Edge-Weighted Hypergraph Transversals & Contextuality

Thomas C. Fraser[1,2,*]

[1] *Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada, N2L 2Y5*
[2] *University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1*

(Dated: September 13, 2017)

This is the abstract.

CONTENTS

I. INTRODUCTION

A. Applications

II. MARGINAL SATISFIABILITY

A. Definitions

To every random variable[1] v there corresponds a prescribed set of **outcomes** $\mathcal{O}_v$ and a set of **events over** v denoted $\Omega(v)$ corresponding to the set of all functions of the form $\omega : \{v\} \to \mathcal{O}_v$. Evidently, $\Omega(v)$ and $\mathcal{O}_v$ are isomorphic structures and their distinction can be confounding. There is rarely any harm in referring synonymously to either as outcomes. Nonetheless, a sheaf-theoretic treatment of contextuality [1] demands the distinction. Specifically for this work, the distinction becomes essential for the exploitation of marginal symmetries in Section III D. As a natural generalization we define the event over a collection of random variables $V = \{v_1, \dots, v_n\}$ in a parallel manner:

$$\Omega(V) \equiv \{\omega : V \to \mathcal{O}_V \mid \forall v \in V, \omega(v) \in \mathcal{O}_v\}$$

Furthermore, the **domain** $\mathcal{D}(\omega)$ of an event ω is the set of random variables it valuates, i.e. if $\omega \in \Omega(V)$ then $\mathcal{D}(\omega) = V$.

For every $V' \subset V$ and $\omega \in \Omega(V)$, the **restriction of** ω **onto** V' (denoted $\omega|_{V'}$) corresponds to the unique event in $\Omega(V')$ that agrees with ω for all valuations of variables in V', i.e. $\forall v' \in V' : \omega|_{V'}(v') = \omega(v')$. Using this notational framework, a probability distribution or simply **distribution** p_V is a probability measure on $\Omega(V)$, assigning to each $\omega \in \Omega(V)$ a real number $\mathsf{p}_V(\omega) \in [0,1]$ such that $\sum_{\omega\in\Omega(V)} \mathsf{p}_V(\omega) = 1$. The set of all distributions over $\Omega(V)$ is denoted $\mathcal{P}_V$. Moreover, given $\mathsf{p}_V \in \mathcal{P}_V$ and $V' \subset V$, there is an induced distribution $\mathsf{p}_V|_{V'} \in \mathcal{P}_{V'}$ obtained by *marginalizing* p_V:

$$\mathsf{p}_V|_{V'}(\omega') = \sum_{\substack{\omega\in\Omega(V) \\ \omega|_{V'}=\omega'}} \mathsf{p}_V(\omega) \tag{1}$$

Presently, the reader is equipped with sufficient notation and terminology to comprehend the marginal problem.

Definition 1. The Marginal Problem: Given a collection of m distributions $\{\mathsf{p}_{V_1}, \dots, \mathsf{p}_{V_m}\}$, does there exist a distribution $\mathsf{p}_\Lambda \in \mathcal{P}_\Lambda$ with $\Lambda \equiv \bigcup_{i=1}^m V_m$ such that $\forall i : \mathsf{p}_\Lambda|_{V_i} = \mathsf{p}_{V_i}$?

To facilitate further discussion of this problem, several pieces of nomenclature will be introduced. First, the set $\mathcal{V} = \{V_1, \dots, V_m\}$ is called the **marginal scenario** while its elements are called the **marginal contexts**. The collection of distributions $\mathsf{p}_{*\mathcal{V}} \equiv \{\mathsf{p}_{V_1}, \dots, \mathsf{p}_{V_m}\}$[2] is

* tcfraser@tcfraser.com

[1] Throughout this document, it is assumed that all random variables are discrete and have finite cardinality.

[2] The subscript $*$ preceding ${}_*\mathcal{V}$ is added for clarity; $\mathsf{p}_{*\mathcal{V}}$ is *not* a distribution but a set of distributions over $\mathcal{V}$. The ${}_*\mathcal{V}$ convention is adopted throughout this report.

called the **marginal model** [2][3]. The distribution p_Λ, if it exists, is termed the **joint distribution**. Strictly speaking, as defined by [2], a marginal scenario forms an *abstract simplicial complex*, meaning it satisfies the supplementary requirement that all subsets of contexts are also contexts, i.e. $\forall V \in \mathcal{V} : V' \subset V \implies V' \in \mathcal{V}$. Throughout this work, we exclusively consider (without loss of generality) *maximal* marginal scenarios, restricting our focus to the contexts which are contained in no others. Finally, a marginal model $\mathsf{p}_{*\mathcal{V}}$ is said to be **non-contextual**, and will be denoted $\mathsf{p}_{*\mathcal{V}} \in \mathcal{N} \subseteq \mathcal{P}_{*\mathcal{V}}$ if it admits a joint distribution and **contextual** otherwise ($\mathsf{p}_{*\mathcal{V}} \in \bar{\mathcal{N}}$). Equipped with this additional terminology and notation, the marginal problem now reads: given $\mathsf{p}_{*\mathcal{V}}$, is $\mathsf{p}_{*\mathcal{V}} \in \mathcal{N}$ or not?

B. Linearity

An essential feature of the marginal problem is linearity; the marginalization of p_Λ onto the marginal contexts $\{\mathsf{p}_\Lambda|_V \mid V \in \mathcal{V}\}$ is a linear transformation, requiring only the summations pursuant to Eq. (1). Consequently, it is advantageous to consider the statement of the marginal problem as a matrix multiplication. To this end, for each marginal scenario $\mathcal{V}$ we define a binary matrix $\mathcal{M}$ called the **incidence matrix** which implements this mapping. The columns of $\mathcal{M}$ are indexed by *joint events* $j \in \Omega(\Lambda)$ and the rows are indexed by *marginal events* $m \in \Omega(V)$ for some $V \in \mathcal{V}$. By deliberate abuse of notation, we will denote the set of all marginal events as $\Omega({*\mathcal{V}})$ and is defined as the following disjoint union:

$$\Omega({*\mathcal{V}}) \equiv \coprod_{V \in \mathcal{V}} \Omega(V)$$

The $|\Omega({*\mathcal{V}})| \times |\Omega(\Lambda)|$ incidence matrix $\mathcal{M}$ is then defined element-wise for $m \in \Omega({*\mathcal{V}})$ and $j \in \Omega(\Lambda)$:

$$\mathcal{M}_j^m = \begin{cases} 1 & j|_{\mathcal{D}(m)} = m \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Conceptually, the entries of this matrix are populated with ones whenever the marginal event (row) m is the restriction of some joint event (column) j. For a given marginal scenario $\mathcal{V}$, $\mathcal{M}$ represents the tuple of restriction maps $\mathcal{M} : \Omega(\Lambda) \to \prod_{V \in \mathcal{V}} \Omega(V) :: j \mapsto \{j|_V \mid V \in \mathcal{V}\}$ [1]. Furthermore, note that the component indices of $\mathcal{M}$ in Eq. (2) are deliberately separated. Among other reasons, this is done to allow one to denote the m-th row of $\mathcal{M}$ as $\mathcal{M}^m$ and the j-th column as $\mathcal{M}_j$. For further notational convenience, since $\mathcal{M}$ is a binary matrix, we let $\mathcal{M}^m$ and $\mathcal{M}_j$ analogously correspond their respective *supports*[4], e.g. $m \in \sigma(\mathcal{M}_j)$ if and only if $\mathcal{M}_j^m = 1$. Throughout the remainder of this report, the utility of the incidence matrix $\mathcal{M}$ will be indispensable.

To illustrate this concretely, consider the following example. Let Λ be 3 binary variables $\{a, b, c\}$ and $\mathcal{V}$ be the marginal scenario $\mathcal{V} = \{\{a, b\}, \{b, c\}, \{a, c\}\}$. The incidence matrix for $\mathcal{V}$ becomes:

$$\begin{array}{cc}
(a,b,c) \mapsto & \begin{array}{cccccccc} (0,0,0) & (0,0,1) & (0,1,0) & (0,1,1) & (1,0,0) & (1,0,1) & (1,1,0) & (1,1,1) \end{array} \\
\begin{array}{c}
(a\mapsto 0, b\mapsto 0) \\ (a\mapsto 0, b\mapsto 1) \\ (a\mapsto 1, b\mapsto 0) \\ (a\mapsto 1, b\mapsto 1) \\ (b\mapsto 0, c\mapsto 0) \\ (b\mapsto 0, c\mapsto 1) \\ (b\mapsto 1, c\mapsto 0) \\ (b\mapsto 1, c\mapsto 1) \\ (a\mapsto 0, c\mapsto 0) \\ (a\mapsto 0, c\mapsto 1) \\ (a\mapsto 1, c\mapsto 0) \\ (a\mapsto 1, c\mapsto 1)
\end{array} &
\begin{pmatrix}
\mathbf{1} & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \mathbf{1} & \mathbf{1} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{1} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{1} \\
\mathbf{1} & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\
0 & \mathbf{1} & 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\
0 & 0 & \mathbf{1} & 0 & 0 & 0 & \mathbf{1} & 0 \\
0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & \mathbf{1} \\
\mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\
0 & \mathbf{1} & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \mathbf{1} & 0 & \mathbf{1} & 0 \\
0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & \mathbf{1}
\end{pmatrix}
\end{array} \tag{3}$$

In addition, for any joint distribution $\mathsf{p}_\Lambda \in \mathcal{P}_\Lambda$ we associate a joint distribution *vector* p_Λ (identically denoted) indexed by $j \in \Omega(\Lambda)$, i.e. $\mathsf{p}_\Lambda^j \equiv \mathsf{p}_\Lambda(j)$. Analogously, for each marginal model $\mathsf{p}_{*\mathcal{V}} \in \mathcal{P}_{*\mathcal{V}}$ there is an associated marginal distribution *vector* $\mathsf{p}_{*\mathcal{V}}$ indexed by $m \in \Omega({*\mathcal{V}})$ such that $\mathsf{p}_{*\mathcal{V}}^m \equiv \mathsf{p}_{\mathcal{D}(m)}(m)$. Using these vectors, the marginal problem becomes the following linear program:

Definition 2. **The Marginal Linear Program (MLP)**:

$$\begin{aligned} \text{minimize:} \quad & \emptyset \cdot \mathsf{p}_\Lambda{}^{5} \\ \text{subject to:} \quad & \mathsf{p}_\Lambda \succeq 0 \\ & \mathcal{M} \cdot \mathsf{p}_\Lambda = \mathsf{p}_{*\mathcal{V}} \end{aligned} \tag{4}$$

As such, $\mathsf{p}_{*\mathcal{V}} \in \mathcal{N}$ if and only if MLP is a *feasible* linear program. Importantly, if MLP is feasible, it will return the joint distribution p_Λ. To every linear program, there exists a dual linear program that characterizes the feasibility of the original [3]. Constructing the dual linear program is a well-defined procedure [4].

Definition 3. **The Dual Marginal Linear Program (DMLP)**:

$$\begin{aligned} \text{minimize:} \quad & \gamma \cdot \mathsf{p}_{*\mathcal{V}} \\ \text{subject to:} \quad & \gamma \cdot \mathcal{M} \succeq 0 \end{aligned}$$

By construction, DMLP completely determines the whether or not $\mathsf{p}_{*\mathcal{V}} \in \mathcal{N}$ or not. If $\mathsf{p}_{*\mathcal{V}} \in \mathcal{N}$, then MLP is feasible and the following holds,

$$\gamma \cdot \mathsf{p}_{*\mathcal{V}} = \gamma \cdot (\mathcal{M} \cdot \mathsf{p}_\Lambda) \geq 0 \tag{5}$$

[3] In [1], $\mathsf{p}_{*\mathcal{V}}$ is instead called an *empirical model*.

[4] The *support* $\sigma(f)$ of a mapping f is the subset of its domain $\mathcal{D}(f)$ that is not mapped to a zero element: $\sigma(f) = \{x \in \mathcal{D}(f) \mid f(x) \neq 0\}$.

[5] Note that the primal value of the linear program is of no interest, all that matters is its *feasibility*. Here $\emptyset$ denotes a null vector of all zero entries.

because both $\gamma \cdot \mathcal{M} \succeq 0$ and $\mathsf{p}_\Lambda \succeq 0$. If however, $\gamma \cdot \mathsf{p}_{*\mathcal{V}} < 0$, then Eq. (5) is violated and $\mathsf{p}_{*\mathcal{V}} \notin \mathcal{N}$[6]. In summary, the sign of $d \equiv \min(\gamma \cdot \mathsf{p}_{*\mathcal{V}})$ answers the marginal program; $\mathsf{p}_{*\mathcal{V}} \in \mathcal{N}$ if and only if $d \geq 0$[7].

Corollary 1. *All linear, homogeneous constraints $\gamma \cdot \mathsf{p}_{*\mathcal{V}} \geq 0$ constraining non-contextual marginal models $\mathcal{N} \subseteq \mathcal{P}_{*\mathcal{V}}$ satisfy $\gamma \cdot \mathcal{M} \succeq 0$. Moreover, all vectors γ satisfying $\gamma \cdot \mathcal{M} \succeq 0$ correspond to valid constraints $\gamma \cdot \mathsf{p}_{*\mathcal{V}} \geq 0$ for $\mathcal{N} \subseteq \mathcal{P}_{*\mathcal{V}}$.*

In light of Definitions 2 and 3, when supplied with a particular marginal model $\mathsf{p}_{*\mathcal{V}}$, the marginal problem can be solved computationally by evaluating DMLP to determine the feasibility of MLP. A more difficult variant of the marginal problem is one wherein no particular marginal model is supplied.

Definition 4. The General Marginal Problem (GMP): Given a marginal scenario $\mathcal{V}$ find a set of independent constraints Γ which completely constraint $\mathcal{N} \subseteq \mathcal{P}_{*\mathcal{V}}$; i.e. $\mathsf{p}_{*\mathcal{V}} \in \mathcal{N}$ if and only if it satisfies all constraints in Γ.

The remainder of this paper is concerned with methods for solving (or partially solving) GMP. Specifically, Section II C discusses existing methods for completely solving GMP and outlines some of their disadvantages. Section II D summarizes an existing method for completely solving a possibilistic variant of GMP. Sections II C and II D motivate Section III, wherein a new method for completely solving GMP is presented.

C. Marginal Polytopes

The complete space of marginal models over $\mathcal{V}$ (denoted $\mathcal{P}_{*\mathcal{V}}$) can be partitioned into two spaces: the contextual marginal models ($\bar{\mathcal{N}}$) and the non-contextual marginal models ($\mathcal{N} \equiv \mathcal{P}_{*\mathcal{V}} \setminus \bar{\mathcal{N}}$). Pitowsky [7] demonstrates that $\mathcal{N}$ forms a *convex* polytope commonly referred to as the **marginal polytope** for $\mathcal{V}$. When embedded in $\mathbb{R}^{|\Omega(_*\mathcal{V})|}$, the extremal rays of the marginal polytope correspond to the columns of $\mathcal{M}$ which further correspond to all *deterministic* joint distributions $\mathsf{p}_\Lambda \in \mathcal{P}_\Lambda$[8]. The normalization of p_Λ ($\sum_j \mathsf{p}_\Lambda^j = 1$) defines the convexity of the polytope; each marginal model $\mathsf{p}_{*\mathcal{V}} \in \mathcal{P}_{*\mathcal{V}}$ must be a convex mixture of the deterministic marginal models pursuant to Eq. (4). Consequently, characterizing the contextuality of marginal models is manifestly a problem of polytope description. Notably, the **facets** of a marginal polytope correspond to a finite set of linear inequalities that are complete in the sense that all contextual distributions violate at least one facet inequality [8]. From the perspective of a marginal polytope, convex hull algorithms or linear quantifier elimination can be used to compute a representation of the complete set of facet inequalities and consequently completely solve the GMP. A popular tool for linear quantifier elimination is *Fourier-Motzkin elimination* [9–12]. In this report, we will avoid expounding upon the Fourier-Motzkin procedure and instead recall a few of its notable features and consequences.[9]

Definition 5. [3, Section 12.2] Given a system of linear inequality constraints $\mathcal{S} = \{A \cdot x \leq b\}$ constraining some free variables x, the **Fourier-Motzkin elimination** procedure eliminates some of the variables in x and returns a system of linear inequality constraints $\mathcal{S}' = \{A' \cdot x' \leq b'\}$ over $x' \subset x$ such that any solution x' of $\mathcal{S}'$ will permit at least one compatible solution x of $\mathcal{S}$ (and vice versa).

$$\exists x' : A' \cdot x' \leq b' \iff \exists x : A \cdot x \leq b \tag{6}$$

In particular, the following system of linear inequalities defines the marginal polytope for $\mathcal{V}$:

$$\begin{aligned}
\forall m \in \Omega(_*\mathcal{V}) : \quad & \mathsf{p}_{*\mathcal{V}}^m - \sum\nolimits_j \mathcal{M}_j^m \mathsf{p}_\Lambda^j \geq 0 \\
\forall m \in \Omega(_*\mathcal{V}) : \quad & -\mathsf{p}_{*\mathcal{V}}^m + \sum\nolimits_j \mathcal{M}_j^m \mathsf{p}_\Lambda^j \geq 0 \\
\forall j \in \Omega(\Lambda) : \quad & \mathsf{p}_\Lambda^j \geq 0 \\
& \sum\nolimits_j \mathsf{p}_\Lambda^j \geq 1 \\
& -\sum\nolimits_j \mathsf{p}_\Lambda^j \geq -1
\end{aligned} \tag{7}$$

Using the Fourier-Motzkin elimination procedure, it is possible to eliminate the variables p_Λ^j relating to joint events and obtain a system of linear inequalities constraining only marginal events $\mathsf{p}_{*\mathcal{V}}^m$ which completely characterizes the set of non-contextual marginal models $\mathcal{N} \subseteq \mathcal{P}_{*\mathcal{V}}$.

Lemma 1. [10] *There exists a finite set of integral vectors Γ such that for all $\mathsf{p}_{*\mathcal{V}} \in \mathcal{P}_{*\mathcal{V}}$:*

$$\mathsf{p}_{*\mathcal{V}} \in \mathcal{N} \iff \forall \gamma \in \Gamma : \gamma \cdot \mathsf{p}_{*\mathcal{V}} \geq 0 \tag{8}$$

Proof. The finiteness and existence of Γ is a fundamental property of polytopes [9, 11–14]. The fact that each vector $\gamma \in \Gamma$ need only be integer-valued follows from

---

[6] These observations are collectively a consequence of Farkas's Lemma [5].

[7] In particular, if $d \geq 0$, then $d = 0$ due to the existence of the trivial solution $\gamma = \emptyset$. This observation is an instance of the *Complementary Slackness Property* [6]. Alternatively, if $d < 0$, then it is unbounded $d = -\infty$ due to the *Unbounded Property* [6].

[8] A deterministic distribution p_Λ is a distribution in which a singular event $j \in \Omega(\Lambda)$ occurs with certainty, i.e. $\mathsf{p}_\Lambda^j = 1$ and $\forall j' \neq j : \mathsf{p}_\Lambda^{j'} = 0$.

[9] Applying the Fourier-Motzkin procedure to completely solve GMP is discussed in more detail in Fritz and Chaves [2].

[10] This is a stronger variant of [11, Proposition 7].

the integer-valued coefficients that constrain Eq. (7). Finally, the homogeneity of the constraints in Eq. (8) follows from the assumption that each $\mathsf{p}_{*\mathcal{V}} \in \mathcal{P}_{*\mathcal{V}}$ a priori satisfies normalization constraints context-wise; i.e. $\forall V \in \mathcal{V} : \sum_{m\in\Omega(V)} \mathsf{p}_V^m = 1$[11]. $\square$

D. Logical Contextuality

Let $a \in \Omega(_*\mathcal{V})$ be *any* marginal event and $C = \{c_1, \dots, c_n\} \subseteq \Omega(_*\mathcal{V})$ be a subset of marginal events such that the following logical implication holds for *all* marginal models $\mathsf{p}_{*\mathcal{V}} \in \mathcal{P}_{*\mathcal{V}}$:

$$a \implies c_1 \vee \cdots \vee c_n = \bigvee_{c\in C} c \tag{9}$$

Which can be dictated: *whenever the event a occurs, at least one event in C occurs.* In accordance with the logical form of Eq. (9), a will be referred to as the **antecedent** and C as the **consequent set**. To clarify, a marginal model $\mathsf{p}_{*\mathcal{V}} \in \mathcal{P}_{*\mathcal{V}}$ satisfies Eq. (9) if there always at least one $c \in C$ that is *possible* ($\mathsf{p}_{*\mathcal{V}}^c > 0$) whenever a is possible. A marginal model violates Eq. (9) whenever *none* of events in c are possible while a remains possible. Marginal models that violate logical statements such as Eq. (9) are known as **Hardy Paradoxes** [10, 15, 16]. Motivated by a greater sense of robustness compared to possibilistic constraints, the concept of witnessing quantum contextuality on a logical level has be analyzed thoroughly for decades [11, 17].

All logical implications of the form of Eq. (9) can be derived by first selecting am antecedent marginal event a, then constructing a consequent set C such that Eq. (9) holds. This is accomplished by making use of Lemma 2.

Lemma 2. *Let $m \in \Omega(_*\mathcal{V})$ be a marginal event. Then for all non-contextual marginal models $\mathsf{p}_{*\mathcal{V}} \in \mathcal{N}$,*

$$m \iff \bigvee_{j\in\sigma(\mathcal{M}^m)} j$$

Essentially, if a joint distribution does exist ($\mathsf{p}_{*\mathcal{V}} \in \mathcal{N}$), then event m represents partial knowledge of the entire system of variables Λ; whenever m occurs, exactly one joint event j has occurred in reality and m must be a restriction of j. Applying Lemma 2 to the antecedent $a \in \Omega(_*\mathcal{V})$ and consequent set $C \subseteq \Omega(_*\mathcal{V})$,

$$\begin{aligned} a &\iff \bigvee_{j\in\sigma(\mathcal{M}^a)} j \\ \bigvee_{c\in C} c &\iff \bigvee_{c\in C} \bigvee_{j\in\sigma(\mathcal{M}^c)} j \end{aligned} \tag{10}$$

[11] Specifically, any inhomogeneous constraint $\gamma \cdot \mathsf{p}_{*\mathcal{V}} \geq \alpha$ can be *homogenized* by replacing α with $\sum_{m\in\Omega(V)} \alpha \mathsf{p}_V^m$.

Therefore, if a subset C of $\Omega(_*\mathcal{V})$ (preferably excluding a) can be found such that,

$$\sigma(\mathcal{M}^a) \subseteq \bigcup_{c\in C} \sigma(\mathcal{M}^c) \tag{11}$$

then Eq. (9) follows from Eqs. (10,11).

It is possible to show that for each logical constraint of the form Eq. (9), there exists a corresponding probabilistic constraint that is tighter. Corollary 2[12] generalizes Lemma 2.

Corollary 2. *Let $m \in \Omega(_*\mathcal{V})$ be a marginal event. Then for all non-contextual marginal models $\mathsf{p}_{*\mathcal{V}} \in \mathcal{N}$,*

$$\mathsf{p}_{*\mathcal{V}}^m = \sum_{j\in\sigma(\mathcal{M}^m)} \mathsf{p}_\Lambda^j$$

III. AN OBSERVATION

A. An Antecedent Hierarchy

B. The Antecedent Hypergraph

Given an antecedent multi-set γ where $\gamma \preceq 0$, we identify the **inhibiting set** of joint events $\mathcal{I}(\gamma) \subseteq \Omega(\Lambda)$ preventing $\gamma \cdot \mathcal{M}$ from being positive semi-definite:

$$\mathcal{I}(\gamma) \equiv \left\{ j \in \Omega(\Lambda) \mid (\gamma \cdot \mathcal{M})_j < 0 \right\}$$

The inhibiting set $\mathcal{I}(\gamma)$ of γ completely characterizes the **antecedent hypergraph** $\mathcal{H}(\gamma)$ whose edges $\mathcal{E}_j$ are indexed by the inhibiting events $j \in \mathcal{I}(\gamma)$. Each edge $\mathcal{E}_j \subseteq \Omega(_*\mathcal{V})$ corresponds to the set of the marginal events $m \in \Omega(_*\mathcal{V})$ which are restrictions of j. Specifically,

$$\begin{gathered} \mathcal{H}(\gamma) \equiv \{\mathcal{E}_j \mid j \in \mathcal{I}(\gamma)\} \\ \mathcal{E}_j \equiv \left\{ m \in \Omega(_*\mathcal{V}) \mid m = j|_{\mathcal{D}(m)}, \gamma_m = 0 \right\} \end{gathered}$$

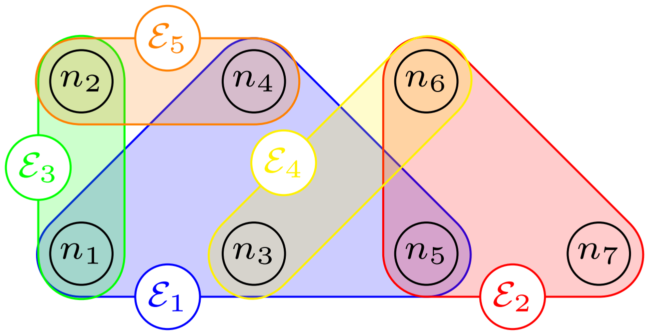
Todo (TC Fraser): Incorporate weights

$$\omega(\gamma) = \left\{ \omega_j = -(\gamma \cdot \mathcal{M})_j \mid j \in \mathcal{I}(\gamma) \right\}$$

Todo (TC Fraser): Define set of right-minimal inequalities

$$I(\gamma) = \left\{ (\gamma + \tilde{\gamma}) \cdot \mathsf{p}_{*\mathcal{V}} \geq 0 \mid \tilde{\gamma} \in \mathrm{Tr}_{\omega(\gamma)}[\mathcal{H}(\gamma)] \right\}$$

[12] Corollary 2 is simply the m-th row of Eq. (4).



$$
\begin{array}{c} \\ n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \\ n_6 \\ n_7 \end{array}
\begin{array}{c} \begin{array}{ccccc} \mathcal{E}_1 & \mathcal{E}_2 & \mathcal{E}_3 & \mathcal{E}_4 & \mathcal{E}_5 \end{array} \\
\begin{pmatrix}
\mathbf{1} & 0 & \mathbf{1} & 0 & 0 \\
0 & 0 & \mathbf{1} & 0 & \mathbf{1} \\
\mathbf{1} & 0 & 0 & \mathbf{1} & 0 \\
\mathbf{1} & 0 & 0 & 0 & \mathbf{1} \\
\mathbf{1} & \mathbf{1} & 0 & 0 & 0 \\
0 & \mathbf{1} & 0 & \mathbf{1} & 0 \\
0 & \mathbf{1} & 0 & 0 & 0
\end{pmatrix} \end{array}
$$

FIG. 1. Dual-representations of a hypergraph $\mathcal{H} = \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4, \mathcal{E}_5\}$.

C. Irreducibility

D. Marginal Symmetries

E. Curated Inequalities

F. Targeted Searches

G. Relaxations

IV. EDGE-WEIGHTED HYPERGRAPH TRANSVERSALS

A. Preliminaries

B. Hypergraph Transversals

C. Adding Weights

D. Unsolved

Consider γ_1, γ_2 and $\gamma_1 + \gamma_2$.

$$I(\gamma_a) = \left\{(\gamma_a + \tilde{\gamma}_a) \cdot \mathsf{p}_* \mathcal{V} \geq 0 \mid \tilde{\gamma}_a \in \mathrm{Tr}_{\omega(\gamma_a)}[\mathcal{H}(\gamma_a)]\right\} \quad a = 1, 2$$
$$I(\gamma_1 + \gamma_2) = \left\{(\gamma_1 + \gamma_2 + \tilde{\gamma}_{12}) \cdot \mathsf{p}_* \mathcal{V} \geq 0 \mid \tilde{\gamma}_{12} \in \mathrm{Tr}_{\omega(\gamma_1)+\omega(\gamma_2)}[\mathcal{H}(\gamma_1 +$$

V. CONCLUSIONS

ACKNOWLEDGMENTS

[1] S. Abramsky and A. Brandenburger, "The Sheaf-Theoretic Structure Of Non-Locality and Contextuality," New J. Phys **13**, 113036 (2011).

[2] T. Fritz and R. Chaves, "Entropic Inequalities and Marginal Problems," IEEE Trans. Info. Theor. **59**, 803 (2011).

[3] A. Schrijver, *Theory of Linear and Integer Programming* (Wiley, 1998).

[4] S. Lahaie, "How to take the Dual of a Linear Program," (2008).

[5] S. Fang and S. Puthenpura, *Linear Optimization and Extensions: Theory and Algorithms* (Englewood Cliffs, NJ: Prentice-Hall, 1993) p. 60.

[6] Bradley, Hax, and Magnanti, *Applied Mathematical Programming* (Addison-Wesley, 1977) Chap. 4, pp. 143–144.

[7] I. Pitowsky, "Correlation polytopes: Their geometry and complexity," Math. Prog. **50**, 395 (1991).

[8] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, "Bell nonlocality," Rev. Mod. Phys **86**, 419 (2013).

[9] G. B. Dantzig and B. C. Eaves, "Fourier-Motzkin elimination and its dual," J. Combin. Theor. A **14**, 288 (1973).

[10] E. Wolfe, R. W. Spekkens, and T. Fritz, "The Inflation Technique for Causal Inference with Latent Variables," (2016), arXiv:1609.00672.

[11] S. Abramsky and L. Hardy, "Logical Bell inequalities," Phys. Rev. A **85**, 062114 (2012).

[12] C. Jones, E. C. Kerrigan, and J. Maciejowski, *Equality set projection: A new algorithm for the projection of polytopes in halfspace representation*, Tech. Rep. (Cambridge University Engineering Dept, 2004).

[13] B. G. Jiri Matousek, *Understanding and Using Linear Programming (Universitext)* (Springer, 2013).

[14] G. M. Ziegler, *Lectures on Polytopes* (Springer New York, 1995).

[15] S. Mansfield and T. Fritz, "Hardy's Non-locality Paradox and Possibilistic Conditions for Non-locality," Found. Phys. **42**, 709 (2011).

[16] L. Mančinska and S. Wehner, "A unified view on Hardy's paradox and the Clauser–Horne–Shimony–Holt inequality," J. Phys. A **47**, 424027 (2014).

[17] D. M. Greenberger, "Bell's theorem without inequalities," Am. J. Phys. **58**, 1131 (1990).