Sets of Marginals and Pearson-Correlation-based CHSH Inequalities for a Two-Qubit System

Yuwen Huang and Pascal O. Vontobel
Department of Information Engineering
The Chinese University of Hong Kong
hy018@ie.cuhk.edu.hk, pascal.vontobel@ieee.org

Abstract

Quantum mass functions (QMFs), which are tightly related to decoherence functionals, were introduced by Loeliger and Vontobel [IEEE Trans. Inf. Theory, 2017, 2020] as a generalization of probability mass functions toward modeling quantum information processing setups in terms of factor graphs.

Simple quantum mass functions (SQMFs) are a special class of QMFs that do *not explicitly* model classical random variables. Nevertheless, classical random variables appear *implicitly* in an SQMF if some marginals of the SQMF satisfy some conditions; variables of the SQMF corresponding to these "emerging" random variables are called classicable variables. Of particular interest are jointly classicable variables.

In this paper we initiate the characterization of the set of marginals given by the collection of jointly classicable variables of a graphical model and compare them with other concepts associated with graphical models like the sets of realizable marginals and the local marginal polytope.

In order to further characterize this set of marginals given by the collection of jointly classicable variables, we generalize the CHSH inequality based on the Pearson correlation coefficients, and thereby prove a conjecture by Pozsgay *et al*. A crucial feature of this inequality is its nonlinearity, which poses difficulties in the proof.

I. Introduction

Graphical models like factor graphs [1]–[3] have been used to represent various statistical models. In the following, we will call a factor graph consisting only of non-negative real-valued local functions a standard factor graph (S-FG). S-FGs have many applications, in particular in communications and coding theory (see, e.g., [4], [5]) and statistical mechanics (see, e.g., [6]). In these applications, factor graphs frequently represent the factorization of the joint probability mass functions (PMFs) of all the relevant random variables. Quantities of interest can then be obtained by exactly or approximately computing marginals of this joint PMF and suitably processing these marginals.

Factor graphs have also been used to represent quantum-mechanical probabilities [7], [8]. In contrast to S-FGs, these factor graphs consist of complex-valued local functions satisfying some constraints. In the following, we will call such factor graphs quantum-probability factor graphs (Q-FGs). A Q-FG is typically used to represent the factorization of the joint quantum mass function (QMF) as introduced in [7].

In this paper, we first discuss similarities and differences between PMFs and QMFs. Some of the features of QMFs will then motivate the study that is carried out in the rest of this paper.

This is an extended version of a paper that was submitted to ISIT 2021 using the same title. This work has been supported in part by the Research Grants Council of the Hong Kong Special Administrative Region, China, under Project CUHK 14209317 and Project CUHK 14207518.

II. PMFs vs. QMFs

In this section we highlight some similarities and crucial differences between PMFs and QMFs. First, we consider a classical setup. In particular, we assume that we are interested in a graphical model that represents the joint PMF $P_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n)$, where $Y_1,\ldots,Y_n$ are some random variables of interest taking value in some alphabets $\mathcal{Y}_1,\ldots,\mathcal{Y}_n$.[1] (In a typical application, we might have observed $Y_1 = y_1,\ldots,Y_{n-1} = y_{n-1}$ and would like to estimate Y_n based on these observations.) In most applications, the PMF $P_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n)$ does *not* have a "nice" factorization in terms of simple factors. However, frequently, with the introduction of suitable auxiliary variables $x_1,\ldots,x_m$ taking values in some alphabets $\mathcal{X}_1,\ldots,\mathcal{X}_m$, respectively, there is a function $p(\boldsymbol{x},\boldsymbol{y})$, where $\boldsymbol{x} := (x_1,\ldots,x_m)$ and $\boldsymbol{y} := (y_1,\ldots,y_n)$, such that

$$p(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}_{\geq 0} \quad \text{(for all } \boldsymbol{x},\boldsymbol{y}) \ ,$$

$$\sum_{\boldsymbol{x},\boldsymbol{y}} p(\boldsymbol{x},\boldsymbol{y}) = 1 \ ,$$

$$\sum_{\boldsymbol{x}} p(\boldsymbol{x},\boldsymbol{y}) = P_{\boldsymbol{Y}}(\boldsymbol{y}) \quad \text{(for all } \boldsymbol{y}) \ ,$$

and such that $p(\boldsymbol{x},\boldsymbol{y})$ has a "nice" factorization. (For example, in a hidden Markov model, the joint PMF of the observations does not have a "nice" factorization, but the joint PMF of the hidden state process and the observations has a "nice" factorization.) Note that the function $p(\boldsymbol{x},\boldsymbol{y})$ can, thanks to its properties, be considered as a joint PMF of some random variables $X_1,\ldots,X_m,Y_1,\ldots,Y_n$.

Second, we consider a quantum-mechanical setup. We assume, again, that we are interested in a graphical model that represents the joint PMF $P_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n)$, where $Y_1,\ldots,Y_n$ are some random variables of interest taking values in some alphabets $\mathcal{Y}_1,\ldots,\mathcal{Y}_n$. Such random variables can, for example, represent the measurements obtained when running some quantum-mechanical experiment, and we might be interested in estimating Y_n based on the observations $Y_1 = y_1,\ldots,Y_{n-1} = y_{n-1}$. As in the classical case, the PMF $P_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n)$ usually does *not* have a "nice" factorization in terms of simple factors. Moreover, standard physical modeling of quantum-mechanical systems shows that introducing a function $p(\boldsymbol{x},\boldsymbol{y})$ as defined above does usually *not* help toward obtaining a function with a "nice" factorization. However, in many quantum-mechanical setups of interest, with the introduction of suitable auxiliary variables $x_1,\ldots,x_m,x'_1,\ldots,x'_m$ taking values in some alphabets $\mathcal{X}_1,\ldots,\mathcal{X}_m$, $\mathcal{X}'_1,\ldots,\mathcal{X}'_m$ (with $\mathcal{X}'_i = \mathcal{X}_i$, $i \in \{1,\ldots,m\}$), there is a function $q(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y})$, called quantum mass function (QMF) [7], such that

$$q(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y}) \in \mathbb{C} \quad \text{(for all } \boldsymbol{x},\boldsymbol{x},\boldsymbol{y}) \ ,$$

$$\sum_{\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y}} q(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y}) = 1 \ ,$$

$$q(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y}) \text{ is a PSD kernel in } (\boldsymbol{x},\boldsymbol{x}') \text{ for every } \boldsymbol{y} \ ,$$

$$\sum_{\boldsymbol{x},\boldsymbol{x}'} q(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y}) = P_{\boldsymbol{Y}}(\boldsymbol{y}) \quad \text{(for all } \boldsymbol{y}) \ ,$$

and such that $q(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y})$ has a "nice" factorization. The major difference between $p(\boldsymbol{x},\boldsymbol{y})$ and $q(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y})$ is the fact that the former takes value in $\mathbb{R}_{\geq 0}$, whereas the latter takes value in $\mathbb{C}$. In particular, $\sum_{\boldsymbol{y}} q(\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y})$ is in general not a PMF over $(\boldsymbol{x},\boldsymbol{x}')$, thereby showing that $\boldsymbol{x},\boldsymbol{x}'$ *cannot* be considered as random variables. (See [7] for more details.)

In [8], the authors discussed an approach to QMFs where $\boldsymbol{y}$ does *not* appear explicitly anymore, but "emerges" from a QMF. More precisely, they first introduced a simple quantum mass function (SQMF) $q(\boldsymbol{x},\boldsymbol{x}')$ that satisfies

$$q(\boldsymbol{x},\boldsymbol{x}') \in \mathbb{C}_{\geq 0} \quad \text{(for all } \boldsymbol{x},\boldsymbol{x}') \ ,$$

[1] For simplicity, in the following all alphabets will be finite.

$$\sum_{\boldsymbol{x},\boldsymbol{x}'} q(\boldsymbol{x},\boldsymbol{x}') = 1 \ ,$$

$$q(\boldsymbol{x},\boldsymbol{x}') \text{ is a PSD kernel in } (\boldsymbol{x},\boldsymbol{x}') \ .$$

Afterwards, they defined "classicable" variables.

Definition 1. *Let $\mathcal{I}$ be a subset of $\{1,\ldots,m\}$ and let $\mathcal{I}^{\mathrm{c}} := \{1,\ldots,m\} \setminus \mathcal{I}$ be its complement. The variables $\boldsymbol{x}_{\mathcal{I}}$ are called jointly classicable if the function*

$$q(\boldsymbol{x}_{\mathcal{I}},\boldsymbol{x}'_{\mathcal{I}}) := \sum_{\boldsymbol{x}_{\mathcal{I}^{\mathrm{c}}},\boldsymbol{x}'_{\mathcal{I}^{\mathrm{c}}}} q(\boldsymbol{x},\boldsymbol{x}')$$

is zero for all $(\boldsymbol{x}_{\mathcal{I}},\boldsymbol{x}'_{\mathcal{I}})$ satisfying $\boldsymbol{x}_{\mathcal{I}} \neq \boldsymbol{x}'_{\mathcal{I}}$.[2]

Note that if $\boldsymbol{x}_{\mathcal{I}}$ are jointly classicable, then one can define the function $p(\boldsymbol{x}_{\mathcal{I}}) := q(\boldsymbol{x}_{\mathcal{I}},\boldsymbol{x}_{\mathcal{I}})$, for which it is straightforward, thanks to the properties of SQMFs, to show that it is a PMF. It is in this sense that random variables $y_1,\ldots,y_n$ that were omitted when going from QMFs to SQMFs can "emerge" again.[3]

Definition 2. *Let $\mathcal{K}$ be a collection of subsets $\mathcal{I}$ of $\{1,\ldots,m\}$ such that $\boldsymbol{x}_{\mathcal{I}}$ is classicable.*

Example 3. *Consider the Q-FG N_4 in Fig. 4, whose global function is an SQMF. In that Q-FG, ρ represents a PSD matrix and U_1, U_2 are unitary matrices. One can show that for all choices of ρ, U_1, and U_2, the collection $\mathcal{K}$ can be chosen to contain the sets $\{1,2\}$, $\{1,4\}$, $\{2,3\}$, and $\{3,4\}$.*[4]

Thoughout this paper, we consider

$$\mathcal{K} = \{\{1,2\},\{1,4\},\{2,3\},\{3,4\}\}. \tag{1}$$

Interestingly enough, the collection of functions $\big\{p(\boldsymbol{x}_{\mathcal{I}})\big\}_{\mathcal{I}\in\mathcal{K}}$ is usually such that there is *no* PMF $p(\boldsymbol{x})$ such that for every $\mathcal{I} \in \mathcal{K}$, the function $p(\boldsymbol{x}_{\mathcal{I}})$ can be obtained as a marginal of $p(\boldsymbol{x})$.[5] In general, we can only guarantee that for two sets $\mathcal{I}_1$, $\mathcal{I}_2 \in \mathcal{K}$ the following consistency constraint holds:

$$\begin{aligned}
\sum_{\boldsymbol{x}_{\mathcal{I}_1\setminus\mathcal{I}_2}} p(\boldsymbol{x}_{\mathcal{I}_1}) &= \sum_{\boldsymbol{x}_{\mathcal{I}_1\setminus\mathcal{I}_2},\boldsymbol{x}'_{\mathcal{I}_1\setminus\mathcal{I}_2}} q(\boldsymbol{x}_{\mathcal{I}_1},\boldsymbol{x}'_{\mathcal{I}_1}) \\
&\overset{(a)}{=} \sum_{\boldsymbol{x},\boldsymbol{x}':\boldsymbol{x}_{\mathcal{I}_1\cap\mathcal{I}_2},\boldsymbol{x}'_{\mathcal{I}_1\cap\mathcal{I}_2}\text{ fixed}} q(\boldsymbol{x},\boldsymbol{x}') \\
&= \sum_{\boldsymbol{x}_{\mathcal{I}_2\setminus\mathcal{I}_1},\boldsymbol{x}'_{\mathcal{I}_2\setminus\mathcal{I}_1}} q(\boldsymbol{x}_{\mathcal{I}_2},\boldsymbol{x}'_{\mathcal{I}_2}) \\
&= \sum_{\boldsymbol{x}_{\mathcal{I}_2\setminus\mathcal{I}_1}} p(\boldsymbol{x}_{\mathcal{I}_2}) \quad (\text{for all } \boldsymbol{x}_{\mathcal{I}_1\cap\mathcal{I}_2}),
\end{aligned} \tag{2}$$

where at step (a) we have used Definition 1.

Let us comment on these special properties of $\big\{p(\boldsymbol{x}_{\mathcal{I}})\big\}_{\mathcal{I}\in\mathcal{K}}$:

- It turns out that these special properties of $\big\{p(\boldsymbol{x}_{\mathcal{I}})\big\}_{\mathcal{I}\in\mathcal{K}}$ are at the heart of quantum mechanical phenomena like Hardy's paradox [12] and the Frauchiger–Renner paradox [13].[6] In fact, the Q-FG N_4 in Fig. 4 can be used to analyze Hardy's paradox. On the side, note that the Q-FG in Fig. 4 also captures the essence of Bell's game [14].

[2] It would be more precise to call this function $q_{\mathcal{I}}$. However, for conciseness, we drop the index $\mathcal{I}$ as it can be inferred from the arguments.

[3] Note that there is a strong connection of SQMFs to the so-called decoherence functional [9], [10], and via this also to the consistent-histories approach to quantum mechanics [11]. However, the starting point of our investigations is quite different.

[4] For special ρ, the set $\mathcal{K}$ contains more elements.

[5] A similar observation is at the origin of the so-called "single-framework" rule in the consistent-histories approach to quantum mechanics.

[6] For a discussion of the latter in terms of SQMFs, see [8].

- Interestingly, these special properties of $\{p(\boldsymbol{x}_{\mathcal{I}})\}_{\mathcal{I}\in\mathcal{K}}$ are very similar to the properties of beliefs in the local marginal polytope of an S-FG (see, e.g., [15]).[7]

The above observations motivate the systematic study of the collection $\{p(\boldsymbol{x}_{\mathcal{I}})\}_{\mathcal{I}\in\mathcal{K}}$ for a given SQMF. Indeed, one key contribution of this paper is to study this collection for the Q-FG in Fig. 4 and compare it with other objects that can be associated with this Q-FG.

III. Motivation and Contributions

Bell inequalities [18] derived by the Bell theorem are useful tools for studying classical variables for PMFs and classicable variables for QMFs. Linear Bell inequalities provide a characterization of joint PMFs [19]–[21], which indicates the Bell inequalities' application in characterizing sets of marginals. The simplest and most well-known Bell inequality is the Clauster-Horne-Shimony-Holt (CHSH) [22] inequality, i.e.,

$$|\mathbb{E}(Z_1 \cdot Z_2) + \mathbb{E}(Z_1 \cdot Z_4) + \mathbb{E}(Z_3 \cdot Z_2) - \mathbb{E}(Z_1 \cdot Z_2)| \leq 2,$$

where $Z_1, \ldots, Z_4$ are binary random variables in $\{-1, 1\}$. In [23], the authors considered the CHSH inequality in terms of the covariance and the Pearson correlation coefficients (PCCs), i.e.,

$$|\mathrm{Cov}(Z_1, Z_2) + \mathrm{Cov}(Z_1, Z_4) + \mathrm{Cov}(Z_3, Z_2) - \mathrm{Cov}(Z_1, Z_2)| \leq \frac{16}{7}, \tag{3}$$

$$|\mathrm{Corr}(Z_1, Z_2) + \mathrm{Corr}(Z_1, Z_4) + \mathrm{Corr}(Z_3, Z_2) - \mathrm{Corr}(Z_1, Z_2)| \leq \frac{5}{2}, \tag{4}$$

where $\mathrm{Cov}(Z_i, Z_j)$ and $\mathrm{Corr}(Z_i, Z_j)$ are covariance and PCC for binary random variables Z_i, $Z_j \in [-1, 1]$ and $\{i, j\} \in \mathcal{K}$, respectively. Note that these inequalities are non-linear for the PMF of $Z_1, \ldots, Z_4$. The authors in [23, Appendix A] proved (3) and conjected (4). They also proved Tsirelson's bound [24] for the covariance measure in [23, Appendix B.1]. The proposed conjecture and the proving techniques they used partially motivate our work.

To better understand classicable variables' marginals, we define the set $\mathcal{M}(\mathsf{N}_4)$, which is the set of the marginals created by the classicable variables in the two-qubit system N_4 in Fig. 4. One of our paper's main topics is to fully characterize $\mathcal{M}(\mathsf{N}_4)$. For comparison, we introduce $\mathcal{LM}(\mathcal{K})$ (the local marginal polytope of the S-FG N_1 in Fig. 1), $\mathcal{M}(\mathsf{N}_1)$ (the set of realizable marginals of N_1), $\mathcal{M}(\mathsf{N}_2)$ (the set of realizable marginals of the Markov chain N_2 in Fig. 2), and $\mathcal{M}(\mathsf{N}_3)$ (the set of realizable marginals of N_3 in Fig. 3). We derive the following results.

- We prove the Venn diagram in Fig. 7 by showing that each part in the diagram is non-empty. We can see that the sets of realizable marginals $\mathcal{M}(\mathsf{N}_3)$ and $\mathcal{M}(\mathsf{N}_4)$ are strict subsets of $\mathcal{LM}(\mathcal{K})$. In particular, both $\mathcal{M}(\mathsf{N}_1)$ and $\mathcal{M}(\mathsf{N}_2)$ have marginals that are not in $\mathcal{M}(\mathsf{N}_4)$; the set $\mathcal{M}(\mathsf{N}_4)$ consists of marginals that are not compatible with any joint PMF.
- We generalize the Clauser-Horne-Shimony-Holt (CHSH) inequality [22] for Pearson correlation coefficients (PCCs), which resolves a conjecture proposed in [23]. Because PCCs are non-linear functions with respect to marginals, the inequality has a non-trivial proof. We suspect that the proof approach is applicable for proving other non-linear Bell inequalities. A violation of this inequality indicates that the associated marginals are not in $\mathcal{M}(\mathsf{N}_3)$.
- We illustrate Hardy's paradox, Bell's game, and the maximum quantum violation of the PCC-based CHSH inequality by the classical variables in N_4.

Besides these specific results, our paper is, more generally, about leveraging tools from factor graphs to understand certain quantities of interest in quantum information processing. In particular, given that factor graphs have been proven very useful in classical information processing, but can also be used for doing quantum information processing, they allow one to understand and appreciate the similarities and the differences between classical and quantum information processing.

[7]Local marginal polytopes are of relevance, for example, when characterizing locally operating message-passing iterative algorithms like the sum-product algorithm [16], [17].

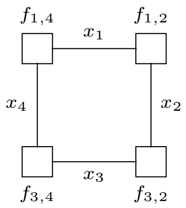

Fig. 1: The S-NFG N_1.

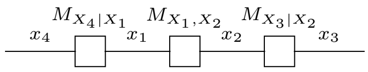

Fig. 2: The Markov chain N_2.

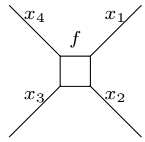

Fig. 3: The S-NFG N_3.

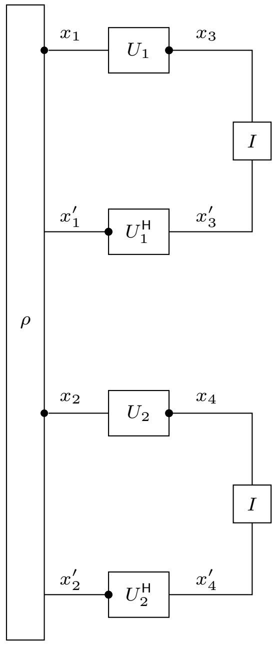

Fig. 4: The Q-NFG N_4.

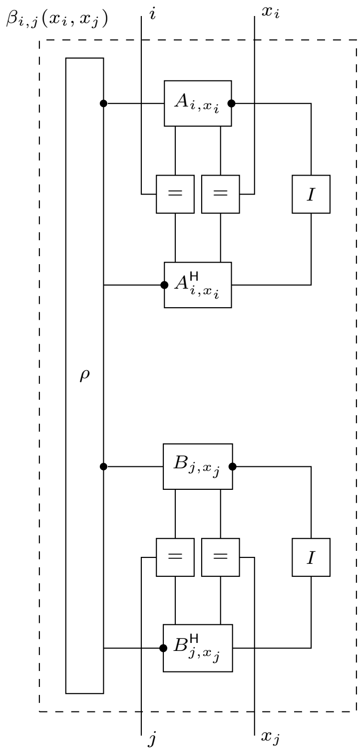

Fig. 5: The NFG representation of $\beta_{i,j}(x_i, x_j)$.

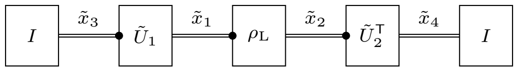

Fig. 6: The DE-NFG N_5.

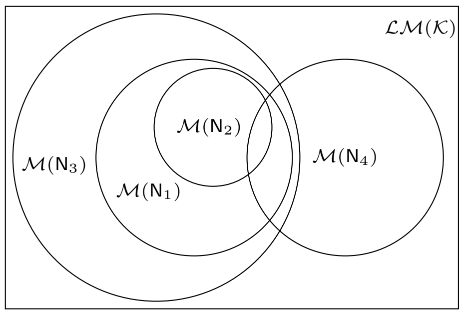

Fig. 7: The Venn diagram for $\mathcal{M}(\mathsf{N}_2)$, $\mathcal{M}(\mathsf{N}_1)$, $\mathcal{M}(\mathsf{N}_3)$, $\mathcal{M}(\mathsf{N}_4)$, and $\mathcal{LM}(\mathcal{K})$.

The rest of this paper is structured as follows. Section IV reviews some basics of S-FGs. Section IV-A discusses some properties for N_3 in Fig. 3. In particular, Section IV-A proves the PCC-based CHSH inequality and discusses the Markov chain in Fig. 2. Section V introduces the factor graphs N_4 and N_5 depicted in Figs. 4 and 6, respectively. Section VI proves the Venn diagram in Fig. 7.

A. Basic Notations and Definitions

The sets $\mathbb{Z}$, $\mathbb{Z}_{\geq 0}$, $\mathbb{Z}_{>0}$, $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$, and $\mathbb{C}$ denote the ring of integers, the set of nonnegative integers, the set of positive integers, the field of real numbers, the set of nonnegative real numbers, the set of positive real numbers, and the field of complex numbers, respectively. An overline denotes complex conjugation. Square brackets are used in two different ways.

Namely, for any $L \in \mathbb{Z}_{>0}$, the function $[L]$ is defined to be the set $[L] := \{1, \ldots, L\}$ with cardinality L and for any statement S, by the Iverson's convention, the function $[S]$ is defined to be $[S] := 1$ if S is true and $[S] := 0$ otherwise. For any vector $\boldsymbol{v} := \begin{pmatrix} v_1 & \ldots & v_N \end{pmatrix} \in \mathbb{C}^N$, we define

$$\mathrm{diag}(\boldsymbol{v}) := \begin{pmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_N \end{pmatrix}.$$

For any matrix $M \in \mathbb{C}^{N \times N}$, $N \in \mathbb{Z}_{>0}$, and $i, j \in [N]$, the vector $M(i, :)$ represents the i-th row of M and $M(:, j)$ represents the j-th column of M. The matrix $M(i_1 : i_2, j_1 : j_2)$ represents the submatrix of M s.t.

$$M(i_1 : i_2, j_1 : j_2) := \begin{pmatrix} M(i_1, j_1) & \cdots & M(i_1, j_2) \\ \vdots & \ddots & \vdots \\ M(i_2, j_1) & \cdots & M(i_2, j_2) \end{pmatrix}, \qquad i_1 < i_2,\ j_1 < j_2.$$

IV. Standard Normal Factor Graphs (S-NFGs)

In this section, we review some basic concepts and properties of an S-NFG. The word "normal" refers to the fact that variables are arguments of only one or two local functions. We use an example to introduce the fundamental concepts of an S-NFG first.

Example 4. *[1], [3] Consider the multivariate function*

$$g_{\mathsf{N}_1}(x_1, \ldots, x_4) := f_{1,2}(x_1, x_2) \cdot f_{1,4}(x_1, x_4) \cdot f_{3,2}(x_3, x_2) \cdot f_{3,4}(x_3, x_4)$$

where g_{N_1}, the so-called global function, is defined to be the product of the so-called local functions $f_{1,2}$, $f_{1,4}$, $f_{3,2}$ and $f_{3,4}$. We can visualize the factorization of g with the help of the S-FG N_1 in Fig. 1. Note that the S-FG N_1 consists of four function nodes $f_{1,2}, \ldots, f_{3,4}$ and four (full) edges with associated variables $x_1, \ldots, x_4$.

For an S-NFG, a half edge is an edge incident on one function node and a full edge is an edge incident on two function nodes.

Definition 5. *The S-NFG $\mathsf{N}(\mathcal{F}(\mathsf{N}), \mathcal{E}(\mathsf{N}), \mathcal{X}(\mathsf{N}))$ consists of:*

1) *The graph $(\mathcal{F}(\mathsf{N}), \mathcal{E}(\mathsf{N}))$ with vertex set $\mathcal{F}(\mathsf{N})$ and edge set $\mathcal{E}(\mathsf{N})$, where $\mathcal{E}(\mathsf{N})$ consists of all full edges and half edges in N. With some slight abuse of notation, an $f \in \mathcal{F}(\mathsf{N})$ will denote a function node and the corresponding local function.*
2) *The alphabet $\mathcal{X}(\mathsf{N}) := \prod_{e \in \mathcal{E}(\mathsf{N})} \mathcal{X}_e$, where $\mathcal{X}_e$ is the alphabet associated with edge $e \in \mathcal{E}(\mathsf{N})$.*

Definition 6. *Given $\mathsf{N}(\mathcal{F}(\mathsf{N}), \mathcal{E}(\mathsf{N}), \mathcal{X}(\mathsf{N}))$, we make the following definitions:*

1) *For every function node $f \in \mathcal{F}(\mathsf{N})$, the set ∂f is the set of edges incident on f. The degree of f is defined to be $|\partial f|$.*
2) *An assignment $\boldsymbol{x} := (x_e)_{e \in \mathcal{E}(\mathsf{N})} \in \mathcal{X}(\mathsf{N})$ is called a configuration of the S-FG. For each $f \in \mathcal{F}(\mathsf{N})$, a configuration $\boldsymbol{x} \in \mathcal{X}(\mathsf{N})$ induces the vector $\boldsymbol{x}_{\partial f}$ with components $\boldsymbol{x}_{\partial f} := (x_e)_{e \in \partial f} \in \prod_{e \in \partial f} \mathcal{X}_e$.*
3) *The local function f associated with function node $f \in \mathcal{F}(\mathsf{N})$ denotes an arbitrary mapping*

$$f : \prod_{e \in \partial f} \mathcal{X}_e \to \mathbb{R}_{\geq 0}.$$

4) *The global function is defined to be*

$$g_{\mathsf{N}}(\boldsymbol{x}) := \prod_{f \in \mathcal{F}(\mathsf{N})} f(x_{\partial f}).$$

5) *A configuration* $\boldsymbol{x}$ *with* $g_{\mathsf{N}}(\boldsymbol{x}) \neq 0$ *is called a valid configuration. The set of all valid configurations, i.e.*

$$\mathcal{C}(\mathsf{N}) := \{\boldsymbol{x} | g_{\mathsf{N}}(\boldsymbol{x}) \neq 0\},$$

is called the global behavior of N*, the full behavior of* N*, or the edge-based code realized by* N*.*

6) *The partition function is defined to be*

$$Z(\mathsf{N}) := \sum_{\boldsymbol{x}} g_{\mathsf{N}}(\boldsymbol{x}),$$

where $\sum_{\boldsymbol{x}}$ *denotes* $\sum_{\boldsymbol{x} \in \mathcal{X}(\mathsf{N})}$.

7) *The probability mass function (PMF) induced on* N *is defined to be the function*

$$p_{\mathsf{N}}(\boldsymbol{x}) := \frac{g_{\mathsf{N}}(\boldsymbol{x})}{Z(\mathsf{N})}.$$

8) *Let* $\mathcal{I}$ *be a subset of* $\mathcal{E}(\mathsf{N})$ *and let* $\mathcal{I}^{\mathrm{c}} := \mathcal{E}(\mathsf{N}) \setminus \mathcal{I}$ *be its complement. The marginal* $p_{\mathsf{N},\mathcal{I}}(\boldsymbol{x}_{\mathcal{I}})$ *is defined to be*

$$p_{\mathsf{N},\mathcal{I}}(\boldsymbol{x}_{\mathcal{I}}) := \sum_{\boldsymbol{x}_{\mathcal{I}^{\mathrm{c}}}} p_{\mathsf{N}}(\boldsymbol{x}), \qquad \boldsymbol{x}_{\mathcal{I}} \in \mathcal{X}_e^{|\mathcal{I}|}.$$

Definition 7. *Considering* $\mathcal{K}$ *given in* (1) *and* $\mathsf{N} \in \{\mathsf{N}_1, \mathsf{N}_2, \mathsf{N}_3\}$*, we make the following definitions:*

1) *The alphabet* $\mathcal{X}_e$ *is* $\mathcal{X}_e := \{0,1\}$ *for all* $e \in \mathcal{E}(\mathsf{N})$.

2) *The matrices* $p_{\mathsf{N},i,j}$ *and* $p_{\mathsf{N},i}$ *are defined to be*

$$p_{\mathsf{N},i,j} := \begin{pmatrix} p_{\mathsf{N},i,j}(0,0) & p_{\mathsf{N},i,j}(0,1) \\ p_{\mathsf{N},i,j}(1,0) & p_{\mathsf{N},i,j}(1,1) \end{pmatrix}, \qquad \{i,j\} \in \mathcal{K},$$

$$p_{\mathsf{N},i} := \begin{pmatrix} p_{\mathsf{N},i}(0) & 0 \\ 0 & p_{\mathsf{N},i}(1) \end{pmatrix}, \qquad i \in \mathcal{E}(\mathsf{N}).$$

3) *The set of matrices* $\boldsymbol{\beta}$ *is defined to be* $\boldsymbol{\beta} := \big((\boldsymbol{\beta}_{i,j})_{\{i,j\}\in\mathcal{K}}, (\boldsymbol{\beta}_i)_{i\in\mathcal{E}(\mathsf{N})}\big)$*, where the matrices* $\boldsymbol{\beta}_{i,j}$ *and* $\boldsymbol{\beta}_i$ *are defined to be*

$$\boldsymbol{\beta}_{i,j} := \begin{pmatrix} \beta_{i,j}(0,0) & \beta_{i,j}(0,1) \\ \beta_{i,j}(1,0) & \beta_{i,j}(1,1) \end{pmatrix} \in \mathbb{R}_{\geq 0}^{2\times 2}, \qquad \{i,j\} \in \mathcal{K},$$

$$\boldsymbol{\beta}_i := \begin{pmatrix} \beta_i(0) & 0 \\ 0 & \beta_i(1) \end{pmatrix} \in \mathbb{R}_{\geq 0}^{2\times 2}, \qquad i \in \mathcal{E}(\mathsf{N}).$$

4) *The set of realizable marginals* $\mathcal{M}(\mathsf{N})$ *is defined to be*

$$\mathcal{M}(\mathsf{N}) := \left\{\boldsymbol{\beta} \,\middle|\, \text{there exists } (p_{\mathsf{N},i,j})_{i,j\in\mathcal{K}} \text{ and } (p_{\mathsf{N},i})_{i\in\mathcal{E}(\mathsf{N})} \text{ such that (6) holds}\right\}, \tag{5}$$

where

$$\boldsymbol{\beta}_{i,j} = p_{\mathsf{N},i,j}, \qquad \boldsymbol{\beta}_i = p_{\mathsf{N},i}, \qquad \boldsymbol{\beta}_j = p_{\mathsf{N},j}, \qquad \{i,j\} \in \mathcal{K}. \tag{6}$$

5) *The set* $\mathcal{LM}(\mathcal{K})$ *is defined to be*

$$\mathcal{LM}(\mathcal{K}) := \{\boldsymbol{\beta} \,|\, \text{(7)–(8) hold}\},$$

where

$$0 \leq \beta_f(x_i, x_j) \leq 1, \quad \sum_{x_j \in \mathcal{X}_e} \beta_{i,j}(x_i, x_j) = \beta_i(x_i), \quad \sum_{x_i \in \mathcal{X}_e} \beta_{i,j}(x_i, x_j) = \beta_j(x_j), \quad x_i, x_j \in \mathcal{X}_e, \ \{i,j\} \in \mathcal{K}, \tag{7}$$

$$\sum_{x_i \in \mathcal{X}_e} \beta_i(x_i) = 1, \qquad i \in \mathcal{E}(\mathsf{N}). \tag{8}$$

The set $\mathcal{LM}(\mathcal{K})$ *contains* $\boldsymbol{\beta}$ *such that (s.t.) for each* $(i,j) \in \mathcal{K}$*, the matrix* $\boldsymbol{\beta}_{i,j}$ *satisfies the normalization condition* (8) *and the marginalization constraints* (7)*, i.e.,* $(\boldsymbol{\beta}_{i,j})_{i,j}$ *are locally consistent PMFs. The set* $\mathcal{LM}(\mathcal{K})$ *is called the local* *marginal polytope of the S-NFG N_1 in Fig. 1. The definition of the local marginal polytope of an S-NFG is given in [15, Section 4.1.1].*

6) *For each $\boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K})$ and $\{i,j\} \in \mathcal{K}$, each marginal $\boldsymbol{\beta}_{i,j}$ can be used to represent the PMF for two random variables Y_1 and Y_2 in $\mathcal{X}_e$ s.t.*

$$\Pr(Y_1 = x_i, Y_2 = x_j) = \beta_{i,j}(x_i, x_j), \qquad x_i, x_j \in \mathcal{X}_e.$$

The covariance of Y_1 and Y_2 equals

$$\mathrm{Cov}(\boldsymbol{\beta}_{i,j}) := \mathbb{E}\Big(\big(Y_1 - \mathbb{E}(Y_1)\big)\cdot\big(Y_2 - \mathbb{E}(Y_2)\big)\Big) = \sum_{x_i,x_j} x_i \cdot x_j \cdot \beta_{i,j}(x_i,x_j) - \left(\sum_{x_i\in\mathcal{X}_e} \beta_i(x_i)\cdot x_i\right)\left(\sum_{x_j\in\mathcal{X}_e} \beta_j(x_j)\cdot x_j\right) \tag{9}$$

The variances of Y_1 and Y_2 equal

$$\mathrm{Var}(\boldsymbol{\beta}_i) := \mathrm{Var}(Y_1) = \mathbb{E}\big((Y_1 - \mathbb{E}(Y_1))\big)^2 = \sum_{x_i\in\mathcal{X}_e} \beta_i(x_i)\cdot x_i - \left(\sum_{x_i\in\mathcal{X}_e} \beta_i(x_i)\cdot x_i\right)^2, \tag{10}$$

$$\mathrm{Var}(\boldsymbol{\beta}_j) := \mathrm{Var}(Y_2) = \mathbb{E}\big((Y_2 - \mathbb{E}(Y_2))\big)^2 = \sum_{x_j\in\mathcal{X}_e} \beta_j(x_j)\cdot x_j - \left(\sum_{x_j\in\mathcal{X}_e} \beta_j(x_j)\cdot x_j\right)^2. \tag{11}$$

When $\mathrm{Var}(Y_1), \mathrm{Var}(Y_2) > 0$, the Pearson correlation coefficient of Y_1 and Y_2 is defined to be

$$\mathrm{Corr}(\boldsymbol{\beta}_{i,j}) := \frac{\mathrm{Cov}(Y_1, Y_2)}{\sqrt{\mathrm{Var}(Y_1)\cdot\mathrm{Var}(Y_2)}}. \tag{12}$$

Note that $\mathcal{E}(\mathsf{N}_1) = \mathcal{E}(\mathsf{N}_2) = \mathcal{E}(\mathsf{N}_3)$. In the rest of this paper, we use $\mathcal{E}(\mathsf{N}_1)$ instead of $\mathcal{E}(\mathsf{N})$ when $\mathsf{N} \in \{\mathsf{N}_1, \mathsf{N}_2, \mathsf{N}_3\}$. For any $\{i,j\} \in \mathcal{K}$, we use $\sum_{x_i,x_j}$ and $\sum_{x_i}$ instead of $\sum_{x_i,x_j\in\mathcal{X}_e}$ and $\sum_{x_i\in\mathcal{X}_e}$ when there is no ambiguity. We also use $\{\cdot\}_{i,j}$, $(\cdot)_{i,j}$, $(\cdot)_i$, and $\{\cdot\}_{\boldsymbol{x}}$ for $\{\cdot\}_{\{i,j\}\in\mathcal{K}}$, $(\cdot)_{\{i,j\}\in\mathcal{K}}$, $(\cdot)_{i\in\mathcal{E}(\mathsf{N}_1)}$, and $\{\cdot\}_{\boldsymbol{x}\in\mathcal{X}(\mathsf{N}_1)}$ respectively.

Because $\mathcal{LM}(\mathcal{K})$ is a convex set by its definition, Carathéodory's theorem [25, Proposition B.6] implies that each element in $\mathcal{LM}(\mathcal{K})$ can be written as a linear combination of the vertices of $\mathcal{LM}(\mathcal{K})$. The vertices of $\mathcal{LM}(\mathcal{K})$ following modulo cyclic symmetries are listed as follows:

	$\boldsymbol{\beta}_{1,4}$	$\boldsymbol{\beta}_{1,2}$	$\boldsymbol{\beta}_{3,2}$	$\boldsymbol{\beta}_{3,4}$		$\boldsymbol{\beta}_{1,4}$	$\boldsymbol{\beta}_{1,2}$	$\boldsymbol{\beta}_{3,2}$	$\boldsymbol{\beta}_{3,4}$
$\boldsymbol{v}_1$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\boldsymbol{v}_5$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$
$\boldsymbol{v}_9$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\boldsymbol{v}_{10}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$
$\boldsymbol{v}_{15}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\boldsymbol{v}_{16}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$
$\boldsymbol{v}_{11}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\boldsymbol{v}_{17}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$
$\boldsymbol{v}_{21}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$					

Modulo cyclic symmetries means that by tranpose and circular shifts of each marginal in $(\boldsymbol{\beta}_{i,j})_{i,j}$ for each vertex listed in the above table, we can obtain all the vertices of $\mathcal{LM}(\mathcal{K})$. The full list of the vertices is in Appendix A.

Lemma 8. *It holds that*

$$\mathrm{Var}(\boldsymbol{\beta}_i) = \sum_{x_i\in\mathcal{X}_e} \beta_i(x_i)\cdot x_i - \left(\sum_{x_i\in\mathcal{X}_e} \beta_i(x_i)\cdot x_i\right)^2 = \det(\boldsymbol{\beta}_i), \qquad i \in \{1,3\},$$

$$\mathrm{Var}(\boldsymbol{\beta}_j) = \sum_{x_j \in \mathcal{X}_e} \beta_j(x_j) \cdot x_j - \left(\sum_{x_j \in \mathcal{X}_e} \beta_j(x_j) \cdot x_j \right)^2 = \det(\boldsymbol{\beta}_j), \qquad j \in \{2, 4\},$$

$$\mathrm{Cov}(\boldsymbol{\beta}_{i,j}) = \sum_{x_i, x_j} x_i \cdot x_j \cdot \beta_{i,j}(x_i, x_j) - \left(\sum_{x_i \in \mathcal{X}_e} \beta_i(x_i) \cdot x_i \right) \left(\sum_{x_j \in \mathcal{X}_e} \beta_j(x_j) \cdot x_j \right) = \det(\boldsymbol{\beta}_{i,j}), \qquad \{i, j\} \in \mathcal{K}. \tag{13}$$

Recall that the definitions of $\mathcal{K}$ *is given in* (1).

Proof. See Appendix B. ■

Corollary 9. *For* $\{i, j\} \in \mathcal{K}$ *and* $0 < \beta_i(0), \beta_j(0) < 1$*, the PCC* $\mathrm{Corr}(\boldsymbol{\beta}_{i,j})$ *satisfies*

$$\mathrm{Cov}(\boldsymbol{\beta}_{i,j}) = \det(\boldsymbol{\beta}_{i,j}), \qquad \mathrm{Corr}(\boldsymbol{\beta}_{i,j}) = \frac{\det(\boldsymbol{\beta}_{i,j})}{\sqrt{\det(\boldsymbol{\beta}_i) \cdot \det(\boldsymbol{\beta}_j)}} = \frac{\beta_{i,j}(0,0) \cdot \beta_{i,j}(1,1) - \beta_{i,j}(0,1) \cdot \beta_{i,j}(1,0)}{\sqrt{\beta_i(0) \cdot \beta_i(1) \cdot \beta_j(0) \cdot \beta_j(1)}}. \tag{14}$$

Proof. It can be proven by Lemma 8. ■

Proposition 10. *For* $\mathrm{Cov}(\boldsymbol{\beta}_{i,j})$ *and* $\mathrm{Corr}(\boldsymbol{\beta}_{i,j})$ *defined in* (9) *and* (12)*, we have*

$$|\mathrm{Cov}(\boldsymbol{\beta}_{i,j})| \leq \frac{1}{4}, \qquad |\mathrm{Corr}(\boldsymbol{\beta}_{i,j})| \leq 1.$$

Proof. The inequality on the LHS can be obtain by

$$|\mathrm{Cov}(\boldsymbol{\beta}_{i,j})| \leq \sqrt{\mathrm{Var}(Y_1) \cdot \mathrm{Var}(Y_2)} \overset{(a)}{=} \sqrt{\det(\boldsymbol{\beta}_i) \cdot \det(\boldsymbol{\beta}_j)} = \sqrt{\beta_i(0)(1 - \beta_i(0))\beta_j(0)(1 - \beta_j(0))} \leq \frac{1}{4}.$$

where at step (a) we have used Lemma 8. One can prove the inequality on the RHS by the Cauchy-Schwarz inequality. ■

Definition 11. *Suppose that* $\boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K})$*, we define two expressions:*

$$\mathrm{CovCHSH}(\boldsymbol{\beta}) := \mathrm{Cov}(\boldsymbol{\beta}_{1,2}) + \mathrm{Cov}(\boldsymbol{\beta}_{1,4}) + \mathrm{Cov}(\boldsymbol{\beta}_{3,2}) - \mathrm{Cov}(\boldsymbol{\beta}_{3,4}), \qquad \boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K}),$$

$$\mathrm{CorrCHSH}(\boldsymbol{\beta}) := \mathrm{Corr}(\boldsymbol{\beta}_{1,2}) + \mathrm{Corr}(\boldsymbol{\beta}_{1,4}) + \mathrm{Corr}(\boldsymbol{\beta}_{3,2}) - \mathrm{Corr}(\boldsymbol{\beta}_{3,4}), \quad 0 < \beta_i(0) < 1, \ i \in \mathcal{E}(\mathsf{N}_1), \ \boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K}), \tag{15}$$

where $\mathrm{Cov}(\boldsymbol{\beta}_{i,j})$ *and* $\mathrm{Corr}(\boldsymbol{\beta}_{i,j})$ *for* $\{i, j\} \in \mathcal{K}$ *are given in* (14).

The constraint $0 < \beta_i(0), \beta_j(0) < 1$ ensures that $\mathrm{Corr}(\boldsymbol{\beta}_{i,j})$ for $\{i, j\} \in \mathcal{K}$ is well-defined.

A. Properties for N_3

In this subsection, we prove inequalities with respect to $\mathrm{CorrCHSH}(\boldsymbol{\beta})$ for $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$. These inequalities genuinely are (nonlinear) Bell inequalities [26] in the usual sense. By definition, one can verify that

$$\mathcal{M}(\mathsf{N}_1) \subseteq \mathcal{M}(\mathsf{N}_3), \qquad \mathcal{M}(\mathsf{N}_2) \subseteq \mathcal{M}(\mathsf{N}_3),$$

which means that any property that holds for $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$ also holds for $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_1) \cup \mathcal{M}(\mathsf{N}_2)$. The set $\mathcal{M}(\mathsf{N}_3)$ contains the set of marginals that can be realized by joint PMFs for four random variables.

Theorem 12. *For any* $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$ *s.t.* $0 < \beta_i(0) < 1$ *for all* $i \in \mathcal{E}(\mathsf{N}_1)$*, we have*

$$|\mathrm{CorrCHSH}(\boldsymbol{\beta})| < 2\sqrt{2}.$$

Proof. See Appendix C. ■

The main idea in the proof of Theorem 12 can be used to verify whether a proposed bound for a function of PMFs is achievable. In particular, we prove it by contradiction. On the one hand, the set $\mathcal{M}(\mathsf{N}_3)$ defined in (5) consists of marginals for binary random variables. On the other hand, to have $\mathrm{CorrCHSH}(\boldsymbol{\beta}) = 2\sqrt{2}$ for $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$, the associated PMF has to be the joint PMF for ternary random variables. It is different from the idea in the proof of the upcoming Theorem 14.

Proposition 13. *There exists a* $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$ *s.t.* $\mathrm{CorrCHSH}(\boldsymbol{\beta}) = 5/2$.

Proof. See Appendix D. ■

Theorem 14. *For any* $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$ *s.t.* $0 < \beta_i(0) < 1$ *for all* $i \in \mathcal{E}(\mathsf{N}_1)$*, we have*

$$|\mathrm{CorrCHSH}(\boldsymbol{\beta})| \le \frac{5}{2}.$$

Proof. See Appendix E. ■

Theorem 14 proves the conjecture stated in [23]. The key idea of the proof is that we consider $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ instead of $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$. Suppose that we want to prove $\mathrm{CorrCHSH}(\boldsymbol{\beta}) \le 5/2$ for $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$ directly. For any $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$, the marginal $\boldsymbol{\beta}_{i,j}$ can be written as a convex combination of some joint PMF for $X_1, \ldots, X_4$, i.e., $\{p_{\mathsf{N}_3}(\boldsymbol{x})\}_{\boldsymbol{x}}$, which makes the expression of $\mathrm{CorrCHSH}(\boldsymbol{\beta})$ non-trivial. By considering a superset of $\mathcal{M}(\mathsf{N}_3)$, i.e., $\mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$, we can simplify $\mathrm{CorrCHSH}(\boldsymbol{\beta})$. We suspect that this idea can be generalized in the proof of other non-linear Bell inequalities.

Corollary 15. *For any* $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$*, we have*

$$|\mathrm{CovCHSH}(\boldsymbol{\beta})| \le \frac{4}{7}, \qquad \boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3).$$

Given that

$$\sum_{x_i,x_j} \beta_{i,j}(x_i,x_j) \cdot (-1)^{x_i+x_j} - \left(\sum_{x_i} \beta_i(x_i) \cdot (-1)^{x_i}\right) \cdot \left(\sum_{x_j} \beta_j(x_j) \cdot (-1)^{x_j}\right) = 4 \cdot \mathrm{Cov}(\boldsymbol{\beta}_{i,j}), \qquad \{i,j\} \in \mathcal{K},$$

we also prove the inequality:

$$\left|\sum_{i,j\in\mathcal{K}} (-1)^{\{i,j\}=\{3,4\}} \cdot \left(\sum_{x_i,x_j} \beta_{i,j}(x_i,x_j) \cdot (-1)^{x_i+x_j} - \left(\sum_{x_i} \beta_i(x_i) \cdot (-1)^{x_i}\right) \cdot \left(\sum_{x_j} \beta_j(x_j) \cdot (-1)^{x_j}\right)\right)\right| \le \frac{16}{7}.$$

Proof. The proof is similar to the proof of Theorem 14 and thus is omitted here. ■

B. Markov Chain N_2 in Fig. 2

In this subsection, we study the Markov chain N_2 in Fig. 2. In particular, we prove some inequalities w.r.t. the correlation coefficients for $(\boldsymbol{\beta}_{i,j})_{i,j}$.

Definition 16. *We make the following definitions for* N_2 *in Fig. 2.*

1) *The matrix* M_{X_1,X_2} *is defined to be*

$$M_{X_1,X_2} := \begin{pmatrix} M_{X_1,X_2}(0,0) & M_{X_1,X_2}(0,1) \\ M_{X_1,X_2}(1,0) & M_{X_1,X_2}(1,1) \end{pmatrix}, \tag{16}$$

where

$$M_{X_1,X_2}(x_1,x_2) \in \mathbb{R}_{\ge 0}, \qquad x_1,x_2 \in \mathcal{X}_e, \qquad \sum_{x_1,x_2\in\mathcal{X}_e} M_{X_1,X_2}(x_1,x_2) = 1$$

2) *For* $\{i,j\} \in \{\{1,4\},\{2,3\}\}$*, the matrix* $M_{X_j|X_i}$ *is defined to be*

$$M_{X_j|X_i} := \begin{pmatrix} M_{X_j|X_i}(0,0) & M_{X_j|X_i}(0,1) \\ M_{X_j|X_i}(1,0) & M_{X_j|X_i}(1,1) \end{pmatrix}, \tag{17}$$

where

$$M_{X_j|X_i}(x_j,x_i) \in \mathbb{R}_{\ge 0}, \qquad \sum_{x_j\in\mathcal{X}_e} M_{X_j|X_i}(x_j,x_i) = 1, \qquad x_i,\ x_j \in \mathcal{X}_e. \tag{18}$$

3) *For $i \in \mathcal{E}(\mathsf{N}_2)$, the diagonal matrix M_{X_i} is defined to be*

$$M_{X_i} := \begin{pmatrix} M_{X_i}(0) & 0 \\ 0 & M_{X_i}(1) \end{pmatrix},$$

where

$$M_{X_i}(x_i) := p_{\mathsf{N}_2,i}(x_i), \qquad x_i \in \mathcal{X}_e. \tag{19}$$

Proposition 17. *The set $\mathcal{M}(\mathsf{N}_2)$ equals*

$$\mathcal{M}(\mathsf{N}_2) = \{\boldsymbol{\beta}\,|\textit{there exists } \mathcal{F}(\mathsf{N}_2) \textit{ s.t. (20)–(22) hold}\},$$

where

$$\boldsymbol{\beta}_{1,2} = M_{X_1,X_2}, \qquad \boldsymbol{\beta}_{3,2} = M_{X_3|X_2} \cdot M_{X_2}, \tag{20}$$

$$\boldsymbol{\beta}_{1,4} = M_{X_1} \cdot \left(M_{X_4|X_1}\right)^{\mathsf{T}}, \qquad \boldsymbol{\beta}_{3,4} = M_{X_3|X_2} \cdot (M_{X_4|X_1} \cdot M_{X_1,X_2})^{\mathsf{T}}, \tag{21}$$

$$\boldsymbol{\beta}_i = M_{X_i}, \qquad i \in \mathcal{E}(\mathsf{N}_1). \tag{22}$$

Proof. It can proven directly by the definition of $\mathcal{M}(\mathsf{N}_2)$ in (5) and thus it is omitted here. ∎

Theorem 18. *For the Markov chain N_2 in Fig. 2, we have*

$$\mathrm{Corr}(\boldsymbol{\beta}_{3,4}) = \mathrm{Corr}(\boldsymbol{\beta}_{3,2}) \cdot \mathrm{Corr}(\boldsymbol{\beta}_{1,2}) \cdot \mathrm{Corr}(\boldsymbol{\beta}_{1,4}).$$

Proof. See [27, Corollary 19]. ∎

Corollary 19. *For the Markov chain N_2 in Fig. 2, it holds that*

$$|\mathrm{Corr}(\boldsymbol{\beta}_{3,4})| \leq |\mathrm{Corr}(\boldsymbol{\beta}_{1,2})| \leq 1.$$

Proof. It can be proven using Theorem 18 and Proposition 10. ∎

We prove another variation of the CHSH inequality [22] for the Markov chain N_2 in Fig. 2.

Proposition 20. *For the Markov chain N_2 in Fig. 2, we have*

$$|\mathrm{Cov}(\boldsymbol{\beta}_{1,2}) + \mathrm{Cov}(\boldsymbol{\beta}_{2,4}) + \mathrm{Cov}(\boldsymbol{\beta}_{1,3}) - \mathrm{Cov}(\boldsymbol{\beta}_{3,4})| \leq \frac{1}{2}. \tag{23}$$

$$|\mathrm{Corr}(\boldsymbol{\beta}_{1,2}) + \mathrm{Corr}(\boldsymbol{\beta}_{2,4}) + \mathrm{Corr}(\boldsymbol{\beta}_{1,3}) - \mathrm{Corr}(\boldsymbol{\beta}_{3,4})| \leq 2, \tag{24}$$

Proof. Concerning the inequality (23), we have

$$\mathrm{Cov}(\boldsymbol{\beta}_{1,3}) \overset{(a)}{=} \det(M_{X_1,X_2}) \cdot \det\!\big(M_{X_3|X_2}\big), \qquad \mathrm{Cov}(\boldsymbol{\beta}_{2,4}) \overset{(a)}{=} \det(M_{X_1,X_2}) \cdot \det\!\big(M_{X_4|X_1}\big), \qquad \mathrm{Cov}(\boldsymbol{\beta}_{3,4}) \overset{(a)}{=} \det(M_{X_1,X_2}) \cdot \det\!\big(M_{X}$$

where at step (a) we have used (20) and (21). By the property of $M_{X_j|X_i}$ for $\{i,j\} \in \{\{1,4\},\{2,3\}\}$ in (18), we have

$$\det\!\big(M_{X_j|X_i}\big) = M_{X_j|X_i}(0,0) \cdot M_{X_j|X_i}(1,1) - (1 - M_{X_j|X_i}(0,0)) \cdot (1 - M_{X_j|X_i}(1,1)) = M_{X_j|X_i}(0,0) + M_{X_j|X_i}(1,1) - 1 \leq 1.$$

Then we have

$$\begin{aligned}
&\Big|\mathrm{Cov}(\boldsymbol{\beta}_{1,2}) \cdot \Big(1 + \mathrm{Cov}(\boldsymbol{\beta}_{1,4}) + \mathrm{Cov}(\boldsymbol{\beta}_{2,3}) \cdot \Big(1 - \mathrm{Cov}(\boldsymbol{\beta}_{1,4})\Big)\Big)\Big| \\
&\overset{(a)}{\leq} \frac{1}{4}\Big|1 + \mathrm{Cov}(\boldsymbol{\beta}_{1,4})\Big| + \frac{1}{4}\Big|1 - \mathrm{Cov}(\boldsymbol{\beta}_{1,4})\Big| \\
&\leq \frac{1}{2}.
\end{aligned}$$

where at step (a) we have used Proposition 10.

Now we turn to prove the inequality (24). Similar to the proof of [27, Corollary 19], the Markov property shown in Fig. 2 implies

$$\mathrm{Corr}(\boldsymbol{\beta}_{1,3}) = \mathrm{Corr}(\boldsymbol{\beta}_{1,2}) \cdot \mathrm{Corr}(\boldsymbol{\beta}_{2,3}), \qquad \mathrm{Corr}(\boldsymbol{\beta}_{2,4}) = \mathrm{Corr}(\boldsymbol{\beta}_{1,2}) \cdot \mathrm{Corr}(\boldsymbol{\beta}_{1,4}).$$

Then we have

$$\begin{aligned}
&\left|\mathrm{Corr}(\boldsymbol{\beta}_{1,2}) \cdot \Big(1 + \mathrm{Corr}(\boldsymbol{\beta}_{1,4}) + \mathrm{Corr}(\boldsymbol{\beta}_{2,3}) \cdot \Big(1 - \mathrm{Corr}(\boldsymbol{\beta}_{1,4})\Big)\Big)\right| \\
&\overset{(a)}{\leq} \left|1 + \mathrm{Corr}(\boldsymbol{\beta}_{1,4})\right| + \left|1 - \mathrm{Corr}(\boldsymbol{\beta}_{1,4})\right| \\
&\leq 2.
\end{aligned}$$

where at step (a) we have used Proposition 10.

■

V. Quantum-Probability Normal Factor Graphs (Q-NFGs)

This section considers a quantum system represented by the Q-NFG N_4 in Fig. 4. Such Q-NFGs have been discussed thoroughly in [7], [8]. For each degree-2 function node in NFGs, we can associate it with a matrix. In quantum information processing systems, the degree of function nodes is usually 2. If a function node has degree more than 2, we can associated it with a collection of matrices by setting extra variables to be the indices of matrices (see, e.g., the collections of matrices $\{A_{i,x_i}\}_{i,x_i}$ and $\{B_{j,x_j}\}_{j,x_j}$ in Fig. 5). In Figs. 4, and 5, the row index of a matrix is marked by a ciliation. Recall that the definitions of QMFs and the associated classicable variables are given in Definitions 1 and 2. We present the details of N_4 in the following for completeness.

Definition 21. *For N_4, we make the following definitions.*

1) *The set of edges is defined to be the set $\mathcal{E}(\mathsf{N}_4) := \mathcal{E}(\mathsf{N}_1)$.*
2) *The alphabet for N_4 is $\mathcal{X}(\mathsf{N}_4) := \prod_{i \in \mathcal{E}(\mathsf{N}_1)} \mathcal{X}_i^2$, where $\mathcal{X}_i^2 := \{0,1\}^2$ is the alphabet for the variable $\tilde{x}_i := (x_i, x_i')$.*
3) *An assignment $\tilde{\boldsymbol{x}} := (\tilde{x}_i)_i \in \mathcal{X}(\mathsf{N}_4)$ is called a configuration of N_4.*
4) *The matrix ρ with row index (x_1, x_2) and column index (x_1', x_2') is defined to be*

$$\rho := \begin{pmatrix} \rho_{1,1} & \rho_{1,2} & \rho_{1,3} & \rho_{1,4} \\ \rho_{2,1} & \rho_{2,2} & \rho_{2,3} & \rho_{2,4} \\ \rho_{3,1} & \rho_{3,2} & \rho_{3,3} & \rho_{3,4} \\ \rho_{4,1} & \rho_{4,2} & \rho_{4,3} & \rho_{4,4} \end{pmatrix},$$

where the first row of ρ is indexed by $(0,0)$, the second row of ρ is indexed by $(0,1)$, the third row of ρ is indexed by $(1,0)$, and the forth row of ρ is indexed by $(1,1)$. The columns of ρ are indexed similarly. We require that the matrix ρ is a Hermitian, positive semi-definite (PSD) matrix with trace 1, i.e., a density matrix, which means that

$$\rho = \begin{pmatrix} \rho_{1,1} & \overline{\rho_{2,1}} & \overline{\rho_{3,1}} & \overline{\rho_{4,1}} \\ \rho_{2,1} & \rho_{2,2} & \overline{\rho_{3,2}} & \overline{\rho_{4,2}} \\ \rho_{3,1} & \rho_{3,2} & \rho_{3,3} & \overline{\rho_{4,3}} \\ \rho_{4,1} & \rho_{4,2} & \rho_{4,3} & \rho_{4,4} \end{pmatrix}, \qquad \rho_{i,j} = \overline{\rho_{j,i}}, \qquad i,j \in []. \tag{25}$$

5) *For $i \in \{1,2\}$, the 2-by-2 unitary matrix U_i is defined to be*

$$U_i := \begin{pmatrix} U_i(0,0) & U_i(0,1) \\ -\exp(\iota\varphi_i)\overline{U_i(0,1)} & \exp(\iota\varphi_i)\overline{U_i(0,0)} \end{pmatrix}, \quad |U_i(0,0)|^2 + |U_i(0,1)|^2 = 1,\ U_i(0,0),\ U_i(0,1) \in \mathbb{C},\ \varphi_i \in \mathbb{R},$$

where ι is the imaginary unit. In particular, the row indices of U_1 and U_2 are x_3 and x_4, respectively; the column indices of U_1 and U_2 are x_1 and x_2, respectively.

6) *The global function for* N_4 *is defined to be*

$$g_{\mathsf{N}_4}(\tilde{\boldsymbol{x}}) := \rho\big((x_1,x_2),(x_1',x_2')\big)\cdot U_1(x_3,x_1)\cdot\overline{U_1(x_3',x_1')}\cdot U_2(x_4,x_2)\cdot\overline{U_2(x_4',x_2')}\cdot[x_3=x_3',x_4=x_4'].$$

7) *The partition function of* N_4 *is defined to be*

$$Z(\mathsf{N}_4) := \sum_{\tilde{\boldsymbol{x}}} g_{\mathsf{N}_4}(\tilde{\boldsymbol{x}}), \tag{26}$$

where $\sum_{\tilde{\boldsymbol{x}}}$ *denotes* $\sum_{\tilde{\boldsymbol{x}}\in\mathcal{X}(\mathsf{N}_4)}$.

8) *The QMF*[8] *induce on* N_4 *is*

$$q_{\mathsf{N}_4}(\tilde{\boldsymbol{x}}) := \frac{g_{\mathsf{N}_4}(\tilde{\boldsymbol{x}})}{Z(\mathsf{N}_4)}.$$

9) *Let* $\mathcal{I}$ *be a subset of* $\{1,\ldots,m\}$ *and let* $\mathcal{I}^{\mathrm{c}} := \{1,\ldots,m\}\setminus\mathcal{I}$ *be its complement. The marginal* $q_{\mathcal{I}}(\tilde{\boldsymbol{x}}_{\mathcal{I}})$ *is defined to be*

$$q_{\mathcal{I}}(\tilde{\boldsymbol{x}}_{\mathcal{I}}) := \sum_{\tilde{\boldsymbol{x}}_{\mathcal{I}^{\mathrm{c}}}} q_{\mathsf{N}_4}(\tilde{\boldsymbol{x}}), \qquad \tilde{\boldsymbol{x}}_{\mathcal{I}} \in \prod_{i=1}^{|\mathcal{I}|}\mathcal{X}_i^2.$$

Proposition 22. *The partition function* $Z(\mathsf{N}_4)$ *defined in* (26) *satisfies* $Z(\mathsf{N}_4)=1$. *Therefore, we have* $q_{\mathsf{N}_4}(\tilde{\boldsymbol{x}}) = g_{\mathsf{N}_4}(\tilde{\boldsymbol{x}})$.

Proof. This can be proven directly using the definitions of $Z(\mathsf{N}_4)$ and $g_{\mathsf{N}_4}(\tilde{\boldsymbol{x}})$. ■

Proposition 23. *For any* $\{i,j\}\in\mathcal{K}$ *and* $\tilde{x}_i,\tilde{x}_j\in\mathcal{X}_e^2$, *the marginals* $q_{i,j}(\tilde{x}_i,\tilde{x}_j)$ *and* $q_i(\tilde{x}_i)$ *are non-negative real numbers.*

Proof. This can be proven directly using the definitions of $\mathcal{K}$ and $q_{\mathsf{N}_4}(\tilde{\boldsymbol{x}})$. ■

Then we define the set of marginals of N_4 that can be realized by varying ρ, U_1, and U_2 in N_4.

Definition 24. *With* $\tilde{0} := (0,0)$, $\tilde{1} := (1,1)$, *the matrices* $\boldsymbol{q}_{i,j}$ *and* $\boldsymbol{q}_i$ *induced by* q_{N_4} *are*

$$\boldsymbol{q}_{i,j} := \begin{pmatrix} q_{i,j}(\tilde{0},\tilde{0}) & q_{i,j}(\tilde{0},\tilde{1}) \\ q_{i,j}(\tilde{1},\tilde{0}) & q_{i,j}(\tilde{1},\tilde{1}) \end{pmatrix}, \qquad \boldsymbol{q}_i := \begin{pmatrix} q_i(\tilde{0}) & 0 \\ 0 & q_i(\tilde{1}) \end{pmatrix}.$$

The set of realizable marginals $\mathcal{M}(\mathsf{N}_4)$ *is defined to be the set*

$$\mathcal{M}(\mathsf{N}_4) := \left\{\boldsymbol{\beta}\,\middle|\,\text{There exists } (\boldsymbol{q}_{i,j})_{i,j} \text{ and } (\boldsymbol{q}_i)_i \text{ for } \mathsf{N}_4 \text{ s.t. } \boldsymbol{\beta}_{i,j} = \boldsymbol{q}_{i,j},\ \boldsymbol{\beta}_i = \boldsymbol{q}_i,\ \{i,j\}\in\mathcal{K}\right\}.$$

For any $\boldsymbol{\beta}\in\mathcal{M}(\mathsf{N}_4)$, there exist ρ, U_1, and U_2 s.t.

$$\beta_{i,j}(x_i,x_j) = \mathrm{Tr}\big((A_{i,x_i}\otimes B_{j,x_j})\cdot\rho\cdot(A_{i,x_i}\otimes B_{j,x_j})^{\mathsf{H}}\big), \qquad x_i,x_j\in\mathcal{X}_e, \{i,j\}\in\mathcal{K}, \tag{27}$$

where

$$A_{i,x_i} := E_{x_i}\cdot U_1^{[i=3]}, \qquad B_{j,x_j} := E_{x_j}\cdot U_2^{[j=4]}, \qquad \{i,j\}\in\mathcal{K}, \tag{28}$$

$$E_0 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad E_1 := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{29}$$

Note that the set $\{E_{x_i}\}_{x_i}$ is a set of projection matrices, i.e.,

$$E_{x_i}\cdot E_{x_i}^{\mathsf{H}} = E_{x_i}^2 = E_{x_i}, \qquad x_i\in\mathcal{X}_e, \tag{30}$$

and also denotes the measurement of a single qubit in the computational basis. Then we have

$$\sum_{x_i\in\mathcal{X}_e} A_{i,x_i}^{\mathsf{H}}\cdot A_{i,x_i} = \sum_{x_j\in\mathcal{X}_e} B_{j,x_j}^{\mathsf{H}}\cdot B_{j,x_j} = I, \quad \{i,j\}\in\mathcal{K}.$$

[8]The general definition of QMFs is given in [8, Definition 1].

which means that both $\{A_{i,x_i}\}_{x_i}$ and $\{B_{j,x_j}\}_{x_j}$ are sets of measurement matrices with outcomes x_i and x_j, respectively. It holds that

$$\begin{aligned} &A_{i,x_i} \cdot A_{i,x_i}^{\mathsf{H}} = E_{x_i}, \qquad B_{j,x_j} \cdot B_{j,x_j}^{\mathsf{H}} = E_{x_j}, \qquad x_i, x_j \in \mathcal{X}_e,\ \{i,j\} \in \mathcal{K}, \\ &A_{i,x_i}^{\mathsf{H}} \cdot A_{i,x_i} \overset{(a)}{=} \Big(U_1^{[i=3]}\Big)^{\mathsf{H}} \cdot E_{x_i} \cdot U_1^{[i=3]} = \Big(A_{i,x_i}^{\mathsf{H}} \cdot A_{i,x_i}\Big)^{\mathsf{H}}, \qquad x_i \in \mathcal{X}_e,\ i \in \{1,3\} \\ &B_{j,x_j}^{\mathsf{H}} \cdot B_{j,x_j} \overset{(a)}{=} \Big(U_2^{[j=4]}\Big)^{\mathsf{H}} \cdot E_{x_J} \cdot U_2^{[j=4]} = \Big(B_{j,x_j}^{\mathsf{H}} \cdot B_{j,x_j}\Big)^{\mathsf{H}}, \qquad x_j \in \mathcal{X}_e,\ j \in \{2,4\}, \end{aligned} \tag{31}$$

where at step (a) we have used (30). Then we have

$$A_{i,x_i}^{\mathsf{H}} \cdot A_{i,x_i} \cdot (A_{i,x_i}^{\mathsf{H}} \cdot A_{i,x_i})^{\mathsf{H}} \overset{(a)}{=} A_{i,x_i}^{\mathsf{H}} \cdot E_{x_i} \cdot A_{i,x_i} \overset{(b)}{=} A_{i,x_i}^{\mathsf{H}} \cdot A_{i,x_i}, \tag{32}$$

$$B_{j,x_j}^{\mathsf{H}} \cdot B_{j,x_j} \cdot (B_{j,x_j}^{\mathsf{H}} \cdot B_{j,x_j})^{\mathsf{H}} \overset{(a)}{=} B_{j,x_j}^{\mathsf{H}} \cdot E_{x_j} \cdot B_{j,x_j} \overset{(b)}{=} B_{j,x_j}^{\mathsf{H}} \cdot B_{j,x_j}, \tag{33}$$

where at step (a) we have used (31) and at step (b) we have used (30). The above equations imply that $\{A_{i,x_i} \cdot A_{i,x_i}^{\mathsf{H}}\}_{x_i}$, $\{A_{i,x_i}^{\mathsf{H}} \cdot A_{i,x_i}\}_{x_i}$, $\{B_{j,x_j} \cdot B_{j,x_j}^{\mathsf{H}}\}_{x_j}$, and $\{B_{j,x_j}^{\mathsf{H}} \cdot B_{j,x_j}\}_{x_j}$ are sets of projection matrices for all $\{i,j\} \in \mathcal{K}$.

Fig. 5 illustrates (27). Namely, after closing the dashed box in Fig. 5, i.e., summing over the variables inside the box, we obtain (27). Therefore, the marginals $\boldsymbol{\beta}_{i,j}$ in Fig. 5 represent the probabilities of the outcomes in the following experiment:

- Alice and Bob share two particles whose density matrix is denoted by ρ. They can do some processing and measurements on their own qubits.
- Alice's i-th measurement on her qubit is described by the set of measurement matrices $\{A_{i,x_i}\}_{x_i}$ for $i \in \{1,3\}$.
- Alice does not know which measurement she shall perform. Instead, when she receives the particle, she uses some random method (e.g., a coin flip) to decide which measurement to perform.
- If $i = 1$, only outcome x_1 is accessible for Alice. If $i = 3$, only outcome x_3 is accessible for Alice.
- Bob's j-th measurement on his qubit is described by the set of measurement matrices $\{B_{j,x_j}\}_{x_j}$ for $j \in \{2,4\}$.
- Similarly, Bob measures his qubit based on some random method.
- When Alice chooses i and Bob chooses j, the probability of getting outcome x_i, x_j is $\beta_{i,j}(x_i, x_j)$.

The setup of this experiment is similar to that in Hardy's paradox and Bell's game, which indicates that we can realize Hardy's paradox and Bell's game in Fig. 5, or equivalently, Fig. 4.

Proposition 25. *For any $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4)$, there exist ρ, U_1, and U_2 s.t.*

$$\boldsymbol{\beta}_{1,2} = \mathrm{diag}(\rho), \qquad \boldsymbol{\beta}_{3,2} = \mathrm{diag}\big((U_1 \otimes I) \cdot \rho \cdot (U_1 \otimes I)^{\mathsf{H}}\big), \tag{34}$$

$$\boldsymbol{\beta}_{1,4} = \mathrm{diag}\big((I \otimes U_2) \cdot \rho \cdot (I \otimes U_2)^{\mathsf{H}}\big), \qquad \boldsymbol{\beta}_{3,4} = \mathrm{diag}\big((U_1 \otimes U_2) \cdot \rho \cdot (U_1 \otimes U_2)^{\mathsf{H}}\big). \tag{35}$$

Proof. This can be proven directly using the definitions of ρ, U_1, and U_2. ■

Equation (2) and Proposition 23 imply $\mathcal{M}(\mathsf{N}_4) \subseteq \mathcal{LM}(\mathcal{K})$. Thus for any $\{i,j\} \in \mathcal{K}$, the functions $\mathrm{Cov}(\boldsymbol{\beta}_{i,j})$ and $\mathrm{Corr}(\boldsymbol{\beta}_{i,j})$ defined in Definition 11 are well-defined for $\boldsymbol{\beta}_{i,j}$ s.t. the associated $\boldsymbol{\beta}$ is in $\mathcal{M}(\mathsf{N}_4)$.

Proposition 26. *There exists a $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4)$ s.t. $|\mathrm{Corr}(\boldsymbol{\beta}_{3,4})| > |\mathrm{Corr}(\boldsymbol{\beta}_{1,2})|$.*

Proof. Let us consider the following setting.

$$U_1 = H, \qquad U_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \rho = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 & -1 \end{pmatrix}^{\mathsf{T}}.$$

The set of matrices $(\boldsymbol{\beta}_{i,j})_{i,j}$ obtained by in (34) and (35) satisfies

$$\boldsymbol{\beta}_{1,2} = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,2} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then the correlation coefficients are

$$\mathrm{Corr}(\boldsymbol{\beta}_{1,2}) = 0, \qquad \mathrm{Corr}(\boldsymbol{\beta}_{3,4}) = 1.$$

■

Proposition 27. *For the set $\tilde{\mathcal{X}}_e$ with arbitrary size, it holds that*

$$|\mathrm{CorrCHSH}(\boldsymbol{\beta})| \leq 2\sqrt{2}, \qquad 0 < \beta_i(0), \beta_j(0) < 1,\ \{i,j\} \in \mathcal{K},\ \boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4), \tag{36}$$

$$|\mathrm{CovCHSH}(\boldsymbol{\beta})| \leq \frac{\sqrt{2}}{2}, \qquad \boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4), \tag{37}$$

which are PCC-based and covariance-based Tsirelon bounds, respctively.

Proof. The proof of (36) can be found in [23, Appendix B]. For the proof of (37), see Appendix F. ■

For comparison, we note that Tsirelson's bound is

$$\Big| \sum_{\{i,j\} \in \mathcal{K}} (-1)^{[i=3,j=4]} \cdot \Big(\beta_{i,j}(0,0) + \beta_{i,j}(1,1) - \beta_{i,j}(0,1) - \beta_{i,j}(1,0) \Big) \Big| \leq 2\sqrt{2}, \qquad \boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4),$$

which is a linear inequality w.r.t. $\boldsymbol{\beta}$.

Let us make some comments on Proposition 27:

- The proof approaches of Proposition 27 work for $\tilde{\mathcal{X}}_e$ with arbitrary size, i.e., arbitrary finite-dimensional quantum systems.
- For any $\{i,j\} \in \mathcal{K}$, the matrices $\sum_{x_i} x_i \cdot A_{i,x_i}^{\mathsf{H}} \cdot A_{i,x_i}$ and $\sum_{x_i} x_j \cdot B_{i,x_i}^{\mathsf{H}} \cdot B_{i,x_i}$ represent the observables with eigenvalues 0 and 1, which are different from the observables in [23, Appendix B]. In [23], they considered more general observables with eigenvalues in $[-1, 1]$. This is the reason why we obtain a stricter covariance-based Tsirelson's bound, compared with the bound derived in [23, Appendix B].
- We suspect that the proof approach for Proposition 27 is applicable for proving the maximum quantum violations of other Bell inequalities for both covariance and PCCs.

Proposition 28. *It holds that*

$$\begin{aligned} &|\mathrm{Corr}(\boldsymbol{\beta}_{1,2}) \cdot \mathrm{Corr}(\boldsymbol{\beta}_{1,4}) - \mathrm{Corr}(\boldsymbol{\beta}_{3,2}) \cdot \mathrm{Corr}(\boldsymbol{\beta}_{3,4})| \\ &\leq \sqrt{1 - \big(\mathrm{Corr}(\boldsymbol{\beta}_{1,2})\big)^2} \cdot \sqrt{1 - (\mathrm{Corr}(\boldsymbol{\beta}_{1,4}))^2} + \sqrt{1 - (\mathrm{Corr}(\boldsymbol{\beta}_{3,2}))^2} \cdot \sqrt{1 - (\mathrm{Corr}(\boldsymbol{\beta}_{3,4}))^2}, \qquad \boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4), \end{aligned} \tag{38}$$

$$\begin{aligned} &16 \cdot |\mathrm{Cov}(\boldsymbol{\beta}_{1,2}) \cdot \mathrm{Cov}(\boldsymbol{\beta}_{1,4}) - \mathrm{Cov}(\boldsymbol{\beta}_{3,2}) \cdot \mathrm{Cov}(\boldsymbol{\beta}_{3,4})| \\ &\leq \sqrt{1 - (4\mathrm{Cov}(\boldsymbol{\beta}_{1,2}))^2} \cdot \sqrt{1 - (4\mathrm{Cov}(\boldsymbol{\beta}_{1,4}))^2} + \sqrt{1 - (4\mathrm{Cov}(\boldsymbol{\beta}_{3,2}))^2} \cdot \sqrt{1 - (4\mathrm{Cov}(\boldsymbol{\beta}_{3,4}))^2}, \qquad \boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4). \end{aligned} \tag{39}$$

Proof. We prove (38) first. By the definitions of $\check{\boldsymbol{\alpha}}_{\ell,i}$ in (87) and $\check{\boldsymbol{\gamma}}_{\ell,j}$ in (88) for $k \in \mathcal{E}(\mathsf{N}_1)$ and $\{i,j\} \in \mathcal{K}$, we have

$$\begin{aligned} &\sum_{\ell} \begin{pmatrix} \check{\boldsymbol{\alpha}}_{\ell,1} & \check{\boldsymbol{\alpha}}_{\ell,3} & \check{\boldsymbol{\gamma}}_{\ell,2} & \check{\boldsymbol{\gamma}}_{\ell,4} \end{pmatrix}^{\mathsf{H}} \cdot \begin{pmatrix} \check{\boldsymbol{\alpha}}_{\ell,1} & \check{\boldsymbol{\alpha}}_{\ell,3} & \check{\boldsymbol{\gamma}}_{\ell,2} & \check{\boldsymbol{\gamma}}_{\ell,4} \end{pmatrix} \\ &\overset{(a)}{=} \begin{pmatrix} 1 & \sum_{\ell} (\check{\boldsymbol{\alpha}}_{1,\ell})^{\mathsf{H}} \cdot \check{\boldsymbol{\alpha}}_{3,\ell} & \mathrm{Corr}(\boldsymbol{\beta}_{1,2}) & \mathrm{Corr}(\boldsymbol{\beta}_{1,4}) \\ \sum_{\ell} (\check{\boldsymbol{\alpha}}_{3,\ell})^{\mathsf{H}} \cdot \check{\boldsymbol{\alpha}}_{1,\ell} & 1 & \mathrm{Corr}(\boldsymbol{\beta}_{3,2}) & \mathrm{Corr}(\boldsymbol{\beta}_{3,4}) \\ \mathrm{Corr}(\boldsymbol{\beta}_{1,2}) & \mathrm{Corr}(\boldsymbol{\beta}_{3,2}) & 1 & \sum_{\ell} (\check{\boldsymbol{\gamma}}_{2,\ell})^{\mathsf{H}} \cdot \check{\boldsymbol{\gamma}}_{4,\ell} \\ \mathrm{Corr}(\boldsymbol{\beta}_{1,4}) & \mathrm{Corr}(\boldsymbol{\beta}_{3,4}) & \sum_{\ell} (\check{\boldsymbol{\gamma}}_{4,\ell})^{\mathsf{H}} \cdot \check{\boldsymbol{\gamma}}_{2,\ell} & 1 \end{pmatrix} \succeq 0. \end{aligned}$$

where at step (a) we have used (90). The author in [28] proved that the positive semi-definiteness of the above matrix implies (38).

The proof of (39) is similar and thus is omitted here. Note that to prove (39), we need Proposition 10, i.e. $|4\mathrm{Cov}(\boldsymbol{\beta}_{i,j})| \leq 1$ for all $\{i,j\} \in \mathcal{K}$.

■

Corollary 29. *For $\{i,j\} \in \mathcal{K}$, we define $\theta_{i,j} \in [0,\pi]$ s.t. $\cos(\theta_{i,j}) = \mathrm{Corr}(\boldsymbol{\beta}_{i,j})$ and $\sin(\theta_{i,j}) \geq 0$. We have*

$$\cos(\theta_{1,2} + \theta_{1,4}) \leq \cos(\theta_{3,2} - \theta_{3,4}), \qquad \cos(\theta_{3,2} + \theta_{3,4}) \leq \cos(\theta_{1,2} - \theta_{1,4}).$$

Proof. It can be proven directly by Proposition 28. ■

Example 30. ***Hardy's Paradox*** *Hardy's paradox [12] states that there is no $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$ s.t.*

$$\beta_{1,2}(1,1) = 0, \qquad \beta_{3,2}(1,0) = 0, \qquad \beta_{1,4}(0,1) = 0, \qquad \beta_{3,4}(1,1) > 0. \tag{40}$$

We consider the following matrices for the Q-NFG N_4 in Fig. 4:

$$H := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad U_1 = H, \qquad U_2 = H, \qquad \rho = \frac{1}{3} \cdot \begin{pmatrix} 1 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 & 0 \end{pmatrix}^{\mathsf{T}},$$

where H is called the Hadamard gate. The collection of matrices $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4)$ obtained via (34) *and* (35) *satisfies*

$$\boldsymbol{\beta}_{1,2} = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \frac{1}{6} \begin{pmatrix} 4 & 0 \\ 1 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,2} = \frac{1}{6} \begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \frac{1}{12} \begin{pmatrix} 9 & 1 \\ 1 & 1 \end{pmatrix}.$$

The above marginals satisfy Hardy's paradox in (40).

A. Bell's Game

In this subsection, we illustrate Bell's game [14] in terms of the classicable variables of the QMF $q_{\mathsf{N}_4}(\tilde{\boldsymbol{x}})$. One way to win Bell's game is to let $\boldsymbol{\beta}$ satisfy

$$\boldsymbol{\beta}_{1,2} = \boldsymbol{\beta}_{1,4} = \boldsymbol{\beta}_{3,2} = \frac{1}{8} \begin{pmatrix} 2+\sqrt{2} & 2-\sqrt{2} \\ 2-\sqrt{2} & 2+\sqrt{2} \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \frac{1}{8} \begin{pmatrix} 2-\sqrt{2} & 2+\sqrt{2} \\ 2+\sqrt{2} & 2-\sqrt{2} \end{pmatrix}, \tag{41}$$

By $\mathrm{Corr}(\boldsymbol{\beta}_{i,j})$ in (14) for $\{i,j\} \in \mathcal{K}$, we have

$$\mathrm{Corr}(\boldsymbol{\beta}_{1,4}) = \mathrm{Corr}(\boldsymbol{\beta}_{1,2}) = \mathrm{Corr}(\boldsymbol{\beta}_{3,2}) = \frac{\sqrt{2}}{2}, \qquad \mathrm{Corr}(\boldsymbol{\beta}_{3,4}) = -\frac{\sqrt{2}}{2},$$

which implies

$$\mathrm{CorrCHSH}(\boldsymbol{\beta}) = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} = 2\sqrt{2}.$$

Proposition 31. *The collection of matrices $\boldsymbol{\beta}$ satisfying* (41) *is in $\mathcal{M}(\mathsf{N}_4)$.*

Proof. See Appendix G. ■

B. The Double-Edge Normal Factor Graph (DE-NFG) N_5

Note that the upper half and the lower half of N_4 in Fig. 4 are mirror images of each other, which makes the factor graph redundant in some sense. This redundancy is eliminated in a more compact factor graph namely DE-NFG [29]. In this subsection, we present the details of the DE-NFG N_5 in Fig. 6, which is defined based on N_4.

Definition 32. *Based on N_4 in Fig. 4, we make the following definitions for the DE-NFG N_5 in Fig. 6.*

- *The matrix ρ_{L} with row index $\tilde{x}_1$ and column index $\tilde{x}_2$ is defined to be the Liouville-superoperator representation of ρ, which means*

$$\rho_{\mathrm{L}}(\tilde{x}_1, \tilde{x}_2) := \rho\big((x_1, x_2), (x_1', x_2')\big), \qquad x_1, x_1', x_2, x_2' \in \mathcal{X}_e.$$

Then we have

$$\rho_{\mathrm{L}} = \begin{pmatrix} \rho_{1,1} & \rho_{1,2} & \rho_{2,1} & \rho_{2,2} \\ \rho_{1,3} & \rho_{1,4} & \rho_{2,3} & \rho_{2,4} \\ \rho_{3,1} & \rho_{3,2} & \rho_{4,1} & \rho_{4,2} \\ \rho_{3,3} & \rho_{3,4} & \rho_{4,3} & \rho_{4,4} \end{pmatrix} \overset{(a)}{=} \begin{pmatrix} \rho_{1,1} & \overline{\rho_{2,1}} & \rho_{2,1} & \rho_{2,2} \\ \overline{\rho_{3,1}} & \overline{\rho_{4,1}} & \overline{\rho_{3,2}} & \overline{\rho_{4,2}} \\ \rho_{3,1} & \rho_{3,2} & \rho_{4,1} & \rho_{4,2} \\ \rho_{3,3} & \overline{\rho_{4,3}} & \rho_{4,3} & \rho_{4,4} \end{pmatrix},$$

where at step (a) we have used (25)*.*

- *The entry in the matrix $\tilde{U}_i$ with row index $\tilde{x}_{i_1}$ and column index $\tilde{x}_{i_2}$ is defined to be*

$$\tilde{U}_i(\tilde{x}_{i_1}, \tilde{x}_{i_2}) := U_1(x_{i_1}, x_{i_2}) \cdot \overline{U_1(x'_{i_1}, x'_{i_2})}, \qquad \tilde{x}_{i_1}, \tilde{x}_{i_2} \in \tilde{\mathcal{X}}_e,\ i \in \{1,2\}. \tag{42}$$

where $i_1 = 3$ and $i_2 = 1$ when $i_1 = 4$ and $i_2 = 2$. The matrix $\tilde{U}_i$ can be written as

$$\tilde{U}_i = U_i \otimes \overline{U_i}.$$

In N_4 and N_5, we require $x_{i_2} = x'_{i_2}$. By item 5) in Definition 21, the matrix $\tilde{U}_i$ can be written as

$$\tilde{U}_i = \begin{pmatrix} |U_i(0,0)|^2 & U_i(0,0)\cdot\overline{U_i(0,1)} & U_i(0,1)\cdot\overline{U_i(0,0)} & |U_i(0,1)|^2 \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ |U_i(0,1)|^2 & -U_i(0,0)\cdot\overline{U_i(0,1)} & -U_i(0,1)\cdot\overline{U_i(0,0)} & |U_i(0,0)|^2 \end{pmatrix}, \tag{43}$$

where the entries marked $\times$ are irrelevant.

- *The set of realizable marginals $\mathcal{M}(\mathsf{N}_5)$ is defined to be the set*

$$\mathcal{M}(\mathsf{N}_5) := \left\{ \boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K}) \;\middle|\; \textit{there exist } \rho_{\mathrm{L}},\ \tilde{U}_1, \textit{ and } \tilde{U}_2 \textit{ s.t. (44)–(47) hold} \right\},$$

where

$$\boldsymbol{\beta}_{1,2} = \begin{pmatrix} \rho_{\mathrm{L}}(\tilde{0},\tilde{0}) & \rho_{\mathrm{L}}(\tilde{0},\tilde{1}) \\ \rho_{\mathrm{L}}(\tilde{1},\tilde{0}) & \rho_{\mathrm{L}}(\tilde{1},\tilde{1}) \end{pmatrix}, \tag{44}$$

$$\boldsymbol{\beta}_{3,2} = \begin{pmatrix} \tilde{U}_1(\tilde{0},:)\cdot\rho_{\mathrm{L}}(:,\tilde{0}) & \tilde{U}_1(\tilde{0},:)\cdot\rho_{\mathrm{L}}(:,\tilde{1}) \\ \tilde{U}_1(\tilde{1},:)\cdot\rho_{\mathrm{L}}(:,\tilde{0}) & \tilde{U}_1(\tilde{1},:)\cdot\rho_{\mathrm{L}}(:,\tilde{1}) \end{pmatrix}, \tag{45}$$

$$\boldsymbol{\beta}_{1,4} = \begin{pmatrix} \rho_{\mathrm{L}}(\tilde{0},:)\cdot\big(\tilde{U}_2(\tilde{0},)\big)^{\mathsf{T}} & \rho_{\mathrm{L}}(\tilde{0},:)\cdot\big(\tilde{U}_2(:,\tilde{1})\big)^{\mathsf{T}} \\ \rho_{\mathrm{L}}(\tilde{1},:)\cdot\big(\tilde{U}_2(:,\tilde{0})\big)^{\mathsf{T}} & \rho_{\mathrm{L}}(\tilde{1},:)\cdot\big(\tilde{U}_2(:,\tilde{1})\big)^{\mathsf{T}} \end{pmatrix}, \tag{46}$$

$$\boldsymbol{\beta}_{3,4} = \begin{pmatrix} \tilde{U}_1(\tilde{0},:)\cdot\rho_{\mathrm{L}}\cdot\big(\tilde{U}_2(:,\tilde{0})\big)^{\mathsf{T}} & \tilde{U}_1(\tilde{0},:)\cdot\rho_{\mathrm{L}}\cdot\big(\tilde{U}_2(:,\tilde{1})\big)^{\mathsf{T}} \\ \tilde{U}_1(\tilde{1},:)\cdot\rho_{\mathrm{L}}\cdot\big(\tilde{U}_2(:,\tilde{0})\big)^{\mathsf{T}} & \tilde{U}_1(\tilde{1},:)\cdot\rho_{\mathrm{L}}\cdot\big(\tilde{U}_2(:,\tilde{1})\big)^{\mathsf{T}} \end{pmatrix}. \tag{47}$$

Proposition 33. *It holds that $\mathcal{M}(\mathsf{N}_5) = \mathcal{M}(\mathsf{N}_4)$.*

Proof. By the definitions of $\mathcal{M}(\mathsf{N}_5)$ and $\mathcal{M}(\mathsf{N}_4)$, there is a bijection between the elements of $\mathcal{M}(\mathsf{N}_5)$ and $\mathcal{M}(\mathsf{N}_4)$. ■

Proposition 26 and Proposition 33 imply that there exists $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_5)$ s.t. $|\mathrm{Corr}(\boldsymbol{\beta}_{3,4})| > |\mathrm{Corr}(\boldsymbol{\beta}_{1,2})|$. Combining with Corollary 19, we can see that although N_2 has a topology similar to N_5, the DE-NFG N_5 provides extra marginals by varying ρ_{L}, $\tilde{U}_1$, and $\tilde{U}_2$.

VI. Relationship among the Sets

In this section, we prove that that the Venn diagram in Fig. 7 holds by proving that each part in the Venn diagram is non-mepty.

We prove that $\mathcal{M}(\mathsf{N}_3)$ and $\mathcal{M}(\mathsf{N}_4)$ are strict subsets of $\mathcal{LM}(\mathcal{K})$ first.

Lemma 34. *It holds that*

$$\mathcal{M}(\mathsf{N}_3) \subsetneq \mathcal{LM}(\mathcal{K}).$$

Proof. See Appendix H. ■

Corollary 35. *The vertices $\boldsymbol{v}_{17}, \ldots, \boldsymbol{v}_{24}$ of $\mathcal{LM}(\mathcal{K})$ (see Appendix A) are not in $\mathcal{M}(\mathsf{N}_3)$.*

Proof. The proof is similar to the proof of Lemma 34 and thus is omitted here. ■

Then we study the relationship between $\mathcal{M}(\mathsf{N}_4)$ and $\mathcal{LM}(\mathcal{K})$.

Lemma 36. *For a $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4)$ satisfying*

$$\boldsymbol{\beta}_{1,2} = \begin{pmatrix} 0 & \alpha \\ 1-\alpha & 0 \end{pmatrix}, \qquad \alpha \in \mathbb{R},\ 0 \leq \alpha \leq 1,$$

the associated matrix ρ equals

$$\rho = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \alpha & \overline{\rho_{3,2}} & 0 \\ 0 & \rho_{3,2} & 1-\alpha & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad |\rho_{3,2}| \leq \sqrt{\alpha \cdot (1-\alpha)}.$$

Proof. See Appendix I ■

Similar to the proof of Lemma 34, we prove that there are points in $\mathcal{LM}(\mathcal{K})$ that are not in $\mathcal{M}(\mathsf{N}_4)$.

Lemma 37. *The vertex $\boldsymbol{v}_{18}$ of $\mathcal{LM}(\mathcal{K})$ (see Appendix A) is not in $\mathcal{M}(\mathsf{N}_4)$.*

Proof. See Appendix J. ■

Lemma 38. *The vertices $\boldsymbol{v}_{17}, \ldots, \boldsymbol{v}_{24}$ of $\mathcal{LM}(\mathcal{K})$ (see Appendix A) are not in $\mathcal{M}(\mathsf{N}_4)$.*

Proof. The proof is similar to the proof of Lemma 37 and thus is omitted here. ■

We can further prove that there are sets in $\mathcal{LM}(\mathcal{K})$ that are not in $\mathcal{M}(\mathsf{N}_4)$.

Lemma 39. *It holds that*

$$\mathcal{S}_{3,4,7,10,18,22}(\mathsf{N}_1) := \left\{ \boldsymbol{v} \in \mathcal{LM}(\mathcal{K}) \;\middle|\; \begin{array}{l} \boldsymbol{v} = \alpha_3 \boldsymbol{v}_3 + \alpha_4 \boldsymbol{v}_4 + \alpha_7 \boldsymbol{v}_7 + \alpha_{10} \boldsymbol{v}_{10} + \alpha_{18} \boldsymbol{v}_{18} + \alpha_{23} \boldsymbol{v}_{23}, \\ \alpha_3, \ldots, \alpha_{23} \in \mathbb{R}_{\geq 0},\ \alpha_3 + \alpha_4 + \alpha_7 + \alpha_{10} + \alpha_{18} + \alpha_{23} = 1, \\ 0 < \alpha_3 + \alpha_7 < 1,\ 0 < \alpha_4 + \alpha_{10} < 1 \end{array} \right\} \nsubseteq \mathcal{M}(\mathsf{N}_4).$$

where $\boldsymbol{v}_3, \ldots, \boldsymbol{v}_{23}$ are given in Appendix A.

Proof. See Appendix K. ■

Lemma 40. *It holds that*

$$\mathcal{S}_{5,9,12,13,18,23}(\mathsf{N}_1) := \left\{ \boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K}) \;\middle|\; \begin{array}{l} \boldsymbol{\beta} = \alpha_5 \boldsymbol{v}_5 + \alpha_9 \boldsymbol{v}_9 + \alpha_{12} \boldsymbol{v}_{12} + \alpha_{13} \boldsymbol{v}_{13} + \alpha_{18} \boldsymbol{v}_{18} + \alpha_{23} \boldsymbol{v}_{23}, \\ \alpha_5, \ldots, \alpha_{23} \in \mathbb{R}_{\geq 0},\ \alpha_5 + \alpha_9 + \alpha_{12} + \alpha_{13} + \alpha_{18} + \alpha_{23} = 1, \\ 0 < \alpha_5 + \alpha_{12} < 1,\ 0 < \alpha_8 + \alpha_{13} < 1 \end{array} \right\} \nsubseteq \mathcal{M}(\mathsf{N}_4),$$

$$\mathcal{S}_{2,8,14,15,17,20}(\mathsf{N}_1) := \left\{ \boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K}) \;\middle|\; \begin{array}{l} \boldsymbol{\beta} = \alpha_2 \boldsymbol{v}_2 + \alpha_8 \boldsymbol{v}_8 + \alpha_{14} \boldsymbol{v}_{14} + \alpha_{15} \boldsymbol{v}_{15} + \alpha_{17} \boldsymbol{v}_{17} + \alpha_{20} \boldsymbol{v}_{20} \\ \alpha_2, \ldots, \alpha_{20} \in \mathbb{R}_{\geq 0},\ \alpha_2 + \alpha_8 + \alpha_{14} + \alpha_{15} + \alpha_{17} + \alpha_{20} = 1, \\ 0 < \alpha_2 + \alpha_{14} < 1,\ 0 < \alpha_8 + \alpha_{15} < 1 \end{array} \right\} \nsubseteq \mathcal{M}(\mathsf{N}_4),$$

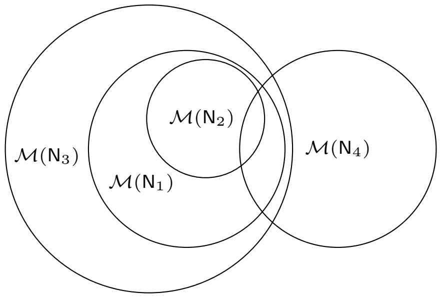

Fig. 8: The Venn diagram of $\mathcal{M}(\mathsf{N}_1)$, $\mathcal{M}(\mathsf{N}_2)$, $\mathcal{M}(\mathsf{N}_3)$ and $\mathcal{M}(\mathsf{N}_4)$.

$$\mathcal{S}_{1,6,11,16,17,20}(\mathsf{N}_1) := \left\{ \boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K}) \;\middle|\; \begin{array}{l} \boldsymbol{\beta} = \alpha_1 \boldsymbol{v}_1 + \alpha_6 \boldsymbol{v}_6 + \alpha_{11} \boldsymbol{v}_{11} + \alpha_{16} \boldsymbol{v}_{16} + \alpha_{17} \boldsymbol{v}_{17} + \alpha_{20} \boldsymbol{v}_{20}, \\ \alpha_1, \ldots, \alpha_{20} \in \mathbb{R}_{\geq 0},\ \alpha_1 + \alpha_6 + \alpha_{11} + \alpha_{16} + \alpha_{17} + \alpha_{20} = 1 \\ 0 < \alpha_1 + \alpha_{11} < 1,\ 0 < \alpha_6 + \alpha_{16} < 1 \end{array} \right\} \not\subseteq \mathcal{M}(\mathsf{N}_4),$$

where $\boldsymbol{v}_2, \ldots, \boldsymbol{v}_{23}$ are given in Appendix A.

Proof. The proof is similar to the proof of Lemma 39 and thus is omitted here. ∎

Corollary 41. *It holds that*

$$\mathcal{M}(\mathsf{N}_4) \subsetneq \mathcal{LM}(\mathcal{K}).$$

Proof. It can be proven by Lemmas 37–40. ∎

A. Proof of the Venn diagram in Fig. 8

In this subsection, we show that the Venn diagram in Fig. 8 holds.

It is easy to verify that

$$\mathcal{M}(\mathsf{N}_1) \subseteq \mathcal{M}(\mathsf{N}_3), \qquad \mathcal{M}(\mathsf{N}_2) \cap \mathcal{M}(\mathsf{N}_4) \neq \emptyset.$$

Lemma 42. *It holds that*

$$\mathcal{M}(\mathsf{N}_3) \setminus \big((\mathcal{M}(\mathsf{N}_1) \cup \mathcal{M}(\mathsf{N}_4)) \cap \mathcal{M}(\mathsf{N}_3) \big) \neq \emptyset.$$

Proof. See Appendix L. ∎

Lemma 43. *It holds that*

$$(\mathcal{M}(\mathsf{N}_3) \cap \mathcal{M}(\mathsf{N}_4)) \setminus \big(\mathcal{M}(\mathsf{N}_1) \cap \mathcal{M}(\mathsf{N}_4) \big) \neq \emptyset.$$

Proof. See Appendix M. ∎

Lemma 44. *It holds that*

$$\mathcal{M}(\mathsf{N}_4) \setminus (\mathcal{M}(\mathsf{N}_4) \cap \mathcal{M}(\mathsf{N}_3)) \neq \emptyset.$$

Proof. In Example 30, we show that there exists a $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4)$ satisfying

$$\boldsymbol{\beta}_{1,2} = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \frac{1}{6} \begin{pmatrix} 4 & 0 \\ 1 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,2} = \frac{1}{6} \begin{pmatrix} 4 & 1 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \frac{1}{12} \begin{pmatrix} 9 & 1 \\ 1 & 1 \end{pmatrix}.$$

On one hand, we have $\beta_{3,4}(1,1) = \frac{1}{12} > 0$. On the other hand, we have

$$\beta_{1,2}(1,1) = 0, \quad \beta_{3,2}(1,0) = 0, \quad \beta_{1,4}(0,1) = 0.$$

If such $\boldsymbol{\beta}$ is in $\mathcal{M}(\mathsf{N}_3)$, there exists a joint PMF $\{p_{\mathsf{N}_3}(\boldsymbol{x})\}_{\boldsymbol{x}}$ s.t.

$$p_{\mathsf{N}_3}(1,1,1,1) = \sum_{x_2} p_{\mathsf{N}_3}(0,x_2,1,1) = \sum_{x_1} p_{\mathsf{N}_3}(x_1,0,1,1) = 0,$$

$$\sum_{x_1,x_2} p_{\mathsf{N}_3}(x_1,x_2,1,1) > 0,$$

which is a contradiction. ■

Lemma 45. *We consider the following vector and matrices for the Q-FG* N_4 *in Fig. 4.*

$$U_1 = H, \quad U_2 = H, \quad \rho = \tfrac{1}{3} \cdot \begin{pmatrix} -1 & 1 & 1 & 0 \end{pmatrix}^{\mathsf{H}} \cdot \begin{pmatrix} -1 & 1 & 1 & 0 \end{pmatrix}.$$

The collection of matrices $\boldsymbol{\beta}$ *obtained via* (34)–(35) *are not in* $\mathcal{M}(\mathsf{N}_3)$.

Proof. The setup above is similar to the setup in Hardy's paradox (see Example 30) except that ρ is different. The collection of matrices $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4)$ obtained via (34) and (35) satisfies

$$\boldsymbol{\beta}_{1,2} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \frac{1}{6}\begin{pmatrix} 0 & 4 \\ 1 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,2} = \frac{1}{6}\begin{pmatrix} 0 & 1 \\ 4 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \frac{1}{12}\begin{pmatrix} 1 & 1 \\ 1 & 9 \end{pmatrix}.$$

Similar to the proof of Lemma 44, we can show that such $\boldsymbol{\beta} \notin \mathcal{M}(\mathsf{N}_3)$. ■

Lemma 46. *It holds that*

$$\mathcal{M}(\mathsf{N}_2) \setminus (\mathcal{M}(\mathsf{N}_4) \cap \mathcal{M}(\mathsf{N}_2)) \neq \emptyset.$$

Proof. See Appendix P. ■

Lemma 47. *It holds that*

$$\mathcal{M}(\mathsf{N}_1) \setminus \big((\mathcal{M}(\mathsf{N}_4) \cup \mathcal{M}(\mathsf{N}_2)) \cap \mathcal{M}(\mathsf{N}_1)\big) \neq \emptyset.$$

Proof. See Appendix N. ■

Lemma 48. *It holds that*

$$\mathcal{M}(\mathsf{N}_2) \subsetneq \mathcal{M}(\mathsf{N}_1), \quad \mathcal{M}(\mathsf{N}_1) \cap \mathcal{M}(\mathsf{N}_4) \setminus (\mathcal{M}(\mathsf{N}_2) \cap \mathcal{M}(\mathsf{N}_1) \cap \mathcal{M}(\mathsf{N}_4)) \neq \emptyset.$$

Proof. See Appendix O. ■

Theorem 49. *The Venn diagram in Fig. 7 holds.*

Proof. It can be proven by combining Lemmas 34–46. ■

We make some remarks on the Venn diagram in Fig. 7:

- On the one hand, the set of realizable marginals $\mathcal{M}(\mathsf{N}_4)$ provides extra marginals that are not in $\mathcal{M}(\mathsf{N}_3)$. For example, by introducing entanglement in the quantum system, one can obtain a set of incompatible marginals (see, e.g., Lemmas 44 and 45).
- On the other hand, the sets $\mathcal{M}(\mathsf{N}_1)$, $\mathcal{M}(\mathsf{N}_2)$, and $\mathcal{M}(\mathsf{N}_3)$ also consist of marginals that are not in $\mathcal{M}(\mathsf{N}_4)$.

Appendix A
Vertices of $\mathcal{LM}(\mathcal{K})$

Given $(\boldsymbol{\beta}_{i,j})_{i,j}$ for $\boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K})$, the matrix $\boldsymbol{\beta}_i$ can be obtained via

$$\sum_{x_j \in \mathcal{X}_e} \beta_{i,j}(x_i,x_j) = \beta_i(x_i), \qquad x_i \in \mathcal{X}_e,\ \{i,j\} \in \mathcal{K}.$$

It is sufficient to provide the values of $(\boldsymbol{\beta}_{i,j})_{i,j}$ to determine a $\boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K})$. Note that $\mathcal{LM}(\mathcal{K})$ is a convex set and using the lrs algorithm [35], a revised version of the reverse search vertex enumeration algorithm proposed in [36], we can find the vertices of $\mathcal{LM}(\mathcal{K})$, which are listed as follows.

	$\boldsymbol{\beta}_{1,4}$	$\boldsymbol{\beta}_{1,2}$	$\beta_{3,2}$	$\beta_{3,4}$		$\boldsymbol{\beta}_{1,4}$	$\boldsymbol{\beta}_{1,2}$	$\beta_{3,2}$	$\beta_{3,4}$
$\boldsymbol{v}_1$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\boldsymbol{v}_2$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$
$\boldsymbol{v}_3$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\boldsymbol{v}_4$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$
$\boldsymbol{v}_5$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\boldsymbol{v}_6$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$
$\boldsymbol{v}_7$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\boldsymbol{v}_8$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$
$\boldsymbol{v}_9$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\boldsymbol{v}_{10}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$
$\boldsymbol{v}_{11}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\boldsymbol{v}_{12}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$
$\boldsymbol{v}_{13}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\boldsymbol{v}_{14}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$	$\begin{pmatrix}0&1\\0&0\end{pmatrix}$
$\boldsymbol{v}_{15}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&0\end{pmatrix}$	$\boldsymbol{v}_{16}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0\\0&1\end{pmatrix}$
$\boldsymbol{v}_{17}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\boldsymbol{v}_{18}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$
$\boldsymbol{v}_{19}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\boldsymbol{v}_{20}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$
$\boldsymbol{v}_{21}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\boldsymbol{v}_{22}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$
$\boldsymbol{v}_{23}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\boldsymbol{v}_{24}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$

Appendix B
Proof of Lemma 8

In this appendix, we prove that

$$\sum_{x_i \in \mathcal{X}_e} \beta_i(x_i) \cdot x_i - \left(\sum_{x_i \in \mathcal{X}_e} \beta_i(x_i) \cdot x_i \right)^2 = \det(\boldsymbol{\beta}_i), \qquad i \in \{1,3\}, \tag{48}$$

$$\sum_{x_j \in \mathcal{X}_e} \beta_j(x_j) \cdot x_j - \left(\sum_{x_j \in \mathcal{X}_e} \beta_j(x_j) \cdot x_j \right)^2 = \det(\boldsymbol{\beta}_j), \qquad j \in \{2,4\},$$

$$\sum_{x_i, x_j} x_i \cdot x_j \cdot \beta_{i,j}(x_i, x_j) - \left(\sum_{x_i \in \mathcal{X}_e} \beta_i(x_i) \cdot x_i \right) \left(\sum_{x_j \in \mathcal{X}_e} \beta_j(x_j) \cdot x_j \right) = \det(\boldsymbol{\beta}_{i,j}), \qquad \{i,j\} \in \mathcal{K}. \tag{49}$$

We prove (48) first. By the definition of $\mathrm{Var}(Y_1)$ in (11), we have

$$\sum_{x_i \in \mathcal{X}_e} \beta_i(x_i) \cdot x_i - \left(\sum_{x_i \in \mathcal{X}_e} \beta_i(x_i) \cdot x_i \right)^2 \overset{(a)}{=} \beta_i(1) - \big(\beta_i(1)\big)^2 = \beta_i(0) \cdot \beta_i(1) = \det(\boldsymbol{\beta}_i).$$

where at step (a) we have used $\mathcal{X}_e = \{0, 1\}$. The proof w.r.t. $\det(\boldsymbol{\beta}_j)$ is similar and thus is omitted here.

In the rest of this proof, we prove (49). Because of $\mathcal{X}_e = \{0, 1\}$, we have

$$\begin{aligned}\sum_{x_i, x_j} x_i \cdot x_j \cdot \beta_{i,j}(x_i, x_j) - \left(\sum_{x_i \in \mathcal{X}_e} \beta_i(x_i) \cdot x_i \right)\left(\sum_{x_j \in \mathcal{X}_e} \beta_j(x_j) \cdot x_j \right) &= \beta_{i,j}(1,1) - (\beta_{i,j}(1,0) + \beta_{i,j}(1,1)) \cdot (\beta_{i,j}(0,1) + \beta_{i,j}(1,1)) \\ &\overset{(a)}{=} \beta_{i,j}(0,0) \cdot \beta_{i,j}(1,1) - \beta_{i,j} \cdot \beta_{i,j}(1,0) \\ &= \det(\boldsymbol{\beta}_{i,j}),\end{aligned}$$

where at step (a) we have used $\sum_{x_i, x_j} \beta_{i,j}(x_i, x_j) = 1$.

Appendix C
Proof of Theorem 12

In this appendix, we prove

$$\mathrm{CorrCHSH}(\boldsymbol{\beta}) < 2\sqrt{2},$$

assuming that

$$0 < \beta_i(0) < 1, \quad i \in \mathcal{E}(\mathsf{N}_1), \quad \boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3).$$

The proof of $\mathrm{CorrCHSH}(\boldsymbol{\beta}) > -2\sqrt{2}$ for the same setup is similar and thus is omitted here.

Note that there is a bijection between the set of all possible joint PMFs of random variables $X_1, \ldots, X_4 \in \{0, 1\}$ and $p_{\mathsf{N}_3}(\boldsymbol{x})$. It means that for any joint PMFs of random variables $X_1, \ldots, X_4$, there is an S-NFG N_3 s.t.

$$\Pr(X_1 = x_1, \ldots, X_4 = x_4) = p_{\mathsf{N}_3}(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathcal{X}_e^{|\mathcal{X}_e|}.$$

The other direction also holds. Thus it is equaivalent to prove that for any joint PMF of random variables $X_1, \ldots, X_4 \in \mathcal{X}_e$ s.t. $0 < \Pr(X_i = 0) < 1$, $i \in \mathcal{E}(\mathsf{N}_1)$, we have

$$\mathrm{Corr}(X_1, X_2) + \mathrm{Corr}(X_1, X_4) + \mathrm{Corr}(X_3, X_2) - \mathrm{Corr}(X_3, X_4) < 2\sqrt{2},$$

where $\mathrm{Corr}(X_i, X_j)$ is the Pearson correlation coefficient of X_i and X_j for $\{i, j\} \in \mathcal{K}$. For simplicity, we define

$$\check{X}_i := \frac{X_i - \mathbb{E}(X_i)}{\sqrt{\mathrm{Var}(X_i)}}, \qquad i \in \mathcal{E}(\mathsf{N}_1).$$

It follows that

$$\mathbb{E}\big((\check{X}_i)^2\big) = 1, \qquad i \in \mathcal{E}(\mathsf{N}_1), \tag{50}$$

$$\mathrm{Corr}(X_1, X_2) + \mathrm{Corr}(X_1, X_4) + \mathrm{Corr}(X_3, X_2) - \mathrm{Corr}(X_3, X_4) = \mathbb{E}\big(\check{X}_2(\check{X}_1 + \check{X}_3)\big) + \mathbb{E}\big(\check{X}_4(\check{X}_3 - \check{X}_1)\big). \tag{51}$$

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\mathbb{E}\big(\check{X}_2(\check{X}_1 + \check{X}_3)\big) &\le \sqrt{\mathbb{E}\big((\check{X}_2)^2\big)} \cdot \sqrt{\mathbb{E}\big((\check{X}_1 + \check{X}_3)^2\big)} \overset{(a)}{=} \sqrt{\mathbb{E}\big((\check{X}_1 + \check{X}_3)^2\big)}, \\ \mathbb{E}\big(\check{X}_4(\check{X}_3 - \check{X}_1)\big) &\le \sqrt{\mathbb{E}\big((\check{X}_4)^2\big)} \cdot \sqrt{\mathbb{E}\big((\check{X}_3 - \check{X}_1)^2\big)} \overset{(a)}{=} \sqrt{\mathbb{E}\big((\check{X}_3 - \check{X}_1)^2\big)},\end{aligned}$$

where at step (a) we have used (50). Then the expression (51) is bounded by

$$\mathbb{E}\big(\check{X}_2(\check{X}_1 + \check{X}_3)\big) + \mathbb{E}\big(\check{X}_4(\check{X}_3 - \check{X}_1)\big) \le \sqrt{\mathbb{E}\big((\check{X}_1 + \check{X}_3)^2\big)} + \sqrt{\mathbb{E}\big((\check{X}_3 - \check{X}_1)^2\big)}. \tag{52}$$

Let us focus on the RHS of the inequality (52). We have

$$\mathbb{E}\big((\check{X}_1+\check{X}_3)^2\big)+\mathbb{E}\big((\check{X}_3-\check{X}_1)^2\big)=2\mathbb{E}\big((\check{X}_1)^2+(\check{X}_3)^2\big)\overset{(a)}{=}4,$$

where at step (a) we have used (50). Note that

$$2\sqrt{2}=\max_{0\le a\le 4}\sqrt{a}+\sqrt{4-a}.$$

Then we have

$$\sqrt{\mathbb{E}\big((\check{X}_1+\check{X}_3)^2\big)}+\sqrt{\mathbb{E}\big((\check{X}_3-\check{X}_1)^2\big)}\le 2\sqrt{2}. \tag{53}$$

In the rest of this proof, we prove that the equality in (53) cannot be achieved. In particular, the Cauchy-Schwarz inequality in (52) holds with equality iff

$$\check{X}_1+\check{X}_3=\mathbb{E}\left(\big(\check{X}_1+\check{X}_3\big)\cdot\check{X}_2\right)\cdot\check{X}_2, \qquad \text{with probability } 1, \tag{54}$$

$$\check{X}_3-\check{X}_1=\mathbb{E}\left(\big(\check{X}_3-\check{X}_1\big)\cdot\check{X}_4\right)\cdot\check{X}_4, \qquad \text{with probability } 1. \tag{55}$$

In particular, we have $\check{X}_3+\check{X}_1\in\mathcal{X}_1$ and $\check{X}_3-\check{X}_1\in\mathcal{X}_2$, where

$$\mathcal{X}_1:=\left\{\frac{x_3-\mathbb{E}(X_3)}{\sqrt{\mathrm{Var}(X_3)}}+\frac{x_1-\mathbb{E}(X_1)}{\sqrt{\mathrm{Var}(X_1)}}\right\}_{x_1,x_3\in\mathcal{X}_e}, \qquad \mathcal{X}_2:=\left\{\frac{x_3-\mathbb{E}(X_1)}{\sqrt{\mathrm{Var}(X_1)}}-\frac{x_1-\mathbb{E}(X_3)}{\sqrt{\mathrm{Var}(X_3)}}\right\}_{x_1,x_3\in\mathcal{X}_e}.$$

The random variables $\check{X}_2$ and $\check{X}_4$ have binary alphabet, which means that we need $|\mathcal{X}_1|=2$ and $|\mathcal{X}_2|=2$ in order to have (54) and (55). There are five conditions w.r.t. X_1 and X_3 that we need to analyse.

1) $X_1=0$ with probability 1. In this case we have

$$\mathbb{E}\big(\check{X}_2(\check{X}_1+\check{X}_3)+\check{X}_4(\check{X}_3-\check{X}_1)\big)=\mathbb{E}\big(\check{X}_3\check{X}_2+\check{X}_3\check{X}_4\big)\le 2.$$

2) $X_1=1$ with probability 1, $X_3=0$ with probability 1, or $X_3=1$ with probability 1. The analysis of these cases is similar to the analysis of the previous case and thus is omitted here.

3) $\Pr((X_1,X_3)=(1,0))=0$ and $\Pr((X_1,X_3)=(0,1))=0$. In this case, we have $X_3=X_1$ with probability 1, which implies $\check{X}_3=\check{X}_1$ with probability 1. Then the LHS of the inequality (53) equals

$$\sqrt{4\cdot\mathbb{E}\big((\check{X}_1)^2\big)}\overset{(a)}{=}2<2\sqrt{2},$$

where at step (a) we have used (50).

4) $\Pr((X_1,X_3)=(0,0))=0$ and $\Pr((X_1,X_3)=(1,1))=0$. The analysis of this case is similar to the analysis of the previous case and thus is omitted here.

5) The support size of $\Pr((X_1,X_3)=(x_1,x_3))$ is larger than or equal to 3. In this case, we cannot have $|\mathcal{X}_1|=|\mathcal{X}_2|=2$.

APPENDIX D
PROOF OF PROPOSITION 13

In this appendix, we prove that there exists $\{p_{\mathsf{N}_3}(\boldsymbol{x})\}_{\boldsymbol{x}}$ s.t. the marginals $(\boldsymbol{\beta}_{i,j})_{i,j}$ computed in (6) for N_3 satisfies

$$\mathrm{CorrCHSH}(\boldsymbol{\beta})=\frac{5}{2}, \qquad 0<\beta_i(0)<1,\ i\in\mathcal{E}(\mathsf{N}_1).$$

We consider a PMF of $X_1,\ldots,X_4$.

$(\boldsymbol{x})$	$(0,0,0,1)$	$(1,1,0,1)$	$(0,1,1,0)$	Otherwise
$p_{\mathsf{N}_3}(\boldsymbol{x})$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0

(56)

Then the marginals computed in (6) are

$$\boldsymbol{\beta}_{1,2}=\begin{pmatrix}\frac{1}{3}&\frac{1}{3}\\0&\frac{1}{3}\end{pmatrix}, \qquad \boldsymbol{\beta}_{3,2}=\begin{pmatrix}\frac{1}{3}&\frac{1}{3}\\0&\frac{1}{3}\end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4}=\begin{pmatrix}\frac{1}{3}&\frac{1}{3}\\0&\frac{1}{3}\end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4}=\begin{pmatrix}0&\frac{2}{3}\\\frac{1}{3}&0\end{pmatrix}, \tag{57}$$

$$\boldsymbol{\beta}_1 = \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix}, \qquad \boldsymbol{\beta}_2 = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{2}{3} \end{pmatrix}, \qquad \boldsymbol{\beta}_3 = \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix}, \qquad \boldsymbol{\beta}_4 = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{2}{3} \end{pmatrix}.$$

The resulting correlation coefficients are

$$\mathrm{Corr}(\boldsymbol{\beta}_{1,2}) = \mathrm{Corr}(\boldsymbol{\beta}_{3,2}) = \mathrm{Corr}(\boldsymbol{\beta}_{1,4}) = \frac{1}{2}, \qquad \mathrm{Corr}(\boldsymbol{\beta}_{3,4}) = -1,$$

which implies

$$\mathrm{CorrCHSH}(\boldsymbol{\beta}) = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + 1 = \frac{5}{2}.$$

Appendix E
Proof of Theorem 14

In this appendix, we prove that

$$\mathrm{CorrCHSH}(\boldsymbol{\beta}) \leq \frac{5}{2}, \qquad\qquad 0 < \beta_i(0) < 1,\ i \in \mathcal{E}(\mathsf{N}_1),\ \boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3). \tag{58}$$

The proof for the inequality $\mathrm{CorrCHSH}(\boldsymbol{\beta}) \geq -\frac{5}{2}$ in the same setup is similar and thus is omitted here.

Definition 50. *The set $\mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ is defined to be*

$$\mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K}) := \{\boldsymbol{\beta} \in \mathcal{LM}(\mathcal{K}) \,|(59) \textit{ and } (60) \textit{ hold}\},$$

where

$$\beta_i(0) \cdot \beta_i(1) > 0, \qquad i \in \mathcal{E}(\mathsf{N}_1), \tag{59}$$

$$\sum_{\{i,j\} \in \mathcal{K}} (-1)^{[i=3,j=4]} \cdot \Big(\beta_{i,j}(0,0) + \beta_{i,j}(1,1) - \beta_{i,j}(0,1) - \beta_{i,j}(1,0)\Big) \leq 2. \tag{60}$$

The inequality (60) *ensures that $\boldsymbol{\beta}$ satisfies the CHSH inequality [22]. We will discuss the details in the proof of Lemma 51.*

Lemma 51. *It holds that for any $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$ s.t. $\beta_i(0) \cdot \beta_i(1) > 0$, $i \in \mathcal{E}(\mathsf{N}_1)$, we have $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$. Recall that $\mathcal{M}(\mathsf{N}_3)$ is defined in* (5)*.*

Proof. By the definition of $\mathcal{M}(\mathsf{N}_3)$, a set of matrices $\boldsymbol{\beta}$ in $\mathcal{M}(\mathsf{N}_3)$ is also in $\mathcal{LM}(\mathcal{K})$. We need to prove that $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$ satisfies (60) as well.

When a $\boldsymbol{\beta}$ is in $\mathcal{M}(\mathsf{N}_3)$, there exists a joint PMF $\big\{p_{\mathsf{N}_3}(\boldsymbol{x})\big\}_{\boldsymbol{x}}$ and the associated marginals $\{\boldsymbol{\beta}_{i,j}\}_{i,j}$ and $\{\boldsymbol{\beta}_i\}_i$ s.t. (6) holds. For any joint PMF of random variables $X_1, \ldots, X_4 \in \mathcal{X}_e$, there is a set of PMFs $\big\{p_{\mathsf{N}_3}(\boldsymbol{x})\big\}_{\boldsymbol{x}}$ realizing it, i.e.,

$$\Pr(X_1 = x_1, \ldots X_4 = x_4) = p_{\mathsf{N}_3}(\boldsymbol{x}), \qquad x_1, \ldots, x_4 \in \mathcal{X}_e. \tag{61}$$

Suppose that we obtain the joint PMF of $X_1, \ldots, X_4$, the well-known CHSH inequality [22] implies that

$$\mathbb{E}\Big((-1)^{X_1} \cdot (-1)^{X_2}\Big) + \mathbb{E}\Big((-1)^{X_3} \cdot (-1)^{X_2}\Big) + \mathbb{E}\Big((-1)^{X_1} \cdot (-1)^{X_4}\Big) - \mathbb{E}\Big((-1)^{X_3} \cdot (-1)^{X_4}\Big) \leq 2. \tag{62}$$

By (61) and (6), inequality (62) implies that (60) holds for $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$. ■

Proposition 52. *It holds that*

$$\mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K}) \backslash \mathcal{M}(\mathsf{N}_3) \neq \emptyset.$$

Proof. Let us consider $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ s.t.

$$\beta_{1,2}(x_1, x_2) = \frac{1}{2} \cdot [x_1 = x_2], \qquad x_1, x_2 \in \mathcal{X}_e, \qquad \beta_{3,2}(x_3, x_2) = \frac{1}{2} \cdot [x_2 \neq x_3], \qquad x_2, x_3 \in \mathcal{X}_e,$$

$$\beta_{1,4}(x_1, x_4) = \frac{1}{2} \cdot [x_1 = x_4], \qquad x_1, x_4 \in \mathcal{X}_e, \qquad \boldsymbol{\beta}_{3,4}(x_3, x_4) = \frac{1}{2} \cdot [x_3 = x_4], \qquad x_3, x_4 \in \mathcal{X}_e.$$

If the above set of marginals is in $\in \mathcal{M}(\mathsf{N}_3)$, there exists a joint PMF $\left\{p_{\mathsf{N}_3}(\boldsymbol{x})\right\}_{\boldsymbol{x}}$ s.t. $\boldsymbol{\beta}$ in (6) for N_3 satisfies the above expressions. In order to have such joint PMF, the valid configurations in $\mathcal{C}(\mathsf{N}_3)$ satisfy

$$x_1 = x_2 = x_3 = x_4, \qquad x_2 \neq x_3.$$

which is a contradiction. Such a $\boldsymbol{\beta}$ is not in $\mathcal{M}(\mathsf{N}_3)$. ■

If we prove

$$\mathrm{CorrCHSH}(\boldsymbol{\beta}) \leq \frac{5}{2}, \qquad \boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K}), \tag{63}$$

then we can prove (58) with the help of Lemma 51. Recall that the definition of $\mathrm{CorrCHSH}(\boldsymbol{\beta})$ is given in (15). We prove (63) for some special cases first and then we generalize the proof. The proof strategy is shown as follows:

1) We formulate an optimization problem where the $\mathrm{CorrCHSH}(\boldsymbol{\beta})$ is maximized over $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ s.t. $\boldsymbol{\beta}$ has a similar structure as $(\boldsymbol{\beta}_{i,j})_{i,j}$ in (57), i.e., having the same number of zero entries as in $(\boldsymbol{\beta}_{i,j})_{i,j}$. Note that $(\boldsymbol{\beta}_{i,j})_{i,j}$ in (57) has four zero entries. This problem is an optimization problem with linear constraints only, which helps determine the optimal solution. We prove $\mathrm{CorrCHSH}(\boldsymbol{\beta}) \leq 5/2$ in this case.
2) Then we consider another optimization problem where the $\mathrm{CorrCHSH}(\boldsymbol{\beta})$ is maximized over $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ s.t. the number of zero entries in $(\boldsymbol{\beta}_{i,j})_{i,j}$ is three. We need to prove that to maximize $\mathrm{CorrCHSH}(\boldsymbol{\beta})$, the optimal $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ needs to have four zero entries in $(\boldsymbol{\beta}_{i,j})_{i,j}$. Then we have $\mathrm{CorrCHSH}(\boldsymbol{\beta}) \leq 5/2$ in this case.
3) By repeating the similar procedure, we prove $\mathrm{CorrCHSH}(\boldsymbol{\beta}) \leq 5/2$ for all $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$.

A. Four Zeros

In the proof of Proposition 13 we show that the PMF in (56) achieves the equality in (63). Based on this observation, in this subsection, we prove (63) for $\boldsymbol{\beta}$ s.t.

$$\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K}), \qquad \beta_{1,2}(1,0) = 0, \qquad \beta_{1,4}(1,0) = 0, \qquad \beta_{3,2}(1,0) = 0, \qquad \beta_{3,4}(0,0) = 0. \tag{64}$$

We define $\boldsymbol{\beta}(p_1, \ldots, p_4) \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ to be the set of matrices satisfying (64) s.t.

$$\boldsymbol{\beta}_{1,2} = \begin{pmatrix} p_1 & 1-p_1-p_2 \\ 0 & p_2 \end{pmatrix}, \quad \boldsymbol{\beta}_{1,4} = \begin{pmatrix} p_3 & 1-p_2-p_3 \\ 0 & p_2 \end{pmatrix}, \quad \boldsymbol{\beta}_{3,2} = \begin{pmatrix} p_1 & 1-p_1-p_4 \\ 0 & p_4 \end{pmatrix}, \quad \boldsymbol{\beta}_{3,4} = \begin{pmatrix} 0 & 1-p_4 \\ p_3 & p_4-p_3 \end{pmatrix}, \tag{65}$$

$$\boldsymbol{\beta}_1 = \mathrm{diag}\begin{pmatrix} 1-p_2 & p_2 \end{pmatrix}, \qquad \boldsymbol{\beta}_2 = \mathrm{diag}\begin{pmatrix} p_1 & 1-p_1 \end{pmatrix}, \qquad \boldsymbol{\beta}_3 = \mathrm{diag}\begin{pmatrix} 1-p_4 & p_4 \end{pmatrix}, \qquad \boldsymbol{\beta}_4 = \mathrm{diag}\begin{pmatrix} p_3 & 1-p_3 \end{pmatrix},$$

$$p_1+p_2 \leq 1,\ p_2+p_3 \leq 1,\ p_1+p_4 \leq 1,\ p_3 \leq p_4,\ 0 < p_1, \ldots, p_4 < 1. \tag{66}$$

Lemma 53. *If $\boldsymbol{\beta}(p_1, \ldots, p_4)$ is in $\mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$, we have*

$$p_1 + p_2 + p_3 \leq 1.$$

Proof. The function $\boldsymbol{\beta}(p_1, \ldots, p_4) \in \mathcal{LM}(\mathcal{K})$ satisfies (59). To ensure that $\boldsymbol{\beta}$ satisfies (60) as well, we need

$$p_1+p_2-1+p_1+p_2+p_3+p_2-1+p_3+p_2+p_1+p_4-1+p_1+p_4+p_3+1-p_4-p_4+p_3 = 4(p_1+p_2+p_3)-2 \leq 2.$$

■

The correlation coefficients for $\boldsymbol{\beta}(p_1, \ldots, p_4)$ are

$$\mathrm{Corr}(\boldsymbol{\beta}_{1,2}) = \sqrt{\frac{p_1}{1-p_1}} \cdot \sqrt{\frac{p_2}{1-p_2}}, \qquad \mathrm{Corr}(\boldsymbol{\beta}_{1,4}) = \sqrt{\frac{p_2}{1-p_2}} \cdot \sqrt{\frac{p_3}{1-p_3}},$$

$$\mathrm{Corr}(\boldsymbol{\beta}_{3,2}) = \sqrt{\frac{p_1}{1-p_1}} \cdot \sqrt{\frac{p_4}{1-p_4}}, \qquad \mathrm{Corr}(\boldsymbol{\beta}_{3,4}) = -\sqrt{\frac{p_3}{1-p_3}} \cdot \sqrt{\frac{1-p_4}{p_4}},$$

We then define an optimization problem:

$$F_1 := \sup_{p_1,p_2,p_3,p_4} \quad f_1(p_1,p_2,p_3,p_4)$$
$$\text{s.t.} \quad p_1+p_4 \le 1,\ p_3 \le p_4,\ p_1+p_2+p_3 \le 1,\ 0 < p_1,p_2,p_3,p_4 < 1, \tag{67}$$

where

$$f_1(p_1,p_2,p_3,p_4) := \sqrt{\frac{p_1}{1-p_1}}\cdot\sqrt{\frac{p_2}{1-p_2}} + \sqrt{\frac{p_2}{1-p_2}}\cdot\sqrt{\frac{p_3}{1-p_3}} + \sqrt{\frac{p_1}{1-p_1}}\cdot\sqrt{\frac{p_4}{1-p_4}} + \sqrt{\frac{p_3}{1-p_3}}\cdot\sqrt{\frac{1-p_4}{p_4}}.$$

If we prove $F_1 \le 5/2$, then we prove (63) for $\boldsymbol{\beta}$ satisfying (64).

Lemma 54. *In order to find F_1, we can set at least one of the following inequalities*

$$p_3 \le p_4, \qquad p_1+p_2+p_3 \le 1, \tag{68}$$

to be an equality.

Proof. The function f_1 is non-decreasing w.r.t. p_3, which means that at least one of the inequalities w.r.t. the upper bound of p_3 can be an equality to maximize f_1. ∎

With the help of Lemma 54, we can solve the problem (67) by considering two cases w.r.t. p_3, respectively. In each case, one of the inequalities in (68) is an equality. The first case is $p_4 = p_3$.

Lemma 55. *It holds that $F_1 = 5/2$ when $p_4 = p_3$.*

Proof. When we let $p_1 = p_2 = p_3 = 1/3$, we have $f_1(p_1,p_2,p_3,p_3) = 5/2$. It is sufficient to prove $F_1 \le 5/2$.

Because f_1 is non-decreasing w.r.t. p_3, we set $p_1+p_2+p_3 = 1$ in order to maximize f_1. We have

$$\begin{aligned} f_1(p_1,p_2,1-p_1-p_2,1-p_1-p_2) &= \sqrt{\frac{p_2}{1-p_1}}\cdot\sqrt{\frac{p_1}{1-p_2}} + \sqrt{\frac{p_2}{p_1+p_2}}\cdot\sqrt{\frac{1-p_1-p_2}{1-p_2}} + \sqrt{\frac{p_1}{p_1+p_2}}\cdot\sqrt{\frac{1-p_1-p_2}{1-p_1}} + 1 \\ &\overset{(a)}{\le} \sqrt{\frac{p_2}{1-p_1}}\cdot\sqrt{\frac{p_1}{1-p_2}} + \sqrt{\left|1-\frac{p_1}{1-p_2}\right| + \left|1-\frac{p_2}{1-p_1}\right|} + 1 \\ &\overset{(b)}{=} \sqrt{\frac{p_2}{1-p_1}}\cdot\sqrt{\frac{p_1}{1-p_2}} + \sqrt{2-\frac{p_1}{1-p_2}-\frac{p_2}{1-p_1}} + 1 \end{aligned}$$

where at step (a) we have used the Cauchy-Schwarz inequality, i.e.,

$$\sqrt{\frac{p_2}{p_1+p_2}}\cdot\sqrt{\frac{1-p_1-p_2}{1-p_2}} + \sqrt{\frac{p_1}{p_1+p_2}}\cdot\sqrt{\frac{1-p_1-p_2}{1-p_1}} \le \underbrace{\sqrt{\left|\sqrt{\frac{p_2}{p_1+p_2}}\right|^2 + \left|\sqrt{\frac{p_1}{p_1+p_2}}\right|^2}}_{=1}\cdot\sqrt{\left|\sqrt{1-\frac{p_1}{1-p_2}}\right|^2 + \left|\sqrt{1-\frac{p_2}{1-p_1}}\right|^2}$$

and at step (b) we have used $p_1+p_2+p_3 = 1$. Considering an optimization problem

$$\sup_{x,y} \quad \sqrt{xy} + \sqrt{2-x-y} + 1$$
$$\text{s.t.} \quad 0 \le x,y \le 1,$$

we have $\sqrt{xy} + \sqrt{2-x-y} + 1 \le \sqrt{2} + 1$ when $x = 0$, $y = 0$, $x = 1$, or $y = 1$. When $x,y > 0$ and $x+y < 2$, we have

$$\frac{\partial}{\partial x}\left(\sqrt{xy} + \sqrt{2-x-y}\right) = \frac{1}{2}\left(\sqrt{\frac{y}{x}} - \frac{1}{\sqrt{2-x-y}}\right), \qquad \frac{\partial}{\partial y}\left(\sqrt{xy} + \sqrt{2-x-y}\right) = \frac{1}{2}\left(\sqrt{\frac{x}{y}} - \frac{1}{\sqrt{2-x-y}}\right).$$

Setting the above expressions equal to zeros, which is the necessary condition for x and y to be the optimal solutions, we have $x = y = 1/2$ and the associated optimal value is $5/2$. Then we prove that $F_1 = 5/2$. ∎

The second case is $p_3 = 1-p_1-p_2$. We obtain the following lemma.

Lemma 56. *When $p_3 = 1-p_1-p_2$, we have $F_1 \le 5/2$.*

Proof. The proof is similar to the proof of Lemma 55 and thus is omitted here. ■

Lemma 57. *Suppose that $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ and for all $\{i,j\} \in \mathcal{K}$, at least one of the entries in $\boldsymbol{\beta}_{i,j}$ is zero, we obtain* (63).

Proof. In this subsection, we prove (63) for $\boldsymbol{\beta}$ satisfying (64). The proof for other cases is similar and thus is omitted here. ■

B. Three Zeros

In this subsection, we prove (63) for $\boldsymbol{\beta}$ s.t.

$$\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K}), \qquad \beta_{1,2}(1,0) = 0, \qquad \beta_{3,2}(1,0) = 0, \qquad \beta_{3,4}(0,0) = 0. \tag{69}$$

We define $\boldsymbol{\beta}(p_1, \ldots, p_5) \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ to be the set of matrices s.t.

$$\boldsymbol{\beta}_{1,2} = \begin{pmatrix} p_1 & 1-p_1-p_2 \\ 0 & p_2 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \begin{pmatrix} p_3 & 1-p_2-p_3 \\ p_2-p_4 & p_4 \end{pmatrix}, \tag{70}$$

$$\boldsymbol{\beta}_{3,2} = \begin{pmatrix} p_1 & 1-p_1-p_5 \\ 0 & p_5 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \begin{pmatrix} 0 & 1-p_5 \\ p_2+p_3-p_4 & p_4+p_5-p_2-p_3 \end{pmatrix}, \tag{71}$$

$$\boldsymbol{\beta}_1 = \mathrm{diag}\begin{pmatrix} 1-p_2 & p_2 \end{pmatrix}, \ \boldsymbol{\beta}_2 = \mathrm{diag}\begin{pmatrix} p_1 & 1-p_1 \end{pmatrix}, \ \boldsymbol{\beta}_3 = \mathrm{diag}\begin{pmatrix} 1-p_5 & p_5 \end{pmatrix}, \ \boldsymbol{\beta}_4 = \mathrm{diag}\begin{pmatrix} p_2+p_3-p_4 & 1+p_4-p_2-p_3 \end{pmatrix},$$

$$p_1+p_2 \le 1, \ p_2+p_3 \le 1, \ p_4 \le p_2, \ p_1+p_5 \le 1, \ p_4 < p_2+p_3 \le p_4+p_5, \ p_3, p_4 \ge 0, \ p_1, p_2, p_5 > 0, \ p_1, \ldots, p_5 < 1. \tag{72}$$

Lemma 58. *If $\boldsymbol{\beta}(p_1, \ldots, p_5)$ is in $\mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$, we have*

$$p_1 + p_2 + p_3 \le 1.$$

Proof. It is similar to the proof of Lemma 53 and thus is omitted here. ■

The correlation coefficients for $\boldsymbol{\beta}(p_1, \ldots, p_5)$ are

$$\mathrm{Corr}\big(\boldsymbol{\beta}_{1,2}\big) = \sqrt{\frac{p_1}{1-p_1}} \cdot \sqrt{\frac{p_2}{1-p_2}}, \qquad \mathrm{Corr}\big(\boldsymbol{\beta}_{1,4}\big) = \frac{p_4(1-p_2) - p_2(1-p_2-p_3)}{\sqrt{p_2(1-p_2)(p_2+p_3-p_4)(1+p_4-p_2-p_3)}},$$

$$\mathrm{Corr}\big(\boldsymbol{\beta}_{3,2}\big) = \sqrt{\frac{p_1}{1-p_1}} \cdot \sqrt{\frac{p_5}{1-p_5}}, \qquad \mathrm{Corr}\big(\boldsymbol{\beta}_{3,4}\big) = -\sqrt{\frac{p_2+p_3-p_4}{1+p_4-p_2-p_3}} \cdot \sqrt{\frac{1-p_5}{p_5}}.$$

We consider an optimization problem:

$$\begin{aligned} F_2 := \sup_{p_1, \ldots, p_5} \quad & f_2(p_1, \ldots, p_5) \\ \text{s.t.} \quad & p_1+p_2+p_3 \le 1, \ p_4 \le p_2, \ p_1+p_5 \le 1, \ p_4 < p_2+p_3 \le p_4+p_5, \ p_3, p_4 \ge 0, \ p_1, p_2, p_5 > 0, \ p_1, \ldots, p_5 < 1. \end{aligned} \tag{73}$$

where

$$\begin{aligned} f_2(p_1, \ldots, p_5) := & \sqrt{\frac{p_1}{1-p_1}} \cdot \sqrt{\frac{p_2}{1-p_2}} + \frac{p_4(1-p_2) - p_2(1-p_2-p_3)}{\sqrt{p_2(1-p_2)(p_2+p_3-p_4)(1+p_4-p_2-p_3)}} + \sqrt{\frac{p_1}{1-p_1}} \cdot \sqrt{\frac{p_5}{1-p_5}} \\ & + \sqrt{\frac{p_2+p_3-p_4}{1+p_4-p_2-p_3}} \cdot \sqrt{\frac{1-p_5}{p_5}}. \end{aligned}$$

If we prove $F_2 \le 5/2$, then we prove (63) for $\boldsymbol{\beta}$ satisfying (69).

Lemma 59. *It holds that $F_2 \le 5/2$.*

Proof. The proof is similar to the proof in Section E-A and thus is omitted here. ■

Lemma 60. *Suppose that $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ and there are (at least) three of the matrices in $(\boldsymbol{\beta}_{i,j})_{i,j}$ s.t. each of these three matrices contains at least one zero entry, we obtain* (63).

Proof. In this subsection, we prove (63) for $\boldsymbol{\beta}$ satisfying (69). The proof for other cases is similar and thus is omitted here. ∎

C. Two Zeros

In this subsection, we prove (63) for $\boldsymbol{\beta}$ s.t.

$$\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K}), \qquad \beta_{3,2}(1,0) = 0, \qquad \beta_{3,4}(0,0) = 0. \tag{74}$$

We define $\boldsymbol{\beta}(p_1, \dots, p_6) \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ to be the set of matrices s.t.

$$\boldsymbol{\beta}_{1,2} = \begin{pmatrix} p_6 & 1 + p_4 - p_1 - p_5 - p_6 \\ p_3 - p_6 & p_1 + p_5 + p_6 - p_3 - p_4 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \begin{pmatrix} p_4 & 1 - p_1 - p_5 \\ p_1 - p_4 & p_5 \end{pmatrix}, \tag{75}$$

$$\boldsymbol{\beta}_{3,2} = \begin{pmatrix} p_3 & p_2 - p_3 \\ 0 & 1 - p_2 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \begin{pmatrix} 0 & p_2 \\ p_1 & 1 - p_1 - p_2 \end{pmatrix}, \tag{76}$$

$$\boldsymbol{\beta}_1 = \mathrm{diag}\begin{pmatrix} 1 + p_4 - p_1 - p_5 & p_1 + p_5 - p_4 \end{pmatrix}, \ \boldsymbol{\beta}_2 = \mathrm{diag}\begin{pmatrix} p_3 & 1 - p_3 \end{pmatrix}, \ \boldsymbol{\beta}_3 = \mathrm{diag}\begin{pmatrix} p_2 & 1 - p_2 \end{pmatrix}, \ \boldsymbol{\beta}_4 = \mathrm{diag}\begin{pmatrix} p_1 & 1 - p_1 \end{pmatrix},$$

$p_6 \le p_3, \ p_3 + p_4 \le p_1 + p_5 + p_6 \le 1 + p_4, \ p_4 \le p_1, \ p_1 + p_5 \le 1, \ p_3 \le p_2, \ p_1 + p_2 \le 1, \ p_1 + p_5 < 1 + p_4, \ p_1 + p_5 > p_4,$
$p_1, \dots, p_6 < 1, \ p_1, p_2, p_3 > 0, \ p_4, p_5, p_6 \ge 0.$

Lemma 61. *If $\boldsymbol{\beta}(p_1, \dots, p_6)$ is in $\mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$, we have*

$$p_1 + p_5 + p_6 \le 1.$$

Proof. It is similar to the proof of Lemma 53 and thus is omitted here. ∎

The correlation coefficients for $\boldsymbol{\beta}(p_1, \dots, p_6)$ are

$$\mathrm{Corr}(\boldsymbol{\beta}_{1,2}) = \frac{p_6 - p_3(1 + p_4 - p_1 - p_5)}{\sqrt{p_3(1-p_3)(p_1 + p_5 - p_4)(1 + p_4 - p_1 - p_5)}}, \qquad \mathrm{Corr}(\boldsymbol{\beta}_{1,4}) = \frac{p_4 - p_1(1 + p_4 - p_1 - p_5)}{\sqrt{p_1(1-p_1)(p_1 + p_5 - p_4)(1 + p_4 - p_1 - p_5)}},$$

$$\mathrm{Corr}(\boldsymbol{\beta}_{3,2}) = \sqrt{\frac{p_3}{1-p_3}} \cdot \sqrt{\frac{1-p_2}{p_2}}, \qquad \mathrm{Corr}(\boldsymbol{\beta}_{3,4}) = -\sqrt{\frac{p_1}{1-p_1}} \cdot \sqrt{\frac{p_2}{1-p_2}}.$$

We consider an optimization problem:

$$\begin{aligned} F_3 := \sup_{p_1, \dots, p_6} \quad & f_3(p_1, \dots, p_6) \\ \text{s.t.} \quad & p_6 \le p_3, \ p_3 + p_4 \le p_1 + p_5 + p_6 \le 1, \ p_4 \le p_1, \ p_1 + p_5 \le 1, \ p_3 \le p_2, \ p_1 + p_2 \le 1, \ p_1 + p_5 < 1 + p_4, \\ & p_1 + p_5 > p_4, \ p_1, \dots, p_6 < 1, \ p_1, p_2, p_3 > 0, \ p_4, p_5, p_6 \ge 0, \end{aligned} \tag{77}$$

where

$$\begin{aligned} f_3(p_1, \dots, p_6) := & \frac{p_6 - p_3(1 + p_4 - p_1 - p_5)}{\sqrt{p_3(1-p_3)(p_1 + p_5 - p_4)(1 + p_4 - p_1 - p_5)}} + \sqrt{\frac{p_3}{1-p_3}} \cdot \sqrt{\frac{1-p_2}{p_2}} \\ & + \frac{p_4 - p_1(1 + p_4 - p_1 - p_5)}{\sqrt{p_1(1-p_1)(p_1 + p_5 - p_4)(1 + p_4 - p_1 - p_5)}} + \sqrt{\frac{p_1}{1-p_1}} \cdot \sqrt{\frac{p_2}{1-p_2}}. \end{aligned} \tag{78}$$

If we can prove that $F_3 \le 5/2$, then we prove (63) for $\boldsymbol{\beta}$ satisfying (74).

Lemma 62. *It holds that $F_3 \le 5/2$.*

Proof. The proof is similar to the proof in Section E-A and thus is omitted here. ∎

Lemma 63. *Suppose that $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ and there exist (at least) two of the matrices in $(\boldsymbol{\beta}_{i,j})_{i,j}$ s.t. each of these two matrices has at least one zero entry, we obtain* (63).

Proof. In this subsection, we prove (63) for $\boldsymbol{\beta}$ satisfying (74). The proof for other cases is similar and thus is omitted here. ∎

D. One Zero

In this subsection, we prove (63) for $\boldsymbol{\beta}$ s.t.

$$\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K}), \qquad \beta_{3,4}(0,0) = 0. \tag{79}$$

We define $\boldsymbol{\beta}(p_1, \ldots, p_7) \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ to be the set of matrices satisfying (79) s.t.

$$\boldsymbol{\beta}_{1,2} = \begin{pmatrix} p_3 & 1-p_3-p_4-p_5 \\ p_4 & p_5 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \begin{pmatrix} p_6 & 1-p_4-p_5-p_6 \\ p_1-p_6 & p_4+p_5+p_6-p_1 \end{pmatrix}, \tag{80}$$

$$\boldsymbol{\beta}_{3,2} = \begin{pmatrix} p_7 & p_2-p_7 \\ p_3+p_4-p_7 & 1+p_7-p_2-p_3-p_4 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \begin{pmatrix} 0 & p_2 \\ p_1 & 1-p_1-p_2 \end{pmatrix}, \tag{81}$$

$$\boldsymbol{\beta}_1 = \mathrm{diag}\begin{pmatrix} 1-p_4-p_5 & p_4+p_5 \end{pmatrix},\ \boldsymbol{\beta}_2 = \mathrm{diag}\begin{pmatrix} p_3+p_4 & 1-p_3-p_4 \end{pmatrix},\ \boldsymbol{\beta}_3 = \mathrm{diag}\begin{pmatrix} p_2 & 1-p_2 \end{pmatrix},\ \boldsymbol{\beta}_4 = \mathrm{diag}\begin{pmatrix} p_1 & 1-p_1 \end{pmatrix},$$

$$p_3+p_4+p_5 \le 1,\ 0 \le p_6 \le p_1 \le p_4+p_5+p_6 \le 1,\ p_7 \le p_2,\ p_7 \le p_3+p_4,\ p_2+p_3+p_4 \le 1+p_7,\ p_1+p_2 \le 1,$$
$$0 < p_3+p_4 < 1,\ 0 < p_4+p_5 < 1,\ p_1, p_2 > 0,\ p_3, \ldots, p_7 \ge 0,\ p_1, \ldots, p_7 < 1.$$

Lemma 64. *If $\boldsymbol{\beta}(p_1, \ldots, p_7)$ is in $\mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$, we have*

$$p_5 + p_6 + p_7 \le 1.$$

Proof. It is similar to the proof of Lemma 53 and thus is omitted here. ∎

The correlation coefficients for $\boldsymbol{\beta}(p_1, \ldots, p_7)$ are

$$\mathrm{Corr}(\boldsymbol{\beta}_{1,2}) = \frac{(p_3+p_4)(p_4+p_5)-p_4}{\sqrt{(p_3+p_4)(1-p_3-p_4)(p_4+p_5)(1-p_4-p_5)}}, \qquad \mathrm{Corr}(\boldsymbol{\beta}_{1,4}) = \frac{p_6-p_1(1-p_4-p_5)}{\sqrt{p_1(1-p_1)(p_4+p_5)(1-p_4-p_5)}},$$
$$\mathrm{Corr}(\boldsymbol{\beta}_{3,2}) = \frac{p_7-p_2(p_3+p_4)}{\sqrt{p_2(1-p_2)(p_3+p_4)(1-p_3-p_4)}}, \qquad \mathrm{Corr}(\boldsymbol{\beta}_{3,4}) = -\sqrt{\frac{p_1}{1-p_1}} \cdot \sqrt{\frac{p_2}{1-p_2}}.$$

We consider an optimization problem.

$$\begin{aligned} F_4 := \sup_{p_1,\ldots,p_7} \quad & f_4(p_1, \ldots, p_7) \\ \text{s.t.} \quad & p_5+p_6+p_7 \le 1,\ p_3+p_4+p_5 \le 1,\ 0 \le p_6 \le p_1 \le p_4+p_5+p_6 \le 1,\ p_7 \le p_3+p_4,\ p_2+p_3+p_4 \le 1+p_7, \\ & p_7 \le p_2,\ p_1+p_2 \le 1,\ 0 < p_3+p_4 < 1,\ 0 < p_4+p_5 < 1,\ p_1, p_2 > 0,\ p_3, \ldots, p_7 \ge 0,\ p_1, \ldots, p_7 < 1 \end{aligned} \tag{82}$$

where

$$\begin{aligned} f_4(p_1, \ldots, p_7) := & \frac{(p_3+p_4)(p_4+p_5)-p_4}{\sqrt{(p_3+p_4)(1-p_3-p_4)(p_4+p_5)(1-p_4-p_5)}} + \frac{p_6-p_1(1-p_4-p_5)}{\sqrt{p_1(1-p_1)(p_4+p_5)(1-p_4-p_5)}} \\ & + \frac{p_7-p_2(p_3+p_4)}{\sqrt{p_2(1-p_2)(p_3+p_4)(1-p_3-p_4)}} + \sqrt{\frac{p_1}{1-p_1}} \cdot \sqrt{\frac{p_2}{1-p_2}}. \end{aligned}$$

If we prove $F_4 \le 5/2$, then we prove (63) for $\boldsymbol{\beta}$ satisfying (79).

Lemma 65. *It holds that $F_4 \le 5/2$ for $p_5 > 0$.*

Proof. The proof is similar to the proof in Section E-A and thus is omitted here. ∎

Lemma 66. *Suppose that $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ and there exist a matrix in $(\boldsymbol{\beta}_{i,j})_{i,j}$ having at least one zero entry, we obtain* (63).

Proof. In this subsection, we prove (63) for $\boldsymbol{\beta}$ satisfying (79). The proof for other cases is similar and thus is omitted here. ∎

E. General Case

In this subsection, we prove (63) for $\boldsymbol{\beta} \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ based on the previously obtained results. We define $\boldsymbol{\beta}(p_1, \ldots, p_8) \in \mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$ to be the set of matrices s.t.

$$\boldsymbol{\beta}_{1,2} = \begin{pmatrix} p_1 & 1-p_1-p_2-p_3 \\ p_3 & p_2 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \begin{pmatrix} p_7 & 1-p_2-p_3-p_7 \\ p_4+p_6-p_7 & p_2+p_3+p_7-p_4-p_6 \end{pmatrix}, \tag{83}$$

$$\boldsymbol{\beta}_{3,2} = \begin{pmatrix} p_8 & p_5+p_6-p_8 \\ p_1+p_3-p_8 & 1+p_8-p_1-p_3-p_5-p_6 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \begin{pmatrix} p_6 & p_5 \\ p_4 & 1-p_4-p_5-p_6 \end{pmatrix}, \tag{84}$$

$$\boldsymbol{\beta}_1 = \mathrm{diag}\begin{pmatrix} 1-p_2-p_3 & p_2+p_3 \end{pmatrix}, \qquad \boldsymbol{\beta}_2 = \mathrm{diag}\begin{pmatrix} p_1+p_3 & 1-p_1-p_3 \end{pmatrix},$$

$$\boldsymbol{\beta}_3 = \mathrm{diag}\begin{pmatrix} p_5+p_6 & 1-p_5-p_6 \end{pmatrix}, \qquad \boldsymbol{\beta}_4 = \mathrm{diag}\begin{pmatrix} p_4+p_6 & 1-p_4-p_6 \end{pmatrix},$$

$$p_1+p_2+p_3 \le 1,\ 0 \le p_7 \le p_4+p_6 \le p_2+p_3+p_7 \le 1,\ p_8 \le p_5+p_6,\ p_8 \le p_1+p_3,\ p_1+p_3+p_5+p_6 \le 1+p_8,$$

$$p_4+p_5+p_6 \le 1,\ 0 < p_1+p_3 < 1,\ 0 < p_2+p_3 < 1,\ 0 < p_4+p_6 < 1,\ 0 < p_5+p_6 < 1,\ p_1, \ldots, p_8 > 0,\ p_7, p_8 < 1.$$

Lemma 67. *If $\boldsymbol{\beta}(p_1, \ldots, p_8)$ is in $\mathcal{LM}_{\mathrm{CHSH}}(\mathcal{K})$, we have*

$$p_2 - p_6 + p_7 + p_8 \le 1.$$

Proof. It is similar to the proof of Lemma 53 and thus is omitted here. ∎

The correlation coefficients for $\boldsymbol{\beta}(p_1, \ldots, p_8)$ are

$$\mathrm{Corr}(\boldsymbol{\beta}_{1,2}) = \frac{(p_1+p_3)(p_2+p_3)-p_3}{\sqrt{(p_1+p_3)(1-p_1-p_3)(p_2+p_3)(1-p_2-p_3)}}, \quad \mathrm{Corr}(\boldsymbol{\beta}_{1,4}) = \frac{p_7-(p_4+p_6)(1-p_2-p_3)}{\sqrt{(p_4+p_6)(1-p_4-p_6)(p_2+p_3)(1-p_2-p_3)}},$$

$$\mathrm{Corr}(\boldsymbol{\beta}_{3,2}) = \frac{p_8-(p_1+p_3)(p_5+p_6)}{\sqrt{(p_1+p_3)(1-p_1-p_3)(p_5+p_6)(1-p_5-p_6)}}, \quad \mathrm{Corr}(\boldsymbol{\beta}_{3,4}) = -\frac{(p_4+p_6)(p_5+p_6)-p_6}{\sqrt{(p_4+p_6)(1-p_4-p_6)(p_5+p_6)(1-p_5-p_6)}}.$$

We then consider an optimization problem:

$$\begin{aligned} F_5 := \sup_{p_1,\ldots,p_8} \quad & f_5(p_1, \ldots, p_8) \\ \text{s.t.} \quad & p_2-p_6+p_7+p_8 \le 1,\ p_1+p_2+p_3 \le 1,\ 0 \le p_7 \le p_4+p_6 \le p_2+p_3+p_7 \le 1,\ p_8 \le p_5+p_6,\ p_8 \le p_1+p_3, \\ & p_1+p_3+p_5+p_6 \le 1+p_8,\ p_4+p_5+p_6 \le 1,\ 0 < p_1+p_3 < 1,\ 0 < p_2+p_3 < 1, \\ & 0 < p_4+p_6 < 1,\ 0 < p_5+p_6 < 1,\ p_1, \ldots, p_8 \ge 0,\ p_7, p_8 < 1 \end{aligned} \tag{85}$$

where

$$\begin{aligned} f_5(p_1, \ldots, p_8) := & \frac{(p_1+p_3)(p_2+p_3)-p_3}{\sqrt{(p_1+p_3)(1-p_1-p_3)(p_2+p_3)(1-p_2-p_3)}} + \frac{p_7-(p_4+p_6)(1-p_2-p_3)}{\sqrt{(p_4+p_6)(1-p_4-p_6)(p_2+p_3)(1-p_2-p_3)}} \\ & + \frac{p_8-(p_1+p_3)(p_5+p_6)}{\sqrt{(p_1+p_3)(1-p_1-p_3)(p_5+p_6)(1-p_5-p_6)}} + \frac{(p_4+p_6)(p_5+p_6)-p_6}{\sqrt{(p_4+p_6)(1-p_4-p_6)(p_5+p_6)(1-p_5-p_6)}}. \end{aligned}$$

Lemma 68. *It holds that $F_5 \le 5/2$.*

Proof. The proof is similar to the proof in Section E-A. First we show that it is sufficient to consider $p_8 = 1 + p_6 - p_2 - p_7$. In particular, we note that when $p_6 = p_2 + p_7$, there is a stricter upper bound w.r.t. p_8. We cannot set $p_8 = 1 + p_6 - p_2 - p_7$ in this case. When $p_6 = p_2 + p_7$, in order to maximize f_5, the following inequalities is activated:

$$p_8 \le \max(p_5 + p_6, p_1 + p_3).$$

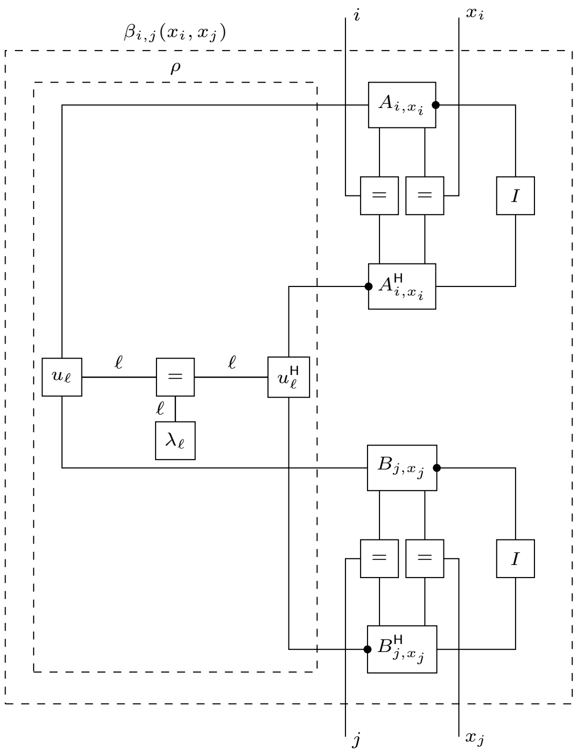

Fig. 9: The NFG representation of $\beta_{i,j}(x_i, x_j)$.

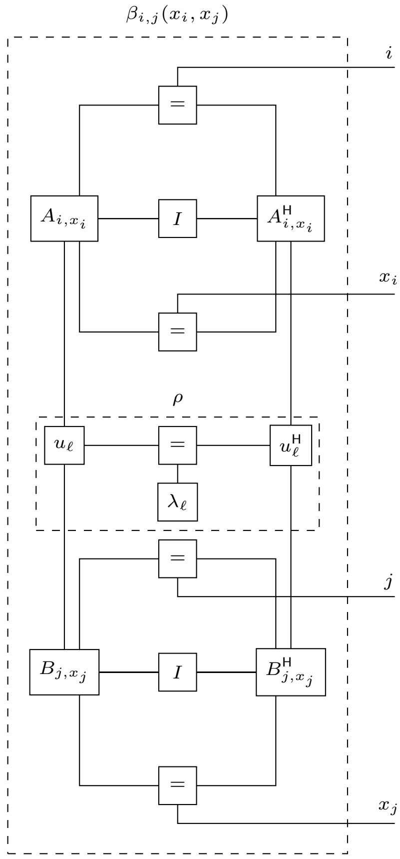

Fig. 10: An alternative NFG representation of $\beta_{i,j}(x_i, x_j)$.

In this case, one of the entries in $\boldsymbol{\beta}_{3,4}$ equals zero. By Lemma 66, we have $F_5 \leq 5/2$.

Then we find F_5 for $p_8 = 1 + p_6 - p_2 - p_7$ and $p_2 + p_7 \geq p_6$. We note that the function f_5 is a linear function w.r.t. p_7. In order to maximize f_5, we can set one of the inequalities w.r.t. p_7 to be an equality. Then one of the entries in $\boldsymbol{\beta}_{1,4}$ or $\boldsymbol{\beta}_{3,2}$ is zero. With the help of Lemma 66, we prove $F_5 \leq 5/2$. The details are omitted here. ■

APPENDIX F
PROOF OF PROPOSITION 27

In this proof, we prove $\mathrm{CovCHSH}(\boldsymbol{\beta}) \leq 2\sqrt{2}$ for $\tilde{\mathcal{X}}_e = \{(0,0),(0,1),(1,0),(1,1)\}$ first. The proof for $\mathrm{CovCHSH}(\boldsymbol{\beta}) \geq -\sqrt{2}/2$ for the same $\tilde{\mathcal{X}}_e$ is similar and thus are omitted here. Our proof approaches are mainly based on the approaches in [23, Appendix B], where the authors proved

$$\mathrm{CovCHSH}(\boldsymbol{\beta}) \leq 2\sqrt{2}, \qquad \mathrm{CorrCHSH}(\boldsymbol{\beta}) \leq 2\sqrt{2}. \tag{86}$$

Note that they considered a more general setup where ρ, U_1, and U_2 have arbitrary finite size. In our setup, we consider a special set of projection matrices $A_{i,0}$ and $B_{j,0}$, as defined in (28), which results in a stricter bound of $\mathrm{CovCHSH}(\boldsymbol{\beta})$ for $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4)$.

For every $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4)$, there exist ρ, U_1, and U_2 s.t. (34)–(35) hold. We suppose that the ρ, U_1, and U_2 in N_4 are given. Because ρ is a density matrix, it has an eigenvalue decomposition, i.e.,

$$\rho = \sum_{\ell} \lambda_\ell \cdot \boldsymbol{u}_\ell \cdot \boldsymbol{u}_\ell^{\mathsf{H}},$$

where $\sum_\ell$ denotes $\sum_{\ell \in [|\tilde{x}_e|]}$, the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{|\tilde{\mathcal{X}}_e|}$ form an orthonormal basis and are eigenvectors of ρ with correspoding eigenvalues $\lambda_1, \ldots, \lambda_{|\tilde{\mathcal{X}}_e|}$,

$$\sum_{\ell} \lambda_\ell = 1, \qquad 0 \leq \lambda_\ell \leq 1, \qquad \ell \in [|\tilde{x}_e|].$$

Then we define

$$\boldsymbol{\alpha}_{\ell,i} := \sqrt{\lambda_\ell} \cdot \left(\left((A_{i,0}^{\mathsf{H}} \cdot A_{i,0}) \otimes I\right) \cdot \boldsymbol{u}_\ell - \beta_i(0) \cdot \boldsymbol{u}_\ell\right), \qquad \check{\boldsymbol{\alpha}}_{\ell,i} := \frac{\boldsymbol{\alpha}_{\ell,i}}{\sqrt{\sum_{\ell'} \|\boldsymbol{\alpha}_{\ell',i}\|^2}}, \qquad \ell \in [|\tilde{x}_e|],\ i \in \{1,3\}, \tag{87}$$

$$\boldsymbol{\gamma}_{\ell,j} := \sqrt{\lambda_\ell} \cdot \left(\left(I \otimes (B_{j,0}^{\mathsf{H}} \cdot B_{j,0})\right) \cdot \boldsymbol{u}_\ell - \beta_j(0) \cdot \boldsymbol{u}_\ell\right), \qquad \check{\boldsymbol{\gamma}}_{\ell,j} := \frac{\boldsymbol{\gamma}_{\ell,j}}{\sqrt{\sum_{\ell'} \|\boldsymbol{\gamma}_{\ell',j}\|^2}}, \qquad \ell \in [|\tilde{x}_e|],\ j \in \{2,4\}, \tag{88}$$

where I is an identity matrix of size $|\mathcal{X}_e| \times |\mathcal{X}_e|$, the matrices $A_{i,0}$ and $B_{j,0}$ are defined in (28), and the sum $\sum_{\ell'}$ denotes $\sum_{\ell' \in [|\tilde{x}_e|]}$. As shown in (32) and (33), the matrices $A_{i,0}^{\mathsf{H}} \cdot A_{i,0}$ and $B_{j,0}^{\mathsf{H}} \cdot B_{j,0}$ are projection matrices. In particular, when $\boldsymbol{\beta}$ satisfies (34)–(35) for given ρ, U_1, and U_2, it holds that

$$\beta_i(0) = \sum_\ell \lambda_\ell \cdot \boldsymbol{u}_\ell^{\mathsf{H}} \cdot \left((A_{i,0}^{\mathsf{H}} \cdot A_{i,0}) \otimes I\right) \cdot \boldsymbol{u}_\ell, \qquad \beta_j(0) = \sum_\ell \lambda_\ell \cdot \boldsymbol{u}_\ell^{\mathsf{H}} \cdot \left(I \otimes (B_{j,0}^{\mathsf{H}} \cdot B_{j,0})\right) \cdot \boldsymbol{u}_\ell, \qquad \{i,j\} \in \mathcal{K},$$

$$\beta_{i,j}(0,0) = \sum_\ell \lambda_\ell \cdot \boldsymbol{u}_\ell^{\mathsf{H}} \cdot \left((A_{i,0}^{\mathsf{H}} \cdot A_{i,0}) \otimes (B_{j,0}^{\mathsf{H}} \cdot B_{j,0})\right) \cdot \boldsymbol{u}_\ell, \qquad \{i,j\} \in \mathcal{K}.$$

As shown in (32) and (33), the matrices $(A_{i,0}^{\mathsf{H}} \cdot A_{i,0}) \otimes I$, $I \otimes (B_{j,0}^{\mathsf{H}} \cdot B_{j,0})$, and $(A_{i,0}^{\mathsf{H}} \cdot A_{i,0}) \otimes (B_{j,0}^{\mathsf{H}} \cdot B_{j,0})$ are projection matrices as well. When $0 < \beta_\ell(0) < 1$ for $\ell \in [|\tilde{x}_e|]$, we have

$$\sum_\ell \|\boldsymbol{\alpha}_{\ell,i}\|^2 = \beta_i(0)(1-\beta_i(0)) \leq \frac{1}{4}, \qquad \sum_\ell \|\boldsymbol{\gamma}_{\ell,j}\|^2 = \beta_j(0)(1-\beta_j(0)) \leq \frac{1}{4}, \qquad \{i,j\} \in \mathcal{K}, \tag{89}$$

$$\sum_\ell \boldsymbol{\alpha}_{\ell,i}^{\mathsf{H}} \cdot \boldsymbol{\gamma}_{\ell,j} = \beta_{i,j}(0,0) - \beta_i(0) \cdot \beta_j(0) \overset{(a)}{=} \mathrm{Cov}(\boldsymbol{\beta}_{i,j}), \qquad \sum_\ell \check{\boldsymbol{\alpha}}_{\ell,i}^{\mathsf{H}} \cdot \check{\boldsymbol{\gamma}}_{\ell,i} = \mathrm{Corr}(\boldsymbol{\beta}_{i,j}) \qquad \{i,j\} \in \mathcal{K}, \tag{90}$$

where at step (a) we have used the definition of $\mathrm{Cov}(\boldsymbol{\beta}_{i,j})$ in (9). Figs. 9 and 10 illustrate $\beta_{i,j}(x_i, x_j)$ for any x_i and x_j in $\mathcal{X}_e$. By suitably rearranging the function nodes in Fig. 9, we obtain Fig. 10, which is equivalent to Fig. 9. The details of Figs. 9 and 10 are listed as follows:

- After closing the smaller dash box in Figs. 9 and 10, respectively, namely summing over the internal variables in this box, we obtain ρ.
- After closing the larger dash box in Figs. 9 and 10, respectively, we obtain $\beta_{i,j}(x_i, x_j)$.
- We obtain $\beta_{i,j}(0,0)$ by setting $x_i = x_j = 0$ and closing the larger dash box in Figs. 9 and 10, respectively.

Then we have

$$\begin{aligned}
\mathrm{CovCHSH}(\boldsymbol{\beta}) &= \sum_\ell \left(\boldsymbol{\alpha}_{\ell,1} \cdot (\boldsymbol{\gamma}_{\ell,2} + \boldsymbol{\gamma}_{\ell,4}) + \boldsymbol{\alpha}_{\ell,3} \cdot (\boldsymbol{\gamma}_{\ell,2} - \boldsymbol{\gamma}_{\ell,4})\right) \\
&\overset{(a)}{\leq} \sqrt{\left(\sum_{\ell,i} \|\boldsymbol{\alpha}_{\ell,i}\|^2\right) \cdot \left(\sum_\ell \|\boldsymbol{\gamma}_{\ell,2} + \boldsymbol{\gamma}_{\ell,3}\|^2 + \sum_\ell \|\boldsymbol{\gamma}_{\ell,2} - \boldsymbol{\gamma}_{\ell,3}\|^2\right)} \\
&\overset{(b)}{=} \sqrt{\frac{1}{2} \cdot \left(\sum_\ell \left(2\|\boldsymbol{\gamma}_{\ell,2}\|^2 + 2\|\boldsymbol{\gamma}_{\ell,4}\|^2\right)\right)} \\
&\overset{(c)}{\leq} \frac{\sqrt{2}}{2},
\end{aligned}$$

where at step (a) we have used the Cauchy-Schwarz inequality and at steps (b) and (c) we have used the inequalities in (89).

Unlike [37], to prove the covariance-based Tsirelon bound here, we need to show that the inequality (86) holds for arbitrary finite-dimensional quantum system. One can verify that our proof in this appendix also works for $\tilde{\mathcal{X}}_e$ with arbitrary size, i.e., our proof works for arbitrary finite-dimensional quantum systems.

Appendix G
Proof of Proposition 31

We prove Proposition 31 by showing that there exist ρ, U_0, and U_1 in N_4 s.t. the associated $\boldsymbol{\beta}$ satisfying

$$\boldsymbol{\beta}_{1,2} = \boldsymbol{\beta}_{1,4} = \boldsymbol{\beta}_{3,2} = \frac{1}{8}\begin{pmatrix} 2+\sqrt{2} & 2-\sqrt{2} \\ 2-\sqrt{2} & 2+\sqrt{2} \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \frac{1}{8}\begin{pmatrix} 2-\sqrt{2} & 2+\sqrt{2} \\ 2+\sqrt{2} & 2-\sqrt{2} \end{pmatrix}, \tag{91}$$

We suppose that the system N_4 in Fig. 4 is prepared in the state

$$\begin{aligned}\rho &= \frac{1}{4}(U\otimes I)\cdot(I\otimes I+\sigma_1\otimes\sigma_1-\sigma_2\otimes\sigma_2+\sigma_3\otimes\sigma_3)\cdot(U\otimes I)^{\mathsf{H}}\\ &= \frac{1}{2}(U\otimes I)\cdot\begin{pmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{pmatrix}\cdot(U\otimes I)^{\mathsf{H}}\\ &= \frac{1}{8}\begin{pmatrix}2+\sqrt{2}&\sqrt{2}&\sqrt{2}&2+\sqrt{2}\\-\sqrt{2}&2-\sqrt{2}&-2+\sqrt{2}&-\sqrt{2}\\\sqrt{2}&-2+\sqrt{2}&2-\sqrt{2}&\sqrt{2}\\2+\sqrt{2}&-\sqrt{2}&\sqrt{2}&2+\sqrt{2}\end{pmatrix},\end{aligned}$$

where

$$\sigma_1 := \begin{pmatrix}0&1\\1&0\end{pmatrix},\qquad \sigma_2 := \begin{pmatrix}0&-\iota\\\iota&0\end{pmatrix},\qquad \sigma_3 := \begin{pmatrix}-1&0\\0&1\end{pmatrix},\qquad U := \frac{1}{2}\begin{pmatrix}\sqrt{2+\sqrt{2}}&-\sqrt{2-\sqrt{2}}\\\sqrt{2-\sqrt{2}}&\sqrt{2+\sqrt{2}}\end{pmatrix}.$$

By $\mathcal{M}(\mathsf{N}_5)=\mathcal{M}(\mathsf{N}_4)$ in Proposition 33, it is equivalent to consider the ρ_{L}, $\tilde{U}_1$ and $\tilde{U}_2$. The associated ρ_{L} is

$$\rho_{\mathrm{L}} = \frac{1}{2}\cdot\tilde{U}\cdot\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix} = \frac{1}{8}\begin{pmatrix}2+\sqrt{2}&-\sqrt{2}&-\sqrt{2}&2-\sqrt{2}\\\sqrt{2}&2+\sqrt{2}&-2+\sqrt{2}&-\sqrt{2}\\\sqrt{2}&-2+\sqrt{2}&2+\sqrt{2}&-\sqrt{2}\\2-\sqrt{2}&\sqrt{2}&\sqrt{2}&2+\sqrt{2}\end{pmatrix},$$

where $\tilde{U}=U\otimes\overline{U}$. We set the matrices $\tilde{U}_1$ and $\tilde{U}_2$ for the N_5 in Fig. 6 to be

$$\tilde{U}_1 = \left(\tilde{U}^2\right)^{\mathsf{H}},\qquad \tilde{U}_2=\overline{\tilde{U}^2}.$$

Then we have

$$\left(\tilde{U}^2\right)^{\mathsf{H}}\cdot\rho_{\mathrm{L}} = \rho_{\mathrm{L}}\cdot\left(\tilde{U}^2\right)^{\mathsf{H}} = \frac{1}{8}\begin{pmatrix}2+\sqrt{2}&\sqrt{2}&\sqrt{2}&2-\sqrt{2}\\-\sqrt{2}&2+\sqrt{2}&-2+\sqrt{2}&\sqrt{2}\\-\sqrt{2}&-2+\sqrt{2}&2+\sqrt{2}&\sqrt{2}\\2-\sqrt{2}&-\sqrt{2}&-\sqrt{2}&2+\sqrt{2}\end{pmatrix},$$

$$\left(\tilde{U}^2\right)^{\mathsf{H}}\cdot\rho_{\mathrm{L}}\cdot\left(\tilde{U}^2\right)^{\mathsf{H}} = \frac{1}{8}\begin{pmatrix}2-\sqrt{2}&\sqrt{2}&\sqrt{2}&2+\sqrt{2}\\-\sqrt{2}&2-\sqrt{2}&-2-\sqrt{2}&\sqrt{2}\\-\sqrt{2}&-2-\sqrt{2}&2-\sqrt{2}&\sqrt{2}\\2+\sqrt{2}&-\sqrt{2}&-\sqrt{2}&2-\sqrt{2}\end{pmatrix}.$$

The marginals $(\boldsymbol{\beta}_{i,j})_{i,j}$ obtained in (44)–(47) are the same as the marginals $(\boldsymbol{\beta}_{i,j})_{i,j}$ in (91).

APPENDIX H
PROOF OF LEMMA 34

One can verify that $\mathcal{M}(\mathsf{N}_3)\subseteq\mathcal{LM}(\mathcal{K})$. In order to prove that $\mathcal{LM}(\mathcal{K})\setminus\mathcal{M}(\mathsf{N}_3)\neq\emptyset$, we consider the vertex $\boldsymbol{v}_{18}$ of $\mathcal{LM}(\mathcal{K})$ (see Appendix A).

	$\boldsymbol{\beta}_{1,4}$	$\boldsymbol{\beta}_{1,2}$	$\boldsymbol{\beta}_{3,2}$	$\boldsymbol{\beta}_{3,4}$
$\boldsymbol{v}_{18}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}$

By the definition of $\mathcal{M}(\mathsf{N}_3)$ in (5), for any $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$, there exists a joint PMF $\{p_{\mathsf{N}_3}(\boldsymbol{x})\}_{\boldsymbol{x}}$ s.t. $\boldsymbol{\beta}$ computed by (6) satisfies the above table for $(\boldsymbol{\beta}_{i,j})_{i,j}$. In order to have such joint PMF, the valid configurations in $\mathcal{C}(\mathsf{N}_3)$ satisfy

$$x_1 = x_4,\ x_1 \neq x_2,\ x_2 = x_3,\ x_3 = x_4.$$

which leads to a contradiction. Such $\boldsymbol{\beta}$ is not in $\mathcal{M}(\mathsf{N}_3)$.

Appendix I
Proof of Lemma 36

Since ρ is a PSD matrix, the principal minors of ρ are all non-negative, which implies the matrix ρ must have the form

$$\rho = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \alpha & \overline{\rho_{3,2}} & 0 \\ 0 & \rho_{3,2} & 1-\alpha & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since ρ is PSD, we need $|\rho_{3,2}| \leq (\alpha(1-\alpha))^{1/2}$.

Appendix J
Proof of Lemma 37

In this appendix, we prove that the vertex $\boldsymbol{v}_{18}$ is not in $\mathcal{M}(\mathsf{N}_4)$ (see Appendix A). Recall that

	$\boldsymbol{\beta}_{1,4}$	$\boldsymbol{\beta}_{1,2}$	$\boldsymbol{\beta}_{3,2}$	$\boldsymbol{\beta}_{3,4}$
$\boldsymbol{v}_{18}$	$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

(92)

By $\mathcal{M}(\mathsf{N}_5) = \mathcal{M}(\mathsf{N}_4)$ in Proposition 33, we know that proving $\boldsymbol{v}_{17} \in \mathcal{M}(\mathsf{N}_4)$ is equivalent to proving $\boldsymbol{v}_{17} \in \mathcal{M}(\mathsf{N}_5)$. In the rest of the proof, we consider the matrices ρ_{L}, $\tilde{U}_1$, and $\tilde{U}_2$ in N_5.

Lemma 36 implies that in this case, the ρ_{L} can be characterized by $\rho_{3,2}$, i.e.,

$$\rho_{\mathrm{L}} = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \overline{\rho_{3,2}} & 0 \\ 0 & \rho_{3,2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix}, \qquad |\rho_{3,2}| \leq \frac{1}{2},$$

The matrix $\boldsymbol{\beta}_{3,2}$ in (45) and the matrix $\boldsymbol{\beta}_{1,4}$ in (46) equal

$$\boldsymbol{\beta}_{3,2} = \frac{1}{2}\begin{pmatrix} |U_1(0,1)|^2 & |U_1(0,0)|^2 \\ |U_1(0,0)|^2 & |U_1(0,1)|^2 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4}\begin{pmatrix} |U_2(0,1)|^2 & |U_2(0,0)|^2 \\ |U_2(0,0)|^2 & |U_2(0,1)|^2 \end{pmatrix}.$$

In order to have $\boldsymbol{\beta}$ in (92), we need

$$|U_1(0,1)| = 1, \qquad U_1(0,0) = 0, \qquad |U_2(0,1)| = 1, \qquad U_2(0,0) = 0.$$

Then the $\tilde{U}_0$ and $\tilde{U}_1$ in (43) become

$$\tilde{U}_1 = \tilde{U}_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix $\boldsymbol{\beta}_{3,4}$ obtained via (47)satisfies $\beta_{3,4}(x_3, x_4) = \frac{1}{2}[x_3 \neq x_4]$, which is a contradiction of the matrix $\boldsymbol{\beta}_{3,4}$ in $\boldsymbol{v}_{17}$.

APPENDIX K
PROOF OF LEMMA 39

In this appendix, we prove that

$$\mathcal{S}_{3,4,7,10,18,22}(\mathsf{N}_1) := \left\{ \boldsymbol{v} \in \mathcal{LM}(\mathcal{K}) \;\middle|\; \begin{array}{l} \boldsymbol{v} = \alpha_3 \boldsymbol{v}_3 + \alpha_4 \boldsymbol{v}_4 + \alpha_7 \boldsymbol{v}_7 + \alpha_{10} \boldsymbol{v}_{10} + \alpha_{18} \boldsymbol{v}_{18} + \alpha_{23} \boldsymbol{v}_{23}, \\ \alpha_3, \ldots, \alpha_{23} \in \mathbb{R}_{\geq 0},\ \alpha_3 + \alpha_4 + \alpha_7 + \alpha_{10} + \alpha_{18} + \alpha_{23} = 1, \\ 0 < \alpha_3 + \alpha_7 < 1,\ 0 < \alpha_4 + \alpha_{10} < 1 \end{array} \right\} \not\subseteq \mathcal{M}(\mathsf{N}_4).$$

where

	$\boldsymbol{\beta}_{1,4}$	$\boldsymbol{\beta}_{1,2}$	$\boldsymbol{\beta}_{3,2}$	$\boldsymbol{\beta}_{3,4}$		$\boldsymbol{\beta}_{1,4}$	$\boldsymbol{\beta}_{1,2}$	$\boldsymbol{\beta}_{3,2}$	$\boldsymbol{\beta}_{3,4}$
$\boldsymbol{v}_3$	$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\boldsymbol{v}_4$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
$\boldsymbol{v}_7$	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\boldsymbol{v}_{10}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$
$\boldsymbol{v}_{18}$	$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\boldsymbol{v}_{23}$	$\frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

Any $\boldsymbol{\beta} \in \mathcal{S}_{2,3,6,9,17,22}(\mathsf{N}_1)$ satisfies

$$\boldsymbol{\beta}_{1,2} = \begin{pmatrix} 0 & \alpha_4 + \alpha_{10} + \frac{\alpha_{18}+\alpha_{23}}{2} \\ \alpha_3 + \alpha_7 + \frac{\alpha_{18}+\alpha_{23}}{2} & 0 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \begin{pmatrix} \alpha_4 + \frac{\alpha_{18}}{2} & \alpha_{10} + \frac{\alpha_{23}}{2} \\ \alpha_3 + \frac{\alpha_{23}}{2} & \alpha_7 + \frac{\alpha_{18}}{2} \end{pmatrix}, \tag{93}$$

$$\boldsymbol{\beta}_{3,2} = \begin{pmatrix} \alpha_3 + \alpha_7 + \frac{\alpha_{18}+\alpha_{23}}{2} & \alpha_4 + \alpha_{10} \\ 0 & \frac{\alpha_{18}+\alpha_{23}}{2} \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \begin{pmatrix} \alpha_3 + \alpha_4 + \frac{\alpha_{18}}{2} & \alpha_7 + \alpha_{10} + \frac{\alpha_{23}}{2} \\ \frac{\alpha_{23}}{2} & \frac{\alpha_{18}}{2} \end{pmatrix}. \tag{94}$$

By $\mathcal{M}(\mathsf{N}_5) = \mathcal{M}(\mathsf{N}_4)$ in Proposition 33, we know that proving $\mathcal{S}_{2,3,6,9,17,22}(\mathsf{N}_1) \not\subseteq \mathcal{M}(\mathsf{N}_4)$ is equivalent to proving $\mathcal{S}_{2,3,6,9,17,22}(\mathsf{N}_1) \not\subseteq \mathcal{M}(\mathsf{N}_5)$. In the rest of the proof, we consider the ρ_{L}, $\tilde{U}_1$, and $\tilde{U}_3$ in N_5.

In order to set the matrix $\boldsymbol{\beta}_{1,2}$ in (44) equal to the matrix $\boldsymbol{\beta}_{1,2}$ in (93), Lemma 36 implies that the ρ_{L} equals

$$\rho_{\mathrm{L}} = \begin{pmatrix} 0 & 0 & 0 & \alpha_4 + \alpha_{10} + \frac{\alpha_{18}+\alpha_{23}}{2} \\ 0 & 0 & \overline{\rho_{3,2}} & 0 \\ 0 & \rho_{3,2} & 0 & 0 \\ \alpha_3 + \alpha_7 + \frac{\alpha_{18}+\alpha_{23}}{2} & 0 & 0 & 0 \end{pmatrix}, \quad \rho_{3,2} \leq \sqrt{\left(\alpha_4 + \alpha_{10} + \frac{\alpha_{18}+\alpha_{23}}{2}\right) \cdot \left(\alpha_3 + \alpha_7 + \frac{\alpha_{18}+\alpha_{23}}{2}\right)}.$$

By (45), the matrix $\boldsymbol{\beta}_{3,2}$ equals

$$\boldsymbol{\beta}_{3,2} = \begin{pmatrix} \left(\alpha_3 + \alpha_7 + \frac{\alpha_{18}+\alpha_{23}}{2}\right) \cdot |U_1(0,1)|^2 & \left(\alpha_4 + \alpha_{10} + \frac{\alpha_{18}+\alpha_{23}}{2}\right) \cdot |U_1(0,0)|^2 \\ \left(\alpha_3 + \alpha_7 + \frac{\alpha_{18}+\alpha_{23}}{2}\right) \cdot |U_1(0,0)|^2 & \left(\alpha_4 + \alpha_{10} + \frac{\alpha_{18}+\alpha_{23}}{2}\right) \cdot |U_1(0,1)|^2 \end{pmatrix}.$$

In order to obtain $\boldsymbol{\beta}_{3,2}$ in (94), we need

$$\left(\alpha_3 + \alpha_7 + \frac{\alpha_{18}+\alpha_{23}}{2}\right) \cdot |U_1(0,1)|^2 = \alpha_3 + \alpha_7 + \frac{\alpha_{18}+\alpha_{23}}{2}, \qquad \left(\alpha_4 + \alpha_{10} + \frac{\alpha_{18}+\alpha_{23}}{2}\right) \cdot |U_1(0,0)|^2 = \alpha_4 + \alpha_{10},$$

$$\left(\alpha_3 + \alpha_7 + \frac{\alpha_{18}+\alpha_{23}}{2}\right) \cdot |U_1(0,0)|^2 = 0, \qquad \left(\alpha_4 + \alpha_{10} + \frac{\alpha_{18}+\alpha_{23}}{2}\right) \cdot |U_1(0,1)|^2 = \frac{\alpha_{18}+\alpha_{23}}{2}.$$

Since $\alpha_3 + \alpha_7 > 0$, the above expressions imply

$$|U_1(0,1)| = \left(\frac{\alpha_{18}+\alpha_{23}}{2(\alpha_4+\alpha_{10}) + \alpha_{18} + \alpha_{23}}\right)^{1/2} = 1, \qquad |U_1(0,0)| = \left(\frac{2(\alpha_4+\alpha_{10})}{2(\alpha_4+\alpha_{10}) + \alpha_{18} + \alpha_{23}}\right)^{1/2} = 0, \qquad \alpha_4 + \alpha_{10} = 0.$$

However, it contradicts to the requirement $\alpha_4 + \alpha_{10} > 0$ in the definition of $\mathcal{S}_{2,3,6,9,17,22}(\mathsf{N}_1)$.

APPENDIX L
PROOF OF LEMMA 42

In this appendix, we prove that

$$\mathcal{M}(\mathsf{N}_3) \setminus \big((\mathcal{M}(\mathsf{N}_1) \cup \mathcal{M}(\mathsf{N}_4)) \cap \mathcal{M}(\mathsf{N}_3)\big) \neq \emptyset.$$

We consider a joint PMF for $\{p_{\mathsf{N}_3}(\boldsymbol{x})\}_{\boldsymbol{x}}$.

$(\boldsymbol{x})$	$(0,0,1,0)$	$(0,0,0,1)$	$(1,1,1,1)$	Otherwise
$p_{\mathsf{N}_3}(\boldsymbol{x})$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0

The associated $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$ in (6) satisfies

$$\boldsymbol{\beta}_{1,2} = \frac{1}{3}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,2} = \frac{1}{3}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \frac{1}{3}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}. \tag{95}$$

If this $\boldsymbol{\beta}$ is in $\mathcal{M}(\mathsf{N}_1)$, the possible valid configurations in $\mathcal{C}(\mathsf{N}_1)$ s.t. the marginals computed by (6) satisfy the above marginals are

$(\boldsymbol{x})$	$(0,0,1,0)$	$(0,0,0,1)$	$(0,0,1,1)$	$(1,1,1,1)$

To have the marginals in (95), we need

$$p_{\mathsf{N}_1}(0,0,1,0) = \frac{1}{Z(\mathsf{N}_1)} \cdot f_{1,2}(0,0) \cdot f_{3,2}(1,0) \cdot f_{1,4}(0,0) \cdot f_{3,4}(1,0) = \frac{1}{3} > 0,$$
$$p_{\mathsf{N}_1}(0,0,0,1) = \frac{1}{Z(\mathsf{N}_1)} \cdot f_{1,2}(0,0) \cdot f_{3,2}(0,0) \cdot f_{1,4}(0,1) \cdot f_{3,4}(0,1) = \frac{1}{3} > 0,$$
$$p_{\mathsf{N}_1}(1,1,1,1) = \frac{1}{Z(\mathsf{N}_1)} \cdot f_{1,2}(1,1) \cdot f_{3,2}(1,1) \cdot f_{1,4}(1,1) \cdot f_{3,4}(1,1) = \frac{1}{3} > 0,$$

which implies

$$p_{\mathsf{N}_1}(0,0,1,1) = \frac{1}{Z(\mathsf{N}_1)} \cdot f_{1,2}(0,0) \cdot f_{3,2}(1,0) \cdot f_{1,4}(0,1) \cdot f_{3,4}(1,1) > 0.$$

It contradicts the definition of $p_{\mathsf{N}_1}(\boldsymbol{x})$, i.e., $\sum_{\boldsymbol{x}} p_{\mathsf{N}_1}(\boldsymbol{x}) = 1$.

In the rest of this proof, we prove that $\boldsymbol{\beta}$ satisfying (95) is not in $\mathcal{M}(\mathsf{N}_4)$. By $\mathcal{M}(\mathsf{N}_5) = \mathcal{M}(\mathsf{N}_4)$ in Proposition 33, we know that it is equivalent to prove that such $\boldsymbol{\beta}$ is not in $\mathcal{M}(\mathsf{N}_5)$.

In order to have $\boldsymbol{\beta}_{1,2}$ in (100), there are two conditions that ρ_{L} needs to satisfy. One is that $\boldsymbol{\beta}_{1,2}$ obtained from (44) satisfy $\boldsymbol{\beta}_{1,2}$ in (95) and the other is that the associated matrix ρ is a density matrix. The resulting ρ_{L} is

$$\rho_{\mathrm{L}} = \begin{pmatrix} \frac{2}{3} & 0 & 0 & 0 \\ 0 & \overline{\rho_{4,1}} & 0 & 0 \\ 0 & 0 & \rho_{4,1} & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{pmatrix}, \qquad |\rho_{4,1}| \leq \frac{\sqrt{2}}{3}.$$

Suppose that $\boldsymbol{\beta}_{3,2} \in \mathcal{M}(\mathsf{N}_5)$, the matrix $\boldsymbol{\beta}_{3,2}$ computed by (45) equals

$$\boldsymbol{\beta}_{3,2} = \frac{1}{3}\begin{pmatrix} 2|U_1(0,0)|^2 & |U_1(0,1)|^2 \\ 2|U_1(0,1)|^2 & |U_1(0,0)|^2 \end{pmatrix}.$$

The matrix $\boldsymbol{\beta}_{3,2}$ in (95) implies that

$$|U_1(0,1)| = \frac{\sqrt{2}}{2}, \qquad |U_1(0,1)| = 0,$$

which is a contradiction.

APPENDIX M
PROOF OF LEMMA 43

In this appendix, we prove that

$$(\mathcal{M}(\mathsf{N}_3) \cap \mathcal{M}(\mathsf{N}_4)) \setminus \big(\mathcal{M}(\mathsf{N}_1) \cap \mathcal{M}(\mathsf{N}_4)\big) \neq \emptyset.$$

We consider a joint PMF for $\{p_{\mathsf{N}_3}(\boldsymbol{x})\}_{\boldsymbol{x}}$.

$(\boldsymbol{x})$	$(0,0,0,0)$	$(0,0,1,0)$	$(0,0,0,1)$	$(1,1,1,0)$	$(1,1,0,1)$	$(1,1,1,1)$	Otherwise
$p_{\mathsf{N}_3}(\boldsymbol{x})$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	0

(96)

The associated $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_3)$ computed by (6) for N_3 satisfies

$$\boldsymbol{\beta}_{1,2} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \frac{1}{8}\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,2} = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \frac{1}{8}\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}. \tag{97}$$

We show that $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4)$. When we let

$$\rho = \frac{1}{2}\cdot\begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}^{\mathsf{H}}\cdot\begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}, \qquad U_1 = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & -\frac{\sqrt{3}}{3}(1+\sqrt{2}\iota) \\ \frac{\sqrt{3}}{3}(1-\sqrt{2}\iota) & 1 \end{pmatrix}, \qquad U_2 = \frac{1}{2}\begin{pmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{pmatrix},$$

the collection of matrices $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_4)$ computed by (34) and (35) satisfies (97).

We want to show that the valid configurations in (96) are the only possible valid configurations in $\mathcal{C}(\mathsf{N}_3)$ to realize the marginals in (97). The possible valid configurations in $\mathcal{C}(\mathsf{N}_3)$ s.t. the marginals computed by (6) satisfy (97), and the range of the associated probabilities are listed as follows.

$(\boldsymbol{x})$	$p_{\mathsf{N}_3}(\boldsymbol{x})$
$(0,0,0,0)$ or $(1,1,0,0)$	$\leq \boldsymbol{\beta}_{3,4}(0,0) = \frac{1}{8}$
$(0,0,1,1)$ or $(1,1,1,1)$	$\leq \boldsymbol{\beta}_{3,4}(1,1) = \frac{1}{8}$
$(0,0,1,0)$	$\leq \boldsymbol{\beta}_{3,2}(1,0) = \frac{1}{4}$
$(0,0,0,1)$	$\leq \boldsymbol{\beta}_{1,4}(0,1) = \frac{1}{8}$
$(1,1,0,1)$	$\leq \boldsymbol{\beta}_{3,2}(0,1) = \frac{1}{4}$
$(1,1,1,0)$	$\leq \boldsymbol{\beta}_{1,4}(1,0) = \frac{1}{8}$

By the definition of the valid configuration, we have

$$\sum_{\boldsymbol{x}\in\mathcal{C}(\mathsf{N}_3)} p_{\mathsf{N}_3}(\boldsymbol{x}) = 1,$$

which implies

$(\boldsymbol{x})$	$p_{\mathsf{N}_3}(\boldsymbol{x})$
$(0,0,0,0)$ or $(1,1,0,0)$	$\frac{1}{8}$
$(0,0,1,1)$ or $(1,1,1,1)$	$\frac{1}{8}$
$(0,0,1,0)$	$\frac{1}{4}$
$(0,0,0,1)$	$\frac{1}{8}$
$(1,1,0,1)$	$\frac{1}{4}$
$(1,1,1,0)$	$\frac{1}{8}$

To have marginals $\beta_{(2,1)}(0,1) = 1/4$ and $\beta_{(2,1)}(1,0) = 1/4$, we need

$$p_{\mathsf{N}_3}(0,0,1,1) = p_{\mathsf{N}_3}(1,1,0,0) = 0,$$

which implies

$$p_{\mathsf{N}_3}(0,0,0,0) = p_{\mathsf{N}_3}(1,1,1,1) = \frac{1}{8}.$$

In the remaining part of the proof, we prove that $\boldsymbol{\beta}$ satisfying (97) is not in $\mathcal{M}(\mathsf{N}_1)$ by contradiction. To have such marginals in $\mathcal{M}(\mathsf{N}_1)$, we need to set the valid configurations in $\mathcal{C}(\mathsf{N}_1)$ satisfy (96), which means that

$$p_{\mathsf{N}_1}(0,0,0,0) = \frac{1}{Z(\mathsf{N}_1)} \cdot f_{1,2}(0,0) \cdot f_{3,2}(0,0) \cdot f_{1,4}(0,0) \cdot f_{3,4}(0,0) = \frac{1}{8} > 0,$$
$$p_{\mathsf{N}_1}(1,1,0,1) = \frac{1}{Z(\mathsf{N}_1)} \cdot f_{1,2}(1,1) \cdot f_{3,2}(0,1) \cdot f_{1,4}(1,1) \cdot f_{3,4}(0,1) = \frac{1}{4} > 0,$$
$$p_{\mathsf{N}_1}(1,1,1,0) = \frac{1}{Z(\mathsf{N}_1)} \cdot f_{1,2}(1,1) \cdot f_{3,2}(1,1) \cdot f_{1,4}(1,0) \cdot f_{3,4}(1,0) = \frac{1}{8} > 0.$$

However, it implies

$$p_{\mathsf{N}_1}(1,1,0,0) = \frac{1}{Z(\mathsf{N}_1)} \cdot f_{1,2}(1,1) \cdot f_{3,2}(0,1) \cdot f_{1,4}(1,0) \cdot f_{3,4}(0,0) > 0,$$

which contradicts to $p_{\mathsf{N}_1}(1,1,0,0) = 0$ as required in (96).

Appendix N
Proof of Lemma 47

In this appendix, we prove that

$$\mathcal{M}(\mathsf{N}_1) \setminus \big((\mathcal{M}(\mathsf{N}_4) \cup \mathcal{M}(\mathsf{N}_2)) \cap \mathcal{M}(\mathsf{N}_1)\big) \neq \emptyset.$$

For N_1, if we let

$$f_{1,2}(x_1,x_2) = [x_1 = x_2 = 0], \quad f_{3,2}(x_2,x_3) = \frac{1}{2}, \quad f_{1,4}(x_4,x_1) = \frac{1}{2}, \quad f_{3,4}(x_3,x_4) = [x_3 \neq x_4], \quad x_1,\ x_2,\ x_3,\ x_4 \in \mathcal{X}_e,$$

then the associated $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_1)$ obtain by (6) satisfies

$$\boldsymbol{\beta}_{1,2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \boldsymbol{\beta}_{1,4} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,2} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,4} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{98}$$

We prove that $\boldsymbol{\beta}$ in (98) is not in $\mathcal{M}(\mathsf{N}_2)$ first. Suppose that $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_2)$ computed in (20) and (21) having $\boldsymbol{\beta}_{1,2}$, $\boldsymbol{\beta}_{1,4}$, and $\boldsymbol{\beta}_{3,2}$ in (98), we have

$$M_{X_1,X_2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad M_{X_4|X_1} = \begin{pmatrix} \frac{1}{2} & \alpha_1 \\ \frac{1}{2} & 1-\alpha_1 \end{pmatrix}, \qquad M_{X_3|X_2} = \begin{pmatrix} \frac{1}{2} & \alpha_2 \\ \frac{1}{2} & 1-\alpha_2 \end{pmatrix}, \qquad 0 \leq \alpha_1, \alpha_2 \leq 1.$$

Then $\boldsymbol{\beta}_{3,4}$ computed in (21) satisfies

$$\boldsymbol{\beta}_{3,4} = M_{X_3|X_2} \cdot (M_{X_4|X_1} \cdot M_{X_1,X_2})^{\mathsf{T}} = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

which is a contradiction to $\boldsymbol{\beta}_{3,4}$ in (98).

By $\mathcal{M}(\mathsf{N}_5) = \mathcal{M}(\mathsf{N}_4)$ in Proposition 33, we know that it is equivalent to prove that $\boldsymbol{\beta}$ in (98) is not in $\mathcal{M}(\mathsf{N}_5)$. In order to have $\boldsymbol{\beta}_{1,2} = (1,0,0,0)$, there are two conditions that ρ_{L} needs to satisfy. One is that $\boldsymbol{\beta}_{1,2}$ obtained from (44) equals $(1,0,0,0)$ and the other is that the associated matrix ρ is a density matrix. Both these two conditions imply that

$$\rho_{\mathrm{L}}(\tilde{x}_1,\tilde{x}_2) = \big[(\tilde{x}_1,\tilde{x}_2) = (\tilde{0},\tilde{0})\big], \qquad \tilde{x}_1,\tilde{x}_2 \in \tilde{\mathcal{X}}_e,$$

Then $\boldsymbol{\beta}_{1,4}$ and $\boldsymbol{\beta}_{3,2}$ obtained via (45) and (46) equal

$$\boldsymbol{\beta}_{1,4} = \begin{pmatrix} |U_2(0,0)|^2 & |U_2(0,1)|^2 \\ 0 & 0 \end{pmatrix}, \qquad \boldsymbol{\beta}_{3,2} = \begin{pmatrix} |U_1(0,0)|^2 & 0 \\ |U_1(0,1)|^2 & 0 \end{pmatrix}.$$

In order to have $\boldsymbol{\beta}_{1,4}$ and $\boldsymbol{\beta}_{3,2}$ in (98), we need

$$|U_1(0,0)|^2 = |U_1(0,1)|^2 = |U_2(0,0)|^2 = |U_1(0,1)|^2 = \frac{1}{2}.$$

Then $\boldsymbol{\beta}_{3,4}$ computed by (47) satisfies $\boldsymbol{\beta}_{3,4}(x_3,x_4) = 1/4$ for all $x_3, x_4 \in \mathcal{X}_e$ and it contradicts to $\boldsymbol{\beta}_{3,4}$ in (98).

APPENDIX O
PROOF OF LEMMA 48

In this appendix, we prove that

$$\mathcal{M}(\mathsf{N}_2) \subsetneq \mathcal{M}(\mathsf{N}_1), \quad \mathcal{M}(\mathsf{N}_1) \cap \mathcal{M}(\mathsf{N}_4) \setminus (\mathcal{M}(\mathsf{N}_2) \cap \mathcal{M}(\mathsf{N}_1) \cap \mathcal{M}(\mathsf{N}_4)) \neq \emptyset.$$

We prove $\mathcal{M}(\mathsf{N}_2) \subseteq \mathcal{M}(\mathsf{N}_1)$ first. For any $\boldsymbol{\beta}$ in $\mathcal{M}(\mathsf{N}_2)$, there exist three matrices M_{X_1,X_2}, $M_{X_3|X_2}$ and $M_{X_4|X_1}$ s.t. (20)–(21) hold. If we let

$$f_{1,2}(x_1,x_2) = M_{X_1,X_2}(x_1,x_2), \quad f_{3,2}(x_3,x_2) = M_{X_3|X_2}(x_3,x_2), \qquad x_1,\ x_2,\ x_3 \in \mathcal{X}_e,$$
$$f_{1,4}(x_4,x_1) = M_{X_4|X_1}(x_4,x_1), \quad f_{3,4}(x_3,x_4) = 1, \qquad x_1,\ x_3,\ x_4 \in \mathcal{X}_e.$$

for the S-FG N_1 in Fig. 1, then $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_1)$ computed by (6) is the same as $\boldsymbol{\beta}$ in (20)–(21).

In the rest of this appendix, we prove that there is a point in $\mathcal{M}(\mathsf{N}_1) \cap \mathcal{M}(\mathsf{N}_4)$ that is not in $\mathcal{M}(\mathsf{N}_2)$. On one hand, if we let

$$f_{1,2}(x_1,x_2) = \frac{[x_1 = x_2]}{2}, \quad f_{3,2}(x_3,x_2) = \frac{1}{4}, \quad f_{1,4}(x_4,x_1) = \frac{1}{4}, \quad f_{3,4}(x_3,x_4) = \frac{[x_3 = x_4]}{2}, \quad x_1,\ x_2,\ x_3,\ x_4 \in \mathcal{X}_e,$$

for N_1 and

$$U_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \qquad U_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \qquad \rho = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix},$$

for N_4, then $\boldsymbol{\beta}$ computed in (6) and (34)–(35), respectively, are the same. In particular, the collection of matrices $\boldsymbol{\beta}$ satisfy

$$\beta_{1,2}(x_1,x_2) = \frac{1}{2}[x_1 = x_2], \qquad \beta_{1,4}(x_1,x_4) = \frac{1}{4}, \qquad \beta_{3,2}(x_1,x_4) = \frac{1}{4}, \qquad \beta_{3,4}(x_1,x_4) = \frac{1}{2}[x_3 = x_4], \quad x_1,\ldots,x_4 \in \mathcal{X}_e. \tag{99}$$

On the other hand, to have $\boldsymbol{\beta}_{1,4}$, $\boldsymbol{\beta}_{1,2}$, and $\boldsymbol{\beta}_{3,2}$ in (99), the expressions (20)–(21) for computing $\boldsymbol{\beta}$ in $\mathcal{M}(\mathsf{N}_2)$ imply that the matrices M_{X_1,X_2}, $M_{X_3|X_2}$ and $M_{X_4|X_1}$ in N_2 satisfy

$$M_{X_1,X_2}(x_1,x_2) = \frac{[x_1 = x_2]}{2}, \qquad M_{X_3|X_2}(x_3,x_2) = \frac{1}{2}, \qquad M_{X_4|X_1}(x_4,x_1) = \frac{1}{2}, \qquad x_1,\ldots,x_4 \in \mathcal{X}_e.$$

Then $\boldsymbol{\beta}_{3,4}$ in (21) satisfies $\boldsymbol{\beta}_{3,4}(x_3,x_4) = 1/4$ for all $x_3, x_4 \in \mathcal{X}_e$, which is a contradiction to $\boldsymbol{\beta}_{3,4}$ in (99).

APPENDIX P
PROOF OF LEMMA 46

In this appendix, we prove

$$\mathcal{M}(\mathsf{N}_2) \setminus (\mathcal{M}(\mathsf{N}_4) \cap \mathcal{M}(\mathsf{N}_2)) \neq \emptyset.$$

by proving that any $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_2)$ with

$$\beta_{1,2}(x_1,x_2) = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \beta_{3,2} = \frac{1}{2} \begin{pmatrix} \alpha_1 & \alpha_2 \\ 1-\alpha_1 & 1-\alpha_2 \end{pmatrix}, \qquad \alpha_1 \neq 1-\alpha_2,\ 0 \leq \alpha_1, \alpha_2 \leq 1 \tag{100}$$

is not in $\mathcal{M}(\mathsf{N}_4)$.

We firstly prove that there exists a $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_2)$ satisfying (100). If we let

$$M_{X_1,X_2} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad M_{X_3|X_2} = \begin{pmatrix} \alpha_1 & \alpha_2 \\ 1-\alpha_1 & 1-\alpha_2 \end{pmatrix}, \qquad \alpha_1 \neq 1-\alpha_2,\ 0 \leq \alpha_1, \alpha_2 \leq 1,$$

in N_2, then $\boldsymbol{\beta} \in \mathcal{M}(\mathsf{N}_2)$ computed by (20) and (21) satisfies (100).

In the rest of this appendix, we prove that $\boldsymbol{\beta}$ satisfying (100) is not in $\mathcal{M}(\mathsf{N}_4)$. By $\mathcal{M}(\mathsf{N}_5) = \mathcal{M}(\mathsf{N}_4)$ in Proposition 33, we know that it is equivalent to prove that this $\boldsymbol{\beta}$ is not in $\mathcal{M}(\mathsf{N}_5)$.

In order to have $\boldsymbol{\beta}_{1,2}$ in (100), there are two conditions that ρ_{L} needs to satisfy. One is that $\boldsymbol{\beta}_{1,2}$ obtained from (44) satisfy $\boldsymbol{\beta}_{1,2}(x_1, x_2) = 1/2 \cdot [x_1 = x_2]$ for all $x_1, x_2 \in \mathcal{X}_e$ and the other is that the associated matrix ρ is a density matrix. The resulting ρ_{L} is

$$\rho_{\mathrm{L}} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & \overline{\rho_{4,1}} & 0 & 0 \\ 0 & 0 & \rho_{4,1} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}, \qquad |\rho_{4,1}| \leq \frac{1}{2}.$$

Suppose that $\boldsymbol{\beta}_{3,2} \in \mathcal{M}(\mathsf{N}_5)$, the matrix $\boldsymbol{\beta}_{3,2}$ computed by (45) equals

$$\boldsymbol{\beta}_{3,2} = \frac{1}{2} \begin{pmatrix} |U_1(0,0)|^2 & |U_1(0,1)|^2 \\ |U_1(0,1)|^2 & |U_1(0,0)|^2 \end{pmatrix}.$$

Because this $\boldsymbol{\beta}_{3,2}$ satisfies (100), we need

$$|U_1(0,0)|^2 = \alpha_1 = 1 - \alpha_2,$$

which is a contradiction to $\alpha_1 \neq 1 - \alpha_2$.

REFERENCES

[1] F. R. Kschischang, B. J. Frey, and H.-. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[2] G. D. Forney, “Codes on graphs: normal realizations,” *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 520–548, Feb. 2001.

[3] H.-A. Loeliger, “An introduction to factor graphs,” *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.

[4] H. Wymeersch, *Iterative Receiver Design*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[5] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[6] M. Mézard and A. Montanari, *Information, Physics and Computation*. Oxford, U.K.: Oxford Univ. Press, 2009.

[7] H.-A. Loeliger and P. O. Vontobel, “Factor graphs for quantum probabilities,” *IEEE Trans. Inf. Theory*, vol. 63, no. 9, pp. 5642–5665, Sep. 2017.

[8] H. Loeliger and P. O. Vontobel, “Quantum measurement as marginalization and nested quantum systems,” *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3485–3499, Jun. 2020.

[9] M. Gell-Mann and J. B. Hartle, “quantum mechanics in the light of quantum cosmology,” in *Proc. Santa Fe Institute Workshop on Com-plexity, Entropy, and the Physics of Information*, May 1989.

[10] H. F. Dowker and J. J. Halliwell, “Quantum mechanics of history: The decoherence functional in quantum mechanics,” *Phys. Rev. D*, vol. 46, pp. 1580–1609, Aug. 1992.

[11] R. B. Griffiths, *Consistent Quantum Theory*. Cambridge Univ. Press, 2002.

[12] L. Hardy, “Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories,” *Phys. Rev. Lett.*, vol. 68, pp. 2981–2984, May 1992.

[13] D. Frauchiger and R. Renner, “Quantum theory cannot consistently describe the use of itself,” *Nature Communications*, vol. 9, no. 3711, 2018.

[14] N. Gisin, *Quantum Chance: Nonlocality, Teleportation and Other Quantum Marvels*. USA: Copernicus, 2014.

[15] M. J. Wainwright and M. I. Jordan, “Graphical models, exponential families, and variational inference,” *Foundation and Trends Machine Learning*, vol. 1, no. 1–2, pp. 1–305, 2008.

[16] J. S. Yedidia, W. T. Freeman, and Y. Weiss, “Constructing free-energy approximations and generalized belief propagation algorithms,” *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2282–2312, Jul. 2005.

[17] P. O. Vontobel, “Counting in graph covers: A combinatorial characterization of the Bethe entropy function,” *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 6018–6048, Sep. 2013.

[18] J. S. Bell, “On the Einstein Podolsky Rosen paradox,” *Physics*, vol. 1, pp. 195–200, Nov. 1964.

[19] I. Pitowsky and K. Svozil, “Optimal tests of quantum nonlocality,” *Phys. Rev. A*, vol. 64, no. 1, Jun. 2001.

[20] D. Collins and N. Gisin, “A relevant two qubit Bell inequality inequivalent to the CHSH inequality,” *J. Phys. A*, vol. 37, no. 5, pp. 1775–1787, Jan. 2004.

[21] C. Śliwa, “Symmetries of the Bell correlation inequalities,” *Phys. Lett. A*, vol. 317, no. 3, pp. 165 – 168, Oct. 2003.

[22] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, “Proposed experiment to test local hidden-variable theories,” *Phys. Rev. Lett.*, vol. 23, pp. 880–884, Oct. 1969.

[23] V. Pozsgay, F. Hirsch, C. Branciard, and N. Brunner, “Covariance Bell inequalities,” *Phys. Rev. A*, vol. 96, p. 062128, Dec. 2017.

[24] B. S. Tsirel'son, "Quantum analogues of the Bell inequalities. the case of two spatially separated domains," *J. Sov. Math.*, vol. 36, no. 4, pp. 557–570, Feb. 1987.

[25] D. Bertsekas, *Nonlinear Programming*, ser. Athena scientific optimization and computation series. Athena Scientific, 2016.

[26] A. Fine, "Hidden variables, joint probability, and the Bell inequalities," *Phys. Rev. Lett.*, vol. 48, pp. 291–295, Feb. 1982. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.48.291

[27] R. Mori, "Loop calculus for non-binary alphabets using concepts from information geometry," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1887–1904, Apr. 2015.

[28] L. J. Landau, "Empirical two-point correlation functions," *Found. Phys.*, vol. 18, no. 4, pp. 449–460, Apr. 1988.

[29] M. X. Cao and P. O. Vontobel, "Double-edge factor graphs: Definition, properties, and examples," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017, pp. 136–140.

[30] A. Horn, "Doubly stochastic matrices and the diagonal of a rotation matrix," *American Journal of Mathematics*, vol. 76, no. 3, pp. 620–630, 1954.

[31] F. Verstraete, J. Dehaene, and B. DeMoor, "Local filtering operations on two qubits," *Phys. Rev. A*, vol. 64, p. 010101, Jun. 2001.

[32] A. Jaffe, "Lorentz transformations rotations and boosts," Nov. 2013. [Online]. Available: http://home.ku.edu.tr/~amostafazadeh/phys517_518/phys517_2016f/Handouts/A_Jaffi_Lorentz_Group.pdf

[33] R. Horodecki and M. Horodecki, "Information-theoretic aspects of quantum inseparability of mixed states," *Phys. Rev. A*, vol. 54, p. 1838, 1996.

[34] F. Verstraete, J. Dehaene, and B. D. Moor, "Lorentz singular-value decomposition and its applications to pure states of three qubits," *Phys. Rev. A*, vol. 65, p. 032308, Feb. 2002.

[35] D. Avis, "lrs: A revised implementation of the reverse search vertex enumeration algorithm," Jan. 1999. [Online]. Available: http://cgm.cs.mcgill.ca/~avis/doc/avis/Av98a.pdf

[36] D. Avis and K. Fukuda, "A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra," *Discrete and Computational Geometry*, vol. 8, pp. 295–313, Sep. 1992.

[37] L. Masanes, "Extremal quantum correlations for n parties with two dichotomic observables per site," 2005. [Online]. Available: https://arxiv.org/abs/quant-ph/0512100