we introduce an extended scenario:

- $\tilde{M} = M_{232} \sqcup \{x_i \oplus y_j : i, j \in \mathbb{Z}_3\}$ ,
- $\tilde{\mathcal{C}} = \{\tilde{C}_{ij} : i, j \in \mathbb{Z}_3\}$  where  $\tilde{C}_{ij} = C_{ij} \sqcup \{x_i \oplus y_j\}$ .

Now we are ready to describe the ext map explicitly. Let  $d = \{d_{C_{ij}}\}_{i,j \in \mathbb{Z}_3}$  be a nonsignaling distribution defined on the  $(2, 3, 2)$  Bell scenario. We define  $\tilde{d}$  as a nonsignaling distribution on  $(\tilde{M}, \tilde{\mathcal{C}})$  by setting

$$\tilde{d}_{\tilde{C}_{ij}}(s) = \begin{cases} d_{C_{ij}}(s|_{C_{ij}}) & s(x_i \oplus y_j) = s(x_i) + s(y_j) \\ 0 & \text{otherwise.} \end{cases}$$

The mapping in Eq. (42) can be used to define an embedding  $M_{232} \subset E$ . Together with the embedding of Eq. (13) we obtain a local vs nonlocal decomposition

$$E = M_{232} \sqcup M.$$

With this convention we will give the explicit form of the ext map on a context of the form

$$C = \{(v, w), (v', w'), (v + v', w + w')\}$$

with  $\omega((v, w), (v', w')) = 0$ . For  $s \in \mathbb{Z}_2^C$  we set  $a = s(v, w)$ ,  $b = s(v', w')$  and  $c = a + b + \beta((v, w), (v', w'))$ . Then we have

$$(\text{ext } d)_C(s) = \frac{1}{2} \left( \tilde{d}|_{\{(v,w)\}}(a) + \tilde{d}|_{\{(v',w')\}}(b) - \tilde{d}|_{\{(v+w,v'+w')\}}(c+1) \right) \quad (44)$$

if  $s(v + w, v' + w') = s(v, w) + s(v', w') + \beta((v, w), (v', w'))$  and  $(\text{ext } d)_C(s) = 0$  otherwise.

**Theorem 7** *The polytope  $\Lambda_2$  is precisely the subpolytope of the nonsignaling polytope  $NS_{232}$  for the  $(2, 3, 2)$  Bell scenario given by those distributions that map to a probability distribution in  $MP_1$  under the ext map given in Eq. (44).*

*Proof.* Using the identification given in Lemma 17 and the operator-theoretic description of  $MP_1^{\mathbb{R}}$  in Lemma 2 the ext map given in Eq. (43) can be identified with the map

$$\Lambda_2^{(1)} \rightarrow MP_1^{\mathbb{R}} \quad (45)$$

obtained by sending  $X$  to the operator  $\bar{X}$  such that  $\langle \bar{X} \rangle_A = \langle X \rangle_A$  for nonlocal Pauli operators  $A$  (including  $\mathbb{1}$ ) and  $\langle \bar{X} \rangle_A = 0$  for the remaining local Pauli operators. Those operators  $X \in \Lambda_2^{(1)}$  which give a probability distribution on the Mermin scenario, instead of a quasiprobability distribution, are precisely those that come from  $\Lambda_2$ .  $\square$

Theorem 7 gives a description of  $\Lambda_2$  in terms of well-understood polytopes:  $NS_{232}$  whose vertices are described in [17] and  $MP_1$  described in Theorem 3. We remark that not every vertex of  $\Lambda_2$  is a vertex of  $NS_{232}$ . Description of the vertices of the former requires the analysis of the ext map in Eq. (44). We leave this analysis for a future work.

We hope that the decomposition described in Theorem 7 will help understand  $\Lambda_n$  for  $n \geq 2$ . The local version  $\Lambda_n^{(l)}$  can be defined for any number of qubits. The  $\Lambda$ -polytope lies inside this local version  $\Lambda_n \subset \Lambda_n^{(l)}$ . One can consider an analogue of the ext map and try to describe  $\Lambda_n$  as a subpolytope of the local version. Since the structure of the non-local Pauli operators and the corresponding contexts are more complicated the analysis will be harder. For  $n \geq 3$  the local version cannot be identified with  $\text{NS}_{n32}$ , the polytope of non-signaling distributions on the  $(n, 3, 2)$  Bell scenario.

## 6 Conclusion

Motivated by a classic example of contextuality known as Mermin's square [7], and its topological realization given in [9], in this paper we considered variations of this scenario parametrized by a function  $\beta$  and studied the corresponding nonsignaling polytopes  $\text{MP}_\beta$ . We showed that these polytopes fall into two equivalence classes, determined by  $[\beta]$ , which has a cohomological interpretation [9]. Among our main results is the characterization of the vertices of  $\text{MP}_\beta$ . We demonstrated that all vertices of  $\text{MP}_0$  are deterministic, which facilitates a novel proof of Fine's theorem [14, 20]. On the other hand,  $\text{MP}_1$  has two types of vertices, both of which are cnc [16]. We also described the graphs associated with the polytopes. In the case of  $\text{MP}_1$ , the edges in this graph are essentially given by the loops on the Mermin torus. These loops correspond to complements of cnc sets and play a significant role throughout the paper.

An important connection is established between  $\text{MP}_1$  and computation through the notion of  $\Lambda$ -simulation [11]. Indeed, if one restricts to just measurements of non-local Pauli operators then one can define a simulation algorithm for  $\text{MP}_1$  in the spirit of [11]; although, since the vertices of  $\text{MP}_1$  are cnc, all resulting quantum computations can be efficiently simulated classically [16]. Alternatively, here we have established that  $\text{MP}_1$  corresponds precisely to the non-local part of the polytope  $\Lambda_2$  [11, 15], with the local part being related to  $\text{NS}_{232}$ , the polytope associated with the  $(2, 3, 2)$  Bell scenario [17]. We expect that this decomposition will be important in understanding the combinatorial structure of  $\Lambda_2$ ; an important first step in analyzing the complexity of classical simulation based on  $\Lambda$ -polytopes, which become numerically intractable even for  $n = 3$  qubits.

An interesting but yet unexplored topic of research is the study more generally of polytopes associated with measurement scenarios, or topological spaces with non-trivial  $[\beta]$ . An interesting example of this is the well-known Mermin star scenario [7], which also has a topological realization as a torus [9]. Particularly appealing about this line of research is that the Mermin's star is closely related to the so-called Greenberger-Horne-Zeilinger (GHZ) paradox [28], which can be exploited for computational advantage; see [29, 30].

## Acknowledgments

This work is supported by the Air Force Office of Scientific Research under award number FA9550-21-1-0002 and the Horizon Europe project FoQaCiA, GA no. 101070558.

## References

pt!

1. D. Bohm, "A suggested interpretation of the quantum theory in terms of hidden variables I", *Physical Review*, vol. 85, no. 2, p. 166, 1952. doi: 10.1103/PhysRev.85.166.

2. D. Schmid, J. H. Selby, and R. W. Spekkens, “Unscrambling the omelette of causation and inference: The framework of causal-inferential theories,” arXiv preprint, 2020. arXiv: 2009.03297.
3. A. Caticha, “Entropic dynamics and quantum ‘measurement’” *Physical Sciences Forum*, vol. 5, p. 36, Multidisciplinary Digital Publishing Institute, 2022. doi: 10.3390/psf2022005036. arXiv: 2208.02156.
4. J. S. Bell, On the problem of hidden variables in quantum mechanics, *Reviews of Modern physics*, vol. 38, no. 3, p. 447, 1966. doi: 10.1103/RevModPhys.38.447.
5. S. Kochen and E. P. Specker, “The problem of hidden variables in quantum mechanics”, in *The logic algebraic approach to quantum mechanics*, pp. 293328, Springer, 1975. doi: 10.1007/978-94-010-1795- 4 17.
6. C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, and J.Å. Larsson, “Kochen-Specker contextuality”, *Reviews of Modern Physics*, vol. 94, no. 4, p. 045007, 2022. doi: 10.1103/RevModPhys.94.045007. arXiv:2102.13036.
7. N. D. Mermin, “Hidden variables and the two theorems of John Bell”, *Reviews of Modern Physics*, vol. 65, no. 3, p. 803, 1993. doi: 10.1103/RevModPhys.65.803. 34
8. R. Cleve and R. Mittal, “Characterization of binary constraint system games,” in *International Colloquium on Automata, Languages, and Programming*, pp. 320331, Springer, 2014. doi: 10.1007/978-3-662-43948-7 27. arXiv: 1209.2729.
9. C. Okay, S. Roberts, S. D. Bartlett, and R. Raussendorf, “Topological proofs of contextuality in quantum mechanics”, *Quantum Information & Computation*, vol. 17, no. 13-14, pp. 11351166, 2017. doi: 10.26421/QIC17.13-14-5. arXiv: 1701.01888.
10. E. Gawrilow and M. Joswig, “Polymake: a framework for analyzing convex polytopes”, *Polytopes Combinatorics and Computation*, p. 4373, 2000. doi: 10.1007/978-3-0348-8438-9 2.
11. M. Zurel, C. Okay, and R. Raussendorf, “Hidden variable model for universal quantum computation with magic states on qubits”, *Physical Review Letters*, vol. 125, no. 26, p. 260404, 2020. doi: 10.1103/PhysRevLett.125.260404. arXiv: 2004.01992.
12. C. Okay, A. Kharoof, and S. Ipek, “Simplicial quantum contextuality”, *Quantum*, vol. 7, p. 1009, 2023. doi: 10.22331/q-2023-05-22-1009. arXiv: 2204.06648.
13. J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, “Proposed experiment to test local hidden variable theories”, *Physical review letters*, vol. 23, no. 15, p. 880, 1969. doi: 10.1103/PhysRevLett.23.880.
14. A. Fine, “Hidden variables, joint probability, and the Bell inequalities”, *Physical Review Letters*, vol. 48, no. 5, p. 291, 1982. doi: 10.1103/PhysRevLett.48.291.
15. C. Okay, M. Zurel, and R. Raussendorf, “On the extremal points of the  $\Lambda$ -polytopes and classical simulation of quantum computation with magic states”, arXiv preprint, 2021. arXiv: 2104.05822.
16. R. Raussendorf, J. Bermejo-Vega, E. Tyhurst, C. Okay, and M. Zurel, Phase space simulation method for quantum computation with magic states on qubits”, *Physical Review A*, vol. 101, no. 1, p. 012350, 2020. doi: 10.1103/PhysRevA.101.012350. arXiv: 1905.05374.
17. N. Jones and L. Masanes, “Interconversion of nonlocal correlations”, *Physical Review A*, p. 4373, 2005. doi: 10.1007/978-3-0348-8438-9 2. arXiv: quant-ph/0506182.
18. Harry Buhrman, Richard Cleve, Monique Laurent, Noah Linden, Alexander Schrijver, and Falk Unger, “New limits on fault-tolerant quantum computation”, *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS’06)*, p. 411, 2006, IEEE.
19. Mark Howard and Jiri Vala, “Nonlocality as a benchmark for universal quantum computation in Ising anyon topological quantum computers”, *Physical Review A*, vol. 85, no. 2, p. 022304, 2012. doi: 10.1103/PhysRevA.85.022304.
20. A. Fine, “Joint distributions, quantum correlations, and commuting observables”, *Journal of Mathematical Physics*, vol. 23, no. 7, pp. 13061310, 1982. doi: 10.1063/1.525514.
21. S. Abramsky and A. Brandenburger, “The sheaf-theoretic structure of non-locality and contextuality”, *New Journal of Physics*, vol. 13, no. 11, p. 113036, 2011. doi: 10.1088/1367-2630/13/11/113036. arXiv: 1102.0264.
22. A. Kharoof, S. Ipek, and C. Okay, “Topological methods for studying contextuality: N-cycle

- scenarios and beyond”, arXiv preprint, 2023. arXiv: 2306.01459.
23. K. Sreekumar and K. Manilal, “Automorphism groups of some families of bipartite graphs”, *Electron. J. Graph Theory Appl.*, vol. 9, no. 1, pp. 6575, 2021. doi: 10.5614/ejgta.2021.9.1.6.
  24. G. M. Ziegler, *Lectures on polytopes*, vol. 152. Springer Science & Business Media, 2012. doi: 10.1007/978-1-4613-8431-1.
  25. V. Chvatal, *Linear programming*, A Series of Books in the Mathematical Sciences, W. H. Freeman and Company, New York, 1983.
  26. C. Godsil and G. F. Royle, *Algebraic graph theory*, vol. 207. Springer Science & Business Media, 2001. doi: 10.1007/978-1-4613-0163-9.
  27. S. Popescu and D. Rohrlich, “Quantum nonlocality as an axiom”, *Foundations of Physics*, vol. 24, no. 3, pp. 379385, 1994. doi: 10.1007/BF02058098. 35
  28. D. M. Greenberger, M. A. Horne, and A. Zeilinger, “Going beyond Bells theorem”, *Bells theorem, quantum theory and conceptions of the universe*, pp. 6972, Springer, 1989. doi: 10.1007/978-94-017-0849-4 10.
  29. J. Anders and D. E. Browne, “Computational power of correlations”, *Physical Review Letters*, vol. 102, no. 5, p. 050502, 2009. doi: 10.1103/PhysRevLett.102.050502. arXiv: 0805.1002.
  30. R. Raussendorf, “Contextuality in measurement-based quantum computation”, *Physical Review A*, vol. 88, no. 2, p. 022322, 2013. doi: 10.1103/PhysRevA.88.022322. arXiv: 0907.5449.
  31. J. Bray, M. Conder, C. Leedham-Green, and E. O'Brien, “Short presentations for alternating and symmetric groups”, *Transactions of the American Mathematical Society*, vol. 363, no. 6, pp. 32773285, 2011. doi: 10.1090/S0002-9947-2011-05231-1.

## Appendix A Proof of Proposition 1

In this section we will prove Proposition 1. For this we will introduce a generalized version of the Mermin polytope (Definition 1). Recall the  $K_{3,3}$  graph associated to the Mermin scenario with vertex set  $\mathcal{C} = \mathcal{C}^h \sqcup \mathcal{C}^v$  and edge set  $M$ ; see Fig. (6). We begin with generalizing the definition of  $\beta$ . Let  $R$  denote the set of pairs  $(C, m) \in \mathcal{C} \times M$  such that  $m \in C$ . We will consider incidence weights on  $K_{3,3}$ , that is functions of the form  $\beta : R \rightarrow \mathbb{Z}_2$ . Let us write  $K_{3,3}^\beta$  to indicate the weight.

**Definition A.1** Let  $\widetilde{\text{MP}}_\beta$  denote the polytope given by the set of functions

$$p : R \rightarrow \mathbb{R}_{\geq 0}$$

satisfying the following conditions:

- (a)  $\sum_{m \in C} p(C, m) \leq 1$  for all  $C \in \mathcal{C}$ ,
- (b) For a context  $C \in \mathcal{C}$  define  $p_C : C \rightarrow \mathbb{R}_{\geq 0}$  by

$$p_C(m) = \begin{cases} \sum_{m' \in C - \{m\}} p(C, m') & \beta(C) = 1, \\ 1 - \sum_{m' \in C - \{m\}} p(C, m') & \beta(C) = 0. \end{cases}$$

Then for all  $m \in M$  and  $C, C' \in \mathcal{C}$  such that  $m \in C \cap C'$  we require that

$$p_C(m) = p_{C'}(m).$$

To have a better idea of this definition consider a context  $C = \{x, y, z\}$  and let

$$a = p(C, x), \quad b = p(C, y), \quad c = p(C, z), \quad d = 1 - (a + b + c). \quad (\text{A.1})$$

Condition (a) says that  $p = \{a, b, c, d\}$  is a probability distribution. The choice of  $\beta$  at  $C$  determines the way  $p$  marginalizes to each single measurement. This is given by condition (b). For example, if  $\beta(C, x) = 1$  then we have  $p_x^0 = b + c$ , but if  $\beta(C, x) = 0$  then  $p_x^0 = a + d = 1 - (b + c)$ . In the notation of (b) the value  $p_C(x)$  coincides with  $p_x^0$ ; similarly for  $y$  and  $z$ . This definition generalizes  $\text{MP}_\beta$ ; for example the  $\beta$  choices given in Fig. (3) can be captured by the weights given in Fig. (A.1).

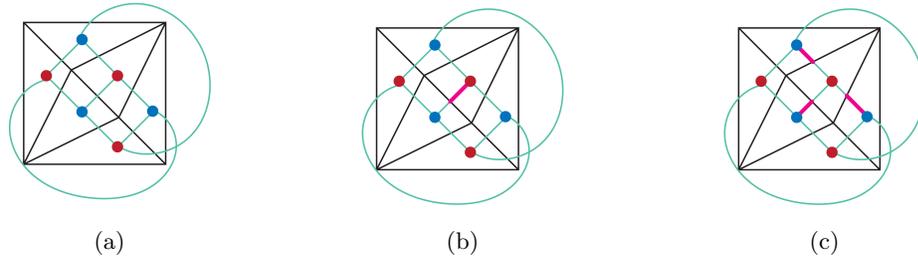
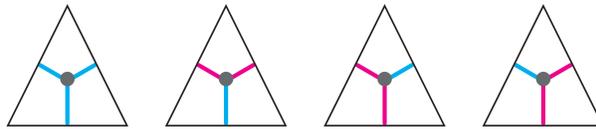


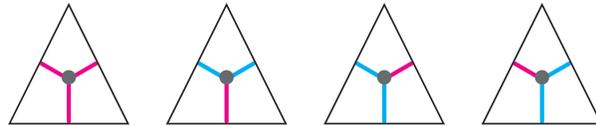
Fig. A.1:  $K_{3,3}^\beta$  for three different choices of  $\beta$ . Pink color on the part of an edge  $x$  incident to  $C$  implies that  $\beta(C, x) = 1$ , otherwise  $\beta$  takes the value of zero.

**Lemma A.1** Let  $\beta'$  be an incidence weight on  $K_{3,3}$  defined in the same way as  $\beta$  except possible at a single context as in (1) and (2), or at two contexts as in (3).

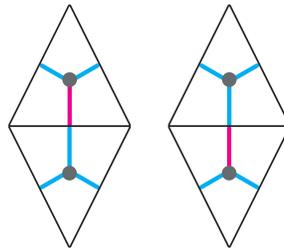
(1) There is a single context on which  $\beta$  and  $\beta'$  are defined as one of the following:



(2) There is a single context on which  $\beta$  and  $\beta'$  are defined as one of the following:



(3) There are two contexts on which  $\beta$  and  $\beta'$  are defined as one of the following:



At each case  $MP_\beta$  is combinatorially isomorphic to  $MP_{\beta'}$ .

*Proof.* The isomorphism is given by permuting the probability coordinates inside the contexts. Let  $C = \{x, y, z\}$  denote the context in cases (1) and (2). (See Fig. (20) for the labeling convention.) In case (1) we can obtain the isomorphism between the polytope corresponding to the first context and the next ones (from right to left) by flipping the outcome of  $x, y$  and both  $x$  and  $y$ ; respectively. Case (2) is similar. In (3) the constrained imposed at the common edge is the same in both cases, hence they specify the same polytope.  $\square$

Proposition 1 is a consequence of the following more general result.

**Proposition 6**  $\widetilde{MP}_\beta$  is combinatorially isomorphic to  $\widetilde{MP}_{\beta'}$  if and only if

$$\sum_{(C,m) \in R} \beta(C,m) = \sum_{(C,m) \in R} \beta'(C,m) \pmod{2}.$$

*Proof.* The main idea of the proof is to use Lemma A.1 to show that every case is either isomorphic to Fig. (A.1a) or Fig. (A.1b). First we observe that applying the transformations in (1) and (2) in Lemma A.1 we can assume that at each context  $C$  either  $\beta(C,m) = 0$  for all  $m \in C$  or  $\beta(C,m) = 1$  for exactly one  $m \in C$  and zero otherwise. Applying the transformation (3) we can assume that every context where  $\beta$  is nonzero is adjacent. Furthermore, again using (3) we can cancel a pair of adjacent contexts with  $\beta = 1$  by first rotating  $\beta(C,m) = 1$  once using (1) and then applying (3) to obtain a pair of contexts where on one of them  $\beta = 0$  and on the other there are two measurements for which  $\beta(C,m) = 1$ . Using (1) the remaining context with two nonzero  $\beta$ 's can be replaced by  $\beta = 0$ . This procedure terminates either at Fig. (A.1a), or, after successive application of the transformation in (3), at Fig. (A.1b).  $\square$

### Appendix B Proof of Proposition 3

Let  $\Sigma_n$  denote the symmetric group on  $n$  letters.

**Proposition 7** The group presentations of  $Cl_1$  and  $G_1$  are given as follows:

$$Cl_1 = \langle H, S \rangle \cong \langle h, s : h^2 = s^4 = (hs)^3 = 1 \rangle \cong \Sigma_4,$$

where  $s$  and  $h$  correspond to  $S$  and  $H$ , and

$$G_1 = \langle Cl_1 \times Cl_1, SWAP \rangle \cong \langle h, s, w : w^2 = h^2 = s^4 = (hs)^3 = [h, whw] = [h, wsw] = [s, whw] = [s, wsw] = 1 \rangle, \quad (B.1)$$

where  $s$ ,  $h$  and  $w$  correspond to  $S \otimes \mathbb{1}$ ,  $H \otimes \mathbb{1}$  and  $SWAP$ ; respectively.

*Proof.* The group  $\langle h, s : h^2 = s^4 = (hs)^3 = 1 \rangle$  is isomorphic to  $\Sigma_4$  (cf. [31, Theorem 8.1]). In particular, it has order 24. Sending  $H \mapsto h$  and  $S \mapsto s$  defines a group homomorphism from  $Cl_1$  since  $H^2 = S^4 = (HS)^3 = \mathbb{1}$ . By comparing the orders of the groups we see that this is an isomorphism.

By the first part, the presentation of  $Cl_1 \times Cl_1$  is given by

$$\begin{aligned} Cl_1 \times Cl_1 &\cong \langle h_1, s_1, h_2, s_2 : h_1^2 = s_1^4 = (h_1 s_1)^3 = h_2^2 = s_2^4 = (h_2 s_2)^3 \\ &= [h_1, h_2] = [h_1, s_2] = [s_1, h_2] = [s_1, s_2] = 1 \rangle \end{aligned}$$

where we identify  $h_1, s_1, h_2, s_2$  with  $H \otimes \mathbb{1}, S \otimes \mathbb{1}, \mathbb{1} \otimes H, \mathbb{1} \otimes S$ ; respectively. The presentation of  $G_1$  is obtained by adding one more generator namely  $w$  (which identify with  $SWAP$ ) and add relations that correspond to the action of  $w$  on  $h_1, s_1, h_2$  and  $s_2$ . Thus we have

$$\begin{aligned} G_1 &\cong \langle h_1, s_1, h_2, s_2, w : w^2 = h_1^2 = s_1^4 = (h_1 s_1)^3 = h_2^2 = s_2^4 = (h_2 s_2)^3 = 1 \\ &[h_1, h_2] = [h_1, s_2] = [s_1, h_2] = [s_1, s_2] = wh_1 wh_2^{-1} = ws_1 ws_2^{-1} = 1 \rangle. \end{aligned}$$

By relations  $wh_1 wh_2^{-1} = ws_1 ws_2^{-1} = 1$ , we can remove the generators  $h_2$  and  $s_2$ . Then we obtain

$$\begin{aligned} G_1 &\cong \langle h_1, s_1, w : w^2 = h_1^2 = s_1^4 = (h_1 s_1)^3 = (wh_1 w)^2 = (ws_1 w)^4 = (wh_1 wws_1 w)^3 = 1 \\ &[h_1, wh_1 w] = [h_1, ws_1 w] = [s_1, wh_1 w] = [s_1, ws_1 w] = 1 \rangle. \end{aligned}$$

Note that the relations  $(wh_1 w)^2 = 1$ ,  $(ws_1 w)^4 = 1$  and  $(wh_1 wws_1 w)^3 = 1$  can be obtained from  $h_1^2 = w^2 = 1$ ,  $w^2 = s_1^4 = 1$  and  $w^2 = (h_1 s_1)^3 = 1$ , respectively. Thus those three relations can be removed. Finally, we obtain

$$G_1 = \langle Cl_1 \times Cl_1, SWAP \rangle \cong \langle h, s, w : w^2 = h^2 = s^4 = (hs)^3 = [h, whw] = [h, wsw] = [s, whw] = [s, wsw] = 1 \rangle$$

where we identify  $h, s, w$  with  $H \otimes \mathbb{1}, S \otimes \mathbb{1}$  and SWAP, respectively.  $\square$

[Proof of Proposition 3] We will construct a function  $\phi : G_1 \rightarrow G_0$ , show that it is a group homomorphism, and makes the following diagram commute:

$$\begin{array}{ccccccc}
 0 & \longrightarrow & \mathbb{Z}^4 & \xrightarrow{\iota} & G_1 & \xrightarrow{p} & p(G_1) \subset Sp_4(\mathbb{Z}_2) \longrightarrow 1 \\
 & & \downarrow f & & \downarrow \phi & & \downarrow g \\
 0 & \longrightarrow & G_l \cong \mathbb{Z}_2^4 & \xrightarrow{\iota'} & G_0 & \xrightarrow{p'} & \text{Aut}(K_{3,3}) \longrightarrow 1
 \end{array}$$

Since  $G_1$  is a subgroup of  $Cl_2$ , the top row of the group extension corresponds to decomposing  $G_1$  into the symplectic part and the Pauli part. Define the following sets:

$$\begin{aligned}
 C_1 &= \{Y \otimes X, X \otimes Y, Z \otimes Z\}, \\
 C_2 &= \{X \otimes X, Y \otimes Y, Z \otimes Z\}, \\
 C_3 &= \{X \otimes Z, Z \otimes X, Y \otimes Y\}, \\
 C_4 &= \{X \otimes Y, Y \otimes Z, Z \otimes X\}, \\
 C_5 &= \{X \otimes X, Y \otimes Z, Z \otimes Y\}, \\
 C_6 &= \{X \otimes Z, Z \otimes Y, Y \otimes X\}.
 \end{aligned}$$

The group  $\text{Aut}(K_{3,3})$  permutes these sets, hence we think of it as a subgroup of  $\Sigma_6$ . We define  $f$  and  $\phi$  as follows

$$\begin{array}{lcl}
 \mathbb{1} \otimes X \mapsto l_{6b} & & H \otimes \mathbb{1} \mapsto (l_{5b}, (16)(23)(45)) \\
 f : \mathbb{1} \otimes Z \mapsto l_{2b} & \text{and } \phi : & S \otimes \mathbb{1} \mapsto (l_{3b}, (12)(34)(56)) \\
 X \otimes \mathbb{1} \mapsto l_{3b} & & \text{SWAP} \mapsto (l_0, (46)) \\
 Z \otimes \mathbb{1} \mapsto l_{4b} & &
 \end{array}$$

where we write  $l_0$  for the trivial element of  $G_l$ . Note that  $\phi$  factors through  $g$  and its surjective. It is clear that  $f$  is an isomorphism. It remains to show that  $\phi$  is group homomorphism and the left square of the diagram commutes. By Proposition 7, we know that the group presentation of  $G_1$  is given by Eq. (B.1). We show that  $\phi$  is a group homomorphism by checking that it respects all the relations.

We will need the products  $\phi(w)\phi(h)\phi(w)$  and  $\phi(w)\phi(s)\phi(w)$ :

$$\begin{aligned}
 \phi(w)\phi(h)\phi(w) &= (l_0, (46))(l_{5b}, (16)(23)(45))(l_0, (46)) \\
 &= (l_{1b}, (1456)(23))(l_0, (46)) \\
 &= (l_{1b}, (14)(23)(56))
 \end{aligned}$$

$$\begin{aligned}
 \phi(w)\phi(s)\phi(w) &= ((l_0, (46))(l_{3b}, (12)(34)(56))(l_0, (46)) \\
 &= (l_{6b}, (12)(3654))(l_0, (46)) \\
 &= (l_{6b}, (12)(36)(45))
 \end{aligned}$$

We check the commutation relation  $[\phi(h), \phi(w)\phi(h)\phi(w)] = 1$ :

$$\begin{aligned}\phi(h)(\phi(w)\phi(h)\phi(w)) &= (l_{5b}, (16)(23)(45))(l_{1b}, (14)(23)(56)) \\ &= (l_{5b} + l_{1b}, (15)(46)) \\ &= (l_{6a}, (15)(46))\end{aligned}$$

$$\begin{aligned}(\phi(w)\phi(h)\phi(w))\phi(h) &= (l_{1b}, (14)(23)(56))(l_{5b}, (16)(23)(45)) \\ &= (l_{1b} + l_{5b}, (15)(46)) \\ &= (l_{6a}, (15)(46)).\end{aligned}$$

The remaining commutation relations  $[\phi(h), \phi(w)\phi(s)\phi(w)] = [\phi(s), \phi(w)\phi(h)\phi(w)] = [\phi(s), \phi(w)\phi(s)\phi(w)] = 1$  can be checked similarly. Next, we check the remaining relations:

$$\begin{aligned}(\phi(h)\phi(s))^3 &= ((l_{5b}, (16)(23)(45))(l_{3b}, (12)(34)(56)))^3 \\ &= (l_{5b} + l_{4b}, (16)(23)(45)(12)(34)(56)) \\ &= (l_{3b}, (135)(264))(l_{3b}, (135)(264))^2 \\ &= (l_{3b} + l_{5b}, (153)(246))(l_{3b}, (135)(264)) \\ &= (l_{4b}, (153)(246))(l_{3b}, (135)(246)) \\ &= (l_{4b} + l_{4b}, ()) \\ &= (l_0, ()).\end{aligned}$$

The relations  $\phi(w)^2 = \phi(h)^2 = \phi(s)^4 = 1$  can be checked similarly.

Finally, we need to check the left square commutes. First, we express all generators of  $\mathbb{Z}_2^4 \subset G_1$  using  $H \otimes \mathbb{1}$ ,  $S \otimes \mathbb{1}$  and SWAP:

$$\begin{aligned}X \otimes \mathbb{1} &= (H \otimes \mathbb{1})(S \otimes \mathbb{1})^2(H \otimes \mathbb{1}) \\ Z \otimes \mathbb{1} &= (S \otimes \mathbb{1})^2 \\ \mathbb{1} \otimes X &= \text{SWAP}(X \otimes \mathbb{1})\text{SWAP} \\ \mathbb{1} \otimes Z &= \text{SWAP}(Z \otimes \mathbb{1})\text{SWAP}.\end{aligned}$$

Then we calculate the image of each generator:

$$\begin{aligned}\phi \circ \iota(Z \otimes \mathbb{1}) &= (l_{3b}, (12)(34)(56))(l_{3b}, (12)(34)(56)) \\ &= (l_{3b} + l_{5b}, ()) \\ &= (l_{4b}, ()) \\ &= \iota' \circ f(Z \otimes \mathbb{1})\end{aligned}$$

and

$$\begin{aligned}\phi \circ \iota(X \otimes \mathbb{1}) &= (l_{5b}, (16)(23)(45))(l_{4b}, ())(l_{5b}, (16)(23)(45)) \\ &= (l_{5b} + l_{3b}, (16)(23)(45))(l_{5b}, (16)(23)(45)) \\ &= (l_{4b}, (16)(23)(45))(l_{5b}, (16)(23)(45)) \\ &= (l_{4b} + l_{5b}, ()) \\ &= (l_{3b}, ()) \\ &= \iota' \circ f(X \otimes \mathbb{1}).\end{aligned}$$

We can verify  $\phi \circ \iota(\mathbb{1} \otimes X) = \iota' \circ f(\mathbb{1} \otimes X)$  and  $\phi \circ \iota(\mathbb{1} \otimes Z) = \iota' \circ f(\mathbb{1} \otimes Z)$  in a similar way.

□

**Appendix C Stabilizers of  $MP_1$  vertices**

*C.1 Stabilizers of type 1 and 2 vertices*

In this section we describe the stabilizers of the vertices of  $MP_1$  in the group  $G_1 \subset Cl_2$ . Recall that  $Cl_2$  is the quotient of the normalizer of the Pauli group by the central subgroup. When we consider a unitary as an element of the Clifford group, we mean the equivalence class up to a scalar, even though this is not indicated in our notation for the sake of simplicity. For the computation of the stabilizers it suffices to choose a representative from each type of vertices. We choose  $q_0$  (type 1) and  $p_0$  (type 2). For the description of the stabilizers we will need the dihedral group whose presentation is given as follows:

$$D_{2n} = \langle a, b : a^n = b^2 = (ba)^2 = 1 \rangle. \tag{C.1}$$

**Lemma C.1** *The stabilizer of  $q_0$  is given by*

$$Stab_{G_1}(q_0) = \langle Q, R \rangle \cong D_{24}$$

where  $Q = YS \otimes X$  and  $R = YH \otimes H$ .

*Proof.* Let  $K = \langle Q, R \rangle$ . It is straight-forward to verify that  $K$  is contained in the stabilizer by explicitly checking that the vertex is fixed by  $Q$  and  $R$ . Hence  $K \subset Stab_{G_1}(q_0)$ . Since there are 48 type 1 vertices and  $G_1$  acts transitively on them by Lemma 4 we have

$$\left| \frac{G_1}{Stab_{G_1}(q_0)} \right| = 48,$$

which implies that  $|Stab_{G_1}(q_0)| = 24$ . We finish the proof by showing that  $K \cong D_{24}$ . Let  $A = QR$  then one can verify that

$$K = \langle A, R : A^{12} = R^2 = (RA)^2 = 1 \rangle \subset G_1.$$

□

**Lemma C.2** *The stabilizer of  $p_0$  is given by*

$$Stab_{G_1}(p_0) = \langle M, SWAP \rangle \cong D_{16}$$

where  $M = X \otimes YS$ .

*Proof.* Proof is similar to Lemma C.1. Let  $L = \langle M, SWAP \rangle$ . First one verifies that the given generators fix  $p_0$ , which implies that  $L$  is contained in the stabilizer. Transitivity of the action of  $G_1$  on the set of type 2 vertices (Lemma 4) implies that  $|Stab_{G_1}(p_0)| = 16$ . To conclude the proof we observe that

$$L = \langle N, SWAP : N^6 = SWAP^2 = (SWAPN)^2 = 1 \rangle \subset G_1,$$

where  $N = M SWAP$ . Therefore  $L \cong D_{16}$ . □

*C.2 Stabilizer action on the neighbors*

**Lemma C.3** *Consider the generator  $a \in D_{2n}$  in the presentation of  $D_{2n}$ ; see Eq. (C.1). If  $a^{n/2} \notin G$ , then either  $D_{2n} \cap G = \{1\}$  or there exists a unique  $i \in \{0, \dots, n-1\}$  such that  $D_{2n} \cap G = \langle a^i b \rangle \cong C_2$ .*

*Proof.* Observe that any non-trivial subgroup of  $\langle a \rangle$  will contains  $a^{n/2}$ . Since  $a^{n/2} \notin G$ , it follows that  $a^i \notin G$  for all  $i \in \{0, 1, \dots, n-1\}$  (otherwise  $a^{n/2} \in \langle a^i \rangle \subset G$ , which is a contradiction). Thus either  $D_{2n} \cap G$  is trivial or  $D_{2n} \cap G$  is generated by elements of form  $a^i b$  where  $i \in \{0, 1, \dots, n-1\}$ . Let  $g = a^i b$  and  $h = a^j b$  be two distinct elements. We have  $gh = a^{i-j}$ , which is a non-trivial elements in  $\langle a \rangle$ . Thus either  $a^i b \notin G$  for all  $i \in \{0, 1, \dots, n-1\}$ , or there exists an unique  $k \in \{0, 1, \dots, n-1\}$

such that  $a^k b \in G$ . This proves the statement.  $\square$

We will consider the following type 2 neighbors of  $p_0$ :  $p_a$  given in Eq. (37),  $p_b$  in Eq. (36) and  $q_0$  in Eq. (38).

**Lemma C.4** *Let  $N = M$  SWAP where  $M = X \otimes YS$ . We have*

1.  $Stab_{G_1}(p_0) \cap Stab_{G_1}(p_a) = \langle SWAP \rangle \cong C_2$ ,
2.  $Stab_{G_1}(p_0) \cap Stab_{G_1}(q_0) = \langle NSWAP \rangle \cong C_2$ ,
3.  $Stab_{G_1}(p_0) \cap Stab_{G_1}(p_b) = \langle N^{-1}SWAP \rangle \cong C_2$ ,
4.  $Stab_{G_1}(p_0) \cap Stab_{G_1}(p_c) = \langle NSWAP \rangle \cong C_2$ .

*Proof.* The table below shows the action of  $N, N^4, N^{-1}, SWAP, NSWAP$  and  $N^{-1}SWAP$  on the non-local Pauli operators. For simplicity we omit the tensor product notation.

$A$ : non-local Pauli	XX	XY	XZ	YX	YY	YZ	ZX	ZY	ZZ
$NAN^\dagger$	XY	-YY	-ZY	XX	-YX	-ZX	-XZ	YZ	ZZ
$N^4A(N^4)^\dagger$	XX	XY	-XZ	YX	YY	-YZ	-ZX	-ZY	ZZ
$N^\dagger AN$	YX	XX	-ZX	-YY	-XY	ZY	-YZ	-XZ	ZZ
$(SWAP)A(SWAP)^\dagger$	XX	YX	ZX	XY	YY	ZY	XZ	YZ	ZZ
$(NSWAP)A(NSWAP)^\dagger$	XY	XX	-XZ	-YY	-YX	YZ	-ZY	-ZX	ZZ
$(N^\dagger SWAP)A(N^\dagger SWAP)^\dagger$	YX	-YY	-YZ	XX	-XY	-XZ	-ZX	ZY	ZZ

Table C.1: The action of some unitaries in  $Stab_{G_1}(p_0)$ .

Using table we can show that  $N^4$  does not fix  $p_a, q_0, q_b$  and  $p_b$ . On the other hand, SWAP,  $N^{-1}SWAP$ , and  $N(SWAP)$  fixes the vertices  $p_a, q_0$ , and  $q_b$  respectively, and  $N(SWAP)$  fixes  $p_b$ . Then the statement follows from Lemma C.3.  $\square$

**Lemma C.5**  *$Stab_{G_1}(q_0)$  acts transitively on the set of neighbor vertices of  $q_0$ .*

*Proof.* By Lemma C.4, we have  $Stab_{G_1}(q_0) \cap Stab_{G_1}(p_0) \cong C_2$ . Then the orbit of  $q_0$  under the  $Stab_{G_1}(q_0)$  action has  $|D_{24}|/|C_2| = 12$  elements, which is the whole set of neighbors of  $q_0$ .  $\square$

**Lemma C.6** *The action of  $Stab_{G_1}(p_0)$  on the set of neighbor vertices of  $p_0$  breaks into three orbits with representatives given by  $q_0$  (type 1),  $p_a$  and  $p_b$  (both type 2).*

*Proof.* By Lemma C.4, we have  $Stab_{G_1}(p_0) \cap Stab_{G_1}(q_0) \cong C_2$ . Then the orbit of  $p_0$  under the action of the stabilizer has  $|D_{16}|/|C_2| = 8$  elements, which is the whole set of type 1 neighbors of  $p_0$ . By Lemma C.4, we have

$$Stab_{G_1}(p_0) \cap Stab_{G_1}(p_b) \cong Stab_{G_1}(p_0) \cap Stab_{G_1}(p_a) \cong C_2.$$

Since there are 16 type 2 neighbors of  $p_0$ , the orbit of  $Stab_{G_1}(p_0)$  on  $p_a$  and  $p_b$  both have size equal

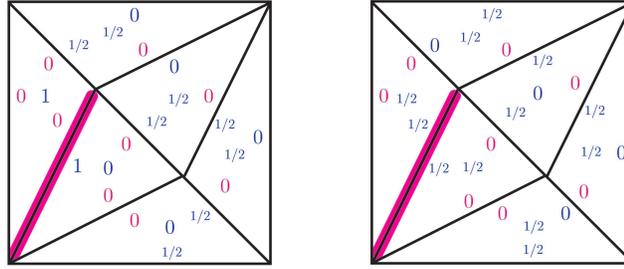


Fig. C.1: We compare the canonical vertex  $p_0$  and  $q_b$ . The 8 overlapping zeros are colored in pink. The 4 zeros on either side of the pink edge cannot all be linearly independent by Lemma 8; see Eq. (29).

to 8. It remains to check that these orbits are distinct. For this we compute the orbit:

$$\begin{aligned}
 Np_bN^\dagger &= \frac{1}{4}(\mathbb{1} + X \otimes Y - Y \otimes Y + Z \otimes X - X \otimes Z - Y \otimes Z) = p_{1b}^- \\
 (N^2)p_b(N^2)^\dagger &= \frac{1}{4}(\mathbb{1} - Y \otimes Y + Y \otimes X - X \otimes Z + Z \otimes Y + Z \otimes X) = p_{5b}^+ \\
 (N^3)p_b(N^3)^\dagger &= \frac{1}{4}(\mathbb{1} + Y \otimes X + X \otimes X + Z \otimes Y + Y \otimes Z - X \otimes Z) = p_{6b}^- \\
 (N^4)p_b(N^4)^\dagger &= \frac{1}{4}(\mathbb{1} + X \otimes X + X \otimes Y + Y \otimes Z - Z \otimes X + Z \otimes Y) = p_{3b}^- \\
 (N^5)p_b(N^5)^\dagger &= \frac{1}{4}(\mathbb{1} + X \otimes Y - Y \otimes Y - Z \otimes X + X \otimes Z + Y \otimes Z) = p_{1b}^+ \\
 (N^6)p_b(N^6)^\dagger &= \frac{1}{4}(\mathbb{1} - Y \otimes Y + Y \otimes X + X \otimes Z - Z \otimes Y - Z \otimes X) = p_a \\
 (N^7)p_b(N^7)^\dagger &= \frac{1}{4}(\mathbb{1} + Y \otimes X + X \otimes X - Z \otimes Y - Y \otimes Z + X \otimes Z) = p_{6b}^+
 \end{aligned}$$

where  $N = (X \otimes YS)(\text{SWAP})$ . Observe that  $p_a$  does not belong to the orbit of  $p_b$ . Thus, the orbits of  $p_a$  and  $p_b$  are distinct.  $\square$

We apply the stabilizer computation to show that the type 1 vertices in Fig. (19b) are not neighbors.

**Lemma C.7** *The vertices in Fig (19b) are not neighbors of  $p_0$ .*

*Proof.* Consider the vertex  $q_b$  given in Eq. (38) from the list of vertices in Fig (19b). By Lemma C.4 part (2), we have  $\text{Stab}_{G_1}(p_0) \cap \text{Stab}_{G_1}(q_b) \cong C_2$ . Then the orbit of  $q_b$  under the  $\text{Stab}_{G_1}(p_0)$  action has  $|D_{16}|/|C_2| = 8$  elements since  $\text{Stab}_{G_1}(p_0) \cong D_{16}$  by Lemma C.2. This covers the whole set of vertices in Fig (19b).

As discussed in Section 4.2, for two distributions  $q_1$  and  $q_2$  to be neighbors they must share 8 linearly independent tight inequalities. Let us consider  $q_1 = p_0$  and the type 1 vertex  $q_2 = q_b$ , and compare the number of overlapping zeros; see Fig. (C.1). There are precisely 8 such zeros. However, by Lemma 8, the two adjacent triangles on either side of the shaded edge cannot have rank 4, thus the overlapping zeros have rank  $< 8$ . Therefore  $p_0$  and  $q_b$  cannot be neighbors. Transitive action of  $\text{Stab}_{G_1}(p_0)$  on the set of Fig. (19b) implies that this holds when  $q_2$  is one of the other vertices listed in Fig (19b) as well.  $\square$

$q \in N_1(p_0)$	$U \in \text{Stab}_{G_1}(p_0)$
$q_0$	$\mathbb{1}, N^7\text{SWAP}$
$q_{4a}^+$	$N^4, N^3\text{SWAP}$
$q_{5a}^+$	$N^7, N^6\text{SWAP}$
$q_{5a}^-$	$N^3, N^2\text{SWAP}$
$q_{8a}^+$	$N^6, N^5\text{SWAP}$
$q_{8a}^-$	$N^2, N\text{SWAP}$
$q_{9a}^+$	$N, \text{SWAP}$
$q_{9a}^-$	$N^5, N^4\text{SWAP}$

(a) The action of  $\text{Stab}_{G_1}(p_0)$  on the type 1 neighbors of  $p_0$ . See Lemma C.4. The left column are  $q \in N_1(p_0)$ , type 1 neighbours of  $p_0$ . The right column are elements  $U \in \text{Stab}_{G_1}(p_0)$  such that  $Uq_0U^\dagger = q$ .

$q$ in Fig (19b)	$U \in \text{Stab}_{G_1}(p_0)$
$q_b$	$\mathbb{1}, N(\text{SWAP})$
$q_0$	$N^2, N^3\text{SWAP}$
$q_{2b}^0$	$N^6, N^7\text{SWAP}$
$q_0$	$N^4, N^5\text{SWAP}$
$q_{4b}^+$	$N^7, \text{SWAP}$
$q_{4b}^-$	$N^5, N^6\text{SWAP}$
$q_{4b}^0$	$N, N^2\text{SWAP}$
$q_0$	$N^3, N^4\text{SWAP}$

(b) The action of  $\text{Stab}_{G_1}(p_0)$  on the vertices in Fig (19b). See Lemma C.4. The left column are  $q$ , which are vertices in Fig (19b). The right column are elements  $U \in \text{Stab}_{G_1}(p_0)$  such that  $Uq_bU^\dagger = q$ .

$p \in N_2(p_0)$	$U \in \text{Stab}_{G_1}(p_0)$	$U \in \text{Stab}_{G_1}(p_0)$
$p_b$	$\mathbb{1}, N\text{SWAP}$	None
$p_{3b}^-$	$N^4, N^5\text{SWAP}$	None
$p_{1b}^+$	$N^5, N^6\text{SWAP}$	None
$p_{1b}$	$N, N^2\text{SWAP}$	None
$p_{5b}^+$	$N^2, N^3\text{SWAP}$	None
$p_a$	$N^6, N^7\text{SWAP}$	None
$p_{6b}^+$	$N^7, \text{SWAP}$	None
$p_{6b}^-$	$N^3, N^4\text{SWAP}$	None
$p_a$	None	$\mathbb{1}, \text{SWAP}$
$p_0$	None	$N^4, N^4\text{SWAP}$
$p_{7a}^+$	None	$N^2, N^2\text{SWAP}$
$p_{7a}^-$	None	$N^6, N^6\text{SWAP}$
$p_{2a}^+$	None	$N, N\text{SWAP}$
$p_{2a}^-$	None	$N^5, N^5\text{SWAP}$
$p_a$	None	$N^7, N^7\text{SWAP}$
$p_{3a}$	None	$N^3, N^3\text{SWAP}$

(c) The action of  $\text{Stab}_{G_1}(p_0)$  on the type 2 neighbors of  $p_0$ . See Lemma C.4. The left column are  $p \in N_2(p_0)$ , vertices of type 2 neighbours of  $p_0$ . The middle column are elements  $U \in \text{Stab}_{G_1}(p_0)$  such that  $Up_bU^\dagger = p$ . The right column are elements  $U \in \text{Stab}_{G_1}(p_0)$  such that  $Up_aU^\dagger = p$ .

$p \in N(q_0)$	$U \in \text{Stab}_{G_1}(q_0)$
$p_0$	$\mathbb{1}, AR$
$p_{4a}^+$	$A^6, A^7R$
$p_{7a}^+$	$A^2, A^3R$
$p_{7a}^-$	$A^8, A^9R$
$p_{9a}^+$	$A^{11}, R$
$p_{9a}^-$	$A^5, A^6R$
$p_{2a}^+$	$A^{10}, A^{11}R$
$p_{2a}^-$	$A^4, A^5R$
$p_{1a}^+$	$A^3, A^4R$
$p_{1a}^-$	$A^9, A^{10}R$
$p_{5a}^+$	$A^7, A^8R$
$p_{5a}^-$	$A, A^2R$

(d) The action of  $\text{Stab}_{G_1}(q_0)$  on the neighbors of  $q_0$ . See Lemma C.1. The left column are vertex  $p \in N(q_0)$ , neighbors of  $q_0$ . The right column are elements of  $U \in \text{Stab}_{G_1}(q_0)$  such that  $Up_0U^\dagger = p$ .

Table C.2: The action of  $\text{Stab}_{G_1}(q_0)$  on the neighbours of  $q_0$  and the action of  $\text{Stab}_{G_1}(p_0)$  on type 1 and type 2 neighbours of  $p_0$  and vertices in Fig (19b); respectively.