FirstPage

The Inflation DAG Technique for Causal Inference with Hidden Variables

Elie Wolfe,[1,*] Robert W. Spekkens,[1,†] and Tobias Fritz[1,2,‡]

[1] *Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada, N2L 2Y5*
[2] *Max Planck Institute for Mathematics in the Sciences, Leipzig, Saxony, Germany, 04103*

(Dated: April 20, 2016)

The fundamental problem of causal inference is to infer from a given probability distribution over observed variables, what causal structures, possibly incorporating hidden variables, could have given rise to that distribution. Given some candidate causal structure, it is therefore valuable to derive infeasibility criteria, such that the hypothesis is not a feasible causal explanation whenever the observed distribution violates an infeasibility criterion. The problem of causal inference via infeasibility criteria comes up in many fields. Special infeasibility criteria are Bell inequalities (which distinguish non-classical from classical distributions) and Tsirelson inequalities (which distinguish quantum from post-quantum distributions), and Pearl's instrumental inequality. All of these are limited to very specific causal structures. Analogues of such inequalities for more-general causal structures, i.e., necessary criteria for either classical or quantum distributions to be realizable from the structure, are highly sought after.

We here introduce a technique for deriving such infeasibility criteria, applicable to any causal structure. It consists of first *inflating* the causal structure and then translating weak constraints on the inflated structure into stronger constraints on the original structure. Moreover, we show how our technique can be tuned to yield either classical criteria (i.e., that may have quantum violations), or post-classical criteria (i.e., that hold even in the context of general probability theories), depending on whether or not the inflation implicitly broadcasts the value of a hidden variable. Concretely, we derive polynomial inequalities for the so-called Triangle scenario, and we show how all Bell inequalities also follow from our method. Furthermore, given both a causal structure and a specific probability distribution, our technique can be used to efficiently witness their inconsistency, even absent explicit inequalities. The inflation technique is therefore both relevant and practical for general causal inference tasks with hidden variables.

* ewolfe@perimeterinstitute.ca
† rspekkens@perimeterinstitute.ca
‡ tobias.fritz@mis.mpg.de

I. INTRODUCTION

Given some hypothesis of causal structure, it is desirable to determine **infeasibility criteria**, i.e., observable constraints such that the their violation implies the invalidity of the hypothesis as an explanation for observational data. Causal infeasibility criteria are used in a wide variety of statistics application, from sussing out biological pathways to enabling machine learning pearl2009causality,spirtes2011causation,studeny2005probabilistic,koller2009probabilistic [? ? ? ?]. ADD SENTENCES ABOUT HOW OUR WORK CONTRIBUTES TO GENERAL CAUSAL INFERENCE TASKS. The foundational role of causal structure in quantum information theory has only recently been appreciated WoodSpekkens,fritz2012bell,pusey2014gdag,BeyondBellII [? ? ? ?].

In contexts other than quantum theory, the latent nodes in causal structures are generally taken to represent hidden variables. This is not fully general, however, so we apply the retronym[1] "classical", as in classical causal structure and classical causal inference. The classical distributions of a given causal structure are defined as those which arise from it while restricting the latent nodes to be arbitrary (classical) random variables. Quantum distributions, by contrast, are those which are realizable if the latent nodes in the causal structure are allowed to be quantum systems. We hereafter take all causal structures and probability distributions to be classical, except where explicitly stated otherwise.

From a physics perspective, therefore, tightly characterizing the set of observable probability distributions realizable from a causal structure is critical, in order to recognize and exploit the existence of distributions that can be realized quantumly but not classically. Few techniques are known for bounding this set of distributions which are simultaneously practical and applicable to general causal structures. Celebrated examples include the use of conditional independence relations (easy) pearl2009causality,spirtes2011causation,studeny2005probabi [? ? ? ?] and entropic inequalities (more advanced) fritz2013...,chaves2014novel,chaves2014informationinfe [? ? ?]. In the presence of hidden variables, these criteria only rarely provide a tight characterization, and frequently fail to witness the non-classicality of quantum distributions.

Distinguishing quantum from classical correlations has historically been achieved through the use of Bell inequalities bell1966lhvm,GisinFramework2012,scarani2012device,Brunner2013Bell,BancalDIApproach [? ? ? ? ?]. Bell inequalities, however, are limited to very special causal scenarios involving *only one* latent common cause variable, i.e. Bell scenarios. A Bell scenario is also very special in that its realizable distributions admit characterization by a finite set of linear inequalities (after conditioning on the setting variables), i.e. its realizable distributions comprise a convex polytope GisinFramework2012,FritzDuality [? ?]. Entirely new techniques, therefore, are required to derive quantum-sensitive infeasibility criteria for more general causal scenarios fritz2012bell,pusey2014gdag,BeyondBellII [? ? ?].

To this end, we here introduce a new technique, applicable to any causal structure, for deriving infeasibility criteria. This technique allows for, but is not limited to, the derivation of polynomial inequalities. These criteria are generally based on the *broadcasting* of the values of a hidden variable, i.e. the assumption that its value can be copied and broadcast at will. The no-broadcasting theorem from quantum theory shows that this is not valid in the non-classical case, and from our perspective this is the reason for the existence of quantum violations of Bell inequalities. Moreover, our technique can also be applied in order to derive criteria that must be satisfied for all distributions that can be generated with latent nodes that are states in quantum theory or any other general probabilistic theory, simply by not assuming the possibility of broadcasting.

II. NOTATION AND DEFINITIONS

finitions

We follow the convention that upper-case letters indicate random variables while lower-case letters indicate some particular value associated with the corresponding random variable. In this convention, for example, a student's score on some exam X might depend probabilistically on the amount of sleep S. The logical proposition, or **event**, $X{=}x, S{=}s$ should be understood as "the student scores x on the exam with a duration of sleep equal to s". Events may be written in lower-case-only shorthand, such as x, s instead of $X{=}x, S{=}s$.

Similarly, we indicate probability distributions using upper-case P, whereas lower-case p is used to indicate the probability of particular events. Thus $P(X, Y)$ is the multivariate probability distribution over the random variables $\{XY\}$, and $p(x, y)$ denotes the joint probability of the two events $X{=}x$ *and* $Y{=}y$. We often omit the comma and just write $P(XY)$ and $p(xy)$, respectively.

A causal structure, for the purpose of this article, is taken to be given by a directed acyclic graph (DAG): each node in the DAG corresponds to a random variable[2], while each edge represents a possible causal influence between variables. In our graphical depictions we follow the convention of representing latent nodes by circles, and observable nodes by triangles pusey2014gdag [?].

---

[1] Retronym (noun): a modification of an original term to distinguish it from a later development retronym [?].

[2] In the quantum context, however, only *observable* nodes correspond to classical random variables, such as the outcomes of measurements. Latent nodes, however, represent quantum systems. Functional dependence is replaced by the action of a quantum operation, and edges then dictate the ways in which quantum systems factor and/or compose. See Refs. pusey2014gdag,leifer2013conditionalstates [? ?]. Rob, add appendix? Classical/Quantum/GPT everything? I'm thinking a table would be nice.

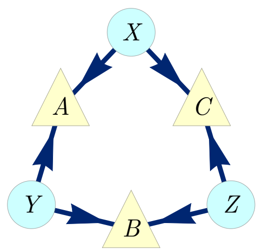

FIG. 1. The causal structure of the Triangle scenario.

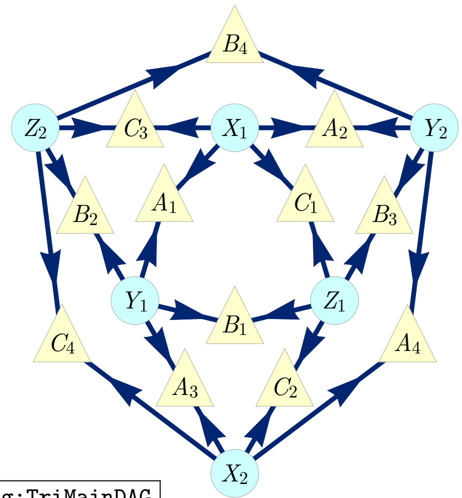

fig:TriMainDAG

FIG. 2. An inflation DAG of the Triangle scenario where each latent node has been duplicated, resulting in four copies of each observable node. fig:TriFullDouble

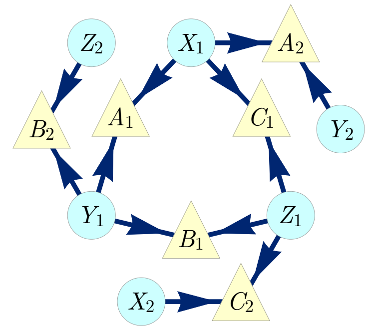

FIG. 3. Another inflation of the Triangle scenario consisting, also notably $\mathsf{AnSubDAG}_{(\text{Fig. 2})}[A_1A_2B_1B_2C_1C_2]$. fig:Tri22

Our terminology for the causal relations between the nodes in a DAG is the standard one. The parents of a node X in a given graph G are defined as those nodes which have directed edges originating at them and terminating at X, i.e. $\mathsf{Pa}_G(X) = \{\, Y \mid Y \to X \,\}$. If $\boldsymbol{U}$ is a set of nodes, then we put $\mathsf{Pa}_G(\boldsymbol{U}) := \bigcup_{X\in\boldsymbol{U}} \mathsf{Pa}_G(X)$. Similarly the children of a node X in a given graph G are defined as those nodes which have have directed edges originating at X and terminating at them, i.e. $\mathsf{Ch}_G(X) = \{\, Y \mid X \to Y \,\}$. The **ancestors** $\mathsf{An}(\boldsymbol{U})$ are defined as those nodes which have a directed *path* to some node in $\boldsymbol{U}$, including the nodes in themselves. Equivalently, $\mathsf{An}(\boldsymbol{U}) := \bigcup_{n\in\mathbb{N}} \mathsf{Pa}^n(\boldsymbol{U})$, where $\mathsf{Pa}^n(\boldsymbol{U})$ is inductively defined via $\mathsf{Pa}^0(\boldsymbol{U}) := \boldsymbol{U}$ and $\mathsf{Pa}^{n+1}(\boldsymbol{U}) := \mathsf{Pa}(\mathsf{Pa}^n(\boldsymbol{U}))$.

We refer to a pair of nodes which not share any common ancestor as being **ancestrally independent**, for which we invent the notation $X \mathrel{\lrcorner\emptyset\llcorner} Y$. Generalizing to sets, $\boldsymbol{U} \mathrel{\lrcorner\emptyset\llcorner} \boldsymbol{V}$ indicates that no node in $\boldsymbol{U}$ shares a common ancestor with any node in $\boldsymbol{V}$, i.e. $\mathsf{An}(\boldsymbol{U}) \cap \mathsf{An}(\boldsymbol{V}) = \emptyset$. It is possible for more than two sets to be ancestrally independent: the notation $\boldsymbol{U} \mathrel{\lrcorner\emptyset\llcorner} \boldsymbol{V} \mathrel{\lrcorner\emptyset\llcorner} \boldsymbol{W}$ should be understood as indicating that the ancestors of $\boldsymbol{U}$,$\boldsymbol{V}$, and $\boldsymbol{W}$ comprise three distinct non-overlapping sets, i.e. $\boldsymbol{U} \mathrel{\lrcorner\emptyset\llcorner} \boldsymbol{V}$ and $\boldsymbol{V} \mathrel{\lrcorner\emptyset\llcorner} \boldsymbol{W}$ and $\boldsymbol{U} \mathrel{\lrcorner\emptyset\llcorner} \boldsymbol{W}$.

Ancestral independence is equivalent to d-separation by the empty set [? ? ? ?] (pearl2009causality,spirtes2011causation,studeny2005probab…). Therefore, for distributions that are Markov with respect to the DAG [? ? ? ?] (pearl2009causality,spirtes2011causation,studeny2005probabilistic,koller2009probabilistic), ancestral independence of nodes implies marginal independence of the random variables. For us, all distributions over the variables represented by the nodes of a DAG will be assumed Markov. Intuitively, we presume that no statistical correlation is possible without causal explanation. Thus if a DAG possesses the feature $\boldsymbol{U} \mathrel{\lrcorner\emptyset\llcorner} \boldsymbol{V} \mathrel{\lrcorner\emptyset\llcorner} \boldsymbol{W}$, we demand factorization of the marginal distributions such that $P(\boldsymbol{U}\boldsymbol{V}\boldsymbol{W}) = P(\boldsymbol{U})P(\boldsymbol{V})P(\boldsymbol{W})$.

$\mathsf{SubDAG}_G[\boldsymbol{V}]$ refers to the induced subgraph of G on a set of nodes $\boldsymbol{V}$. It consists of the nodes $\boldsymbol{V}$ and the edges between pairs of nodes in $\boldsymbol{V}$ per the original graph. Of special importance to us is the **ancestral subgraph**, $\mathsf{AnSubDAG}[\boldsymbol{V}] := \mathsf{SubDAG}[\mathsf{An}(\boldsymbol{V})]$, which is the minimal subgraph containing the full ancestry of $\boldsymbol{V}$.

We now introduce the notion of an **inflation DAG**. The nodes of an inflation DAG are copies $A_1, \ldots, A_k$ of the nodes of the original DAG, where the subscript i is just a dummy index of sorts, and such that the following defining property holds: **upon removing the subscripts indexing the copies, the ancestral subgraph in G' of a node A_i looks like the ancestral subgraph of the corresponding node A in G.** The idea is that at the level of distributions, this property guarantees that all copies of a node in the inflation DAG have the same distribution as the corresponding node in the original DAG *if* the functional dependencies among the variables are taken to be the same.

When two objects (e.g. nodes, sets of nodes, DAGs, etc…) are the same up to dummy indices, then we use $\sim$ to indicate this, so $A_i \sim A_j \sim A$. The formal definition of an inflation DAG is therefore as follows:

$$G' \in \mathsf{Inflations}[G] \quad \text{iff} \quad \forall_{A_i\in G'}\ \mathsf{AnSubDAG}_{G'}[A_i] \sim \mathsf{AnSubDAG}_G[A]. \tag{1}$$

{eq:defin…

As an example, consider the Triangle scenario [? (Fig. E#8), ? (Fig. 18b), ? (Fig. 3), ? (Fig. 6a), ? (Fig. 1a), ? (Fig. 8), ? (Fig. 1b), ? (Fig. 4b)] (pusey2014gdag, WoodSpekkens, fritz2012bell, chaves2014nove…, Chaves2015inf…, QuantumCorrela…, steudel2010ancestors, chaves2014informationinference). The associated DAG, the shape of which explains the name, is depicted here in

Fig. 1. Possible inflations of the Triangle scenario are depicted in Figs. 2 and 3.

Any causal model for the original DAG specifies a corresponding causal model for the inflated DAG, by virtue of

$$\forall_{A_i \in G'}\ P(A_i|\mathsf{Pa}(A_i)) = P(A|\mathsf{Pa}(A)), \tag{2}$$ {eq:funcd

where one identifies the parents of A_i in G' with the parents of A in G. If a causal model on G' is of this form, we call it an **inflation model**. In particular, all copies of exogenous (non-root) nodes in an inflation model share the same functional dependence on their parents, and all copies of endogenous (root) nodes in the inflated DAG are identically *and independently* distributed. For us, the most relevant feature of an inflation model is that all the copies of a single random variable can have the same probability distribution as the corresponding variable in the original DAG,

$$\forall_{A_i \in G'}\ P(A_i) = P(A), \tag{3}$$ {eq:singl

because of Eq. (2) and $\mathsf{AnSubDAG}_{G'}[A_i] \sim \mathsf{AnSubDAG}_G[A]$. Note that the operation of equipping a modified DAG with the same functional dependencies as the original one also appears in the *do calculus* of [pearl2009causality ?].

FiXme Note: non-Shannon

To be perfectly clear however, a_1 and a_2 *do not* refer to two distinct possible outcomes of one random variable; rather $a_1 \wedge a_2$ represents the event in which A_1 and A_2 both have the same outcome a. Moreover, generally $p(a_1a_2) \neq p(a)$, because although $P(A_1) = P(A_2)$, nevertheless A_1 and A_2 may not be perfectly correlated. In the same vein, generally $p(a_1a_2) \neq p(a_1)p(a_2)$, because identically distributed does not mean independently distributed. Indeed, in Fig. 3, for example, A_1 and A_2 share the common ancestor X_1, and hence they are not independent. On the other hand, sometimes two copies of a random variable not share any common ancestor, such as A_1 and A_4 in Fig. 2. Fig. 2 implies, therefore, that $p(a_1a_4) = p(a_1)p(a_4)$.

FiXme Note:

Any subset of nodes of the inflation DAG which contains multiple copies of some node is considered a **redundant** set. Formally, $\boldsymbol{U}$ is redundant if and only if there are A and indices i, j such that $\{A_i, A_j\} \subseteq \boldsymbol{U}$. Otherwise, $\boldsymbol{U}$ is said to be irredundant. Alternatively, $\boldsymbol{U} \subseteq \mathsf{Nodes}[G']$ is irredundant if and only if there is $\boldsymbol{V} \subseteq \mathsf{Nodes}[G]$ such that $\boldsymbol{U} \sim \boldsymbol{V}$. When discussing irredundant sets we hereafter denote the corresponding original-DAG node-set by underscript $\sim$, i.e. $\boldsymbol{U} \sim \tilde{\boldsymbol{U}}$ where $\tilde{\boldsymbol{U}} \subseteq \mathsf{Nodes}[G]$.

Critically, the idea that coinciding causal histories implies coinciding distributions can be generalized from individual nodes to irredundant sets. While generally $P(\boldsymbol{U}) \neq P(\tilde{\boldsymbol{U}})$, $\boldsymbol{U}$ and $\tilde{\boldsymbol{U}}$ *must* have coinciding distribution whenever their respective ancestral subgraphs are equivalent up to dummy indices, as a consequence of Eq. (2). The natural generalization of Eq. (3), therefore, is,

$$P(\boldsymbol{U}) = P(\tilde{\boldsymbol{U}}) \quad \text{whenever} \quad \mathsf{AnSubDAG}_{G'}[\boldsymbol{U}] \sim \mathsf{AnSubDAG}_G[\tilde{\boldsymbol{U}}]. \tag{4}$$ {eq:coinc

For example, any individual node $\boldsymbol{U} = \{X\}$ has this property. General sets of nodes which posses this special property form the inferential link between the inflation DAG and the original DAG. An irredundant set of nodes $\boldsymbol{U}$ in the inflation DAG will be called **injectable** if $\boldsymbol{U}$ is a positive instance of Eq. (4), i.e.

$$\boldsymbol{U} \in \mathsf{Injectables}[G'] \quad \text{iff} \quad \mathsf{AnSubDAG}_{G'}[\boldsymbol{U}] \sim \mathsf{AnSubDAG}_G[\tilde{\boldsymbol{U}}]. \tag{5}$$ {eq:defin

Any redundant set is obviously not injectable; it should be clear from Eq. (5) that $\boldsymbol{U}$ is also not injectable whenever $\mathsf{An}(\boldsymbol{U})$ is redundant. Moreover,

$$\boldsymbol{U} \text{ is injectable} \quad \text{iff} \quad \mathsf{An}(\boldsymbol{U}) \text{ is irredundant}, \tag{6}$$ {eq:defin

because the edge structure is automatically preserved by the definition of an inflation DAG. If $\boldsymbol{U}$ is injectable, then any subset of $\boldsymbol{U}$ is also injectable. If $\boldsymbol{U}$ is not injectable, then any superset of $\boldsymbol{U}$ is not injectable. In Fig. 3, for example, $\{A_1B_1C_1\}$ and $\{A_2C_1\}$ are both injectable, but $\{A_1A_2C_1\}$ and $\{A_2B_1C_1\}$ are both not injectable.

It is useful to consider a looser notion than injectability. A set of nodes in the inflation DAG $\boldsymbol{U}$ will be called **pre-injectable** whenever it is a union of injectable sets with disjoint ancestries. Equivalently, a set of nodes is pre-injectable if and only if its (weakly) connected components are injectable. In particular, every injectable set is also pre-injectable.

For us, the crucial property of a pre-injectable set $\boldsymbol{U}$ is that in an inflation model, $P(\boldsymbol{U})$ is fully determined by the original causal model via Eq. (2). More concretely, if $\boldsymbol{U}_1, \dots, \boldsymbol{U}_n$ are the weakly connected components of $\boldsymbol{U}$, then in an inflation model we must have

$$P(\boldsymbol{U}) = P(\boldsymbol{U}_1) \cdots P(\boldsymbol{U}_n) = P(\tilde{\boldsymbol{U}}_1) \cdots P(\tilde{\boldsymbol{U}}_n). \tag{7}$$ {eq:prein

For this reason, pre-injectable sets will play a role in deriving polynomial inequalities via polytope projection techniques.

It is worth noting that duplicating an outgoing edge in a causal structure means **broadcasting** the value of the random variable. For example in Fig. 4, the information about X which was "sent" to A is effectively broadcast to both A_1 and A_2 in the inflation. This is quite intentional. Quantum theory is governed by a no-broadcasting theorem NoCloningQuantum1996,NoCloningGeneral2006 [? ?]; by electing to embed broadcasting into an inflation DAG we can specifically construct a foil to quantum causal structures. Infeasibility constraints derived from **non-broadcasting inflations** on the other hand, such as Fig. 5, are valid even when relaxing the interpretation of latent nodes to allow for quantum or general probabilistic resources. This contrast is elaborated at length in Sec. VII.

So a non-broadcasting inflation DAG is one in which the set of children of every latent node in the inflation DAG G' is irredundant, i.e.

$$G' \in \mathsf{NonBroadcastingInflations}[G] \quad \text{iff} \quad \forall_{\text{latent } A_i \in G'} \ \mathsf{Ch}_{G'}(A_i) \text{ is an irredundant set.} \tag{8}$$

{eq:nonbr

We also find it useful to define the notion of a non-broadcasting subset of nodes within some larger broadcasting inflation DAG. A set of nodes $\boldsymbol{U}$ is a **non-broadcasting set** iff $\mathsf{AnSubDAG}_{G'}[\boldsymbol{U}]$ is a non-broadcasting inflation DAG. Any inference about the original DAG which can be made by referencing exclusively to non-broadcasting sets hold in both the classical and quantum paradigms. Broadcasting inflation DAGs are therefore especially useful for deriving criteria which distinguish quantum and classical probability distributions, but we anticipate them to be valuable for broader causal inference tasks as well.

In classical causal structures the latent nodes correspond to classical hidden variables. In quantum causal structures, however, the latent nodes are taken to be quantum systems. Some quantum causal structures are famously capable of realizing distributions that would not be possible classically, although the set of quantum distribution is superficially quite similar to the classical subset pusey2014gdag,fritz2012bell [? ?]. For example, classical and quantums distributions alike respect all conditional independence relations implied by the common underlying causal structure pusey2014gdag [?]. Recent work has found that quantum causal structure implies many of the entropic inequalities implied by their classical counterparts pusey2014gdag,Chave [? ? ?]. To-date, no quantum distribution has been witnessed to violate a Shannon-type entropic inequality on observable variables derived from the Markov conditions on all nodes chaves2012entropic,fritz2012bell [? ?]. Fine-graining the scenario by conditioning on discrete settings leads to a different kind of entropic inequality, and these have proven somewhat quantum-sensitive braunstein1988 [? ? ?]. Such braunstein1988entropic [?] type inequalities are still limited, however, in that they rely on root observable nodes[3], and they still fail to detect certain extremal non-classical distributions chaves2014novel,fritz2012bell [? ?].

The insufficiency of entropic inequalities is a pressing concern in quantum information theory since they often fail to detect the infeasibility of quantum distributions, for instance in the Triangle scenario fritz2012bell [? , Prob. 2.17]. The superficial similarity between quantum and classical distributions demands especially sensitive causal infeasibility criteria in order to distinguish them. We hope that polynomial inequalities derived from broadcasting inflation DAGs will be suitable tools for this purpose.

III. THE INFLATION DAG TECHNIQUE

algorithm

Our main contribution here is in recognizing that inferences about the original DAG can be made by proxy, so to speak, using the inflation DAG. If a distribution can arise from the original DAG, then the corresponding specification of the injectable sets must also be realizable from the inflation DAG. **Conversely, one may certify that a distribution is inconsistent with the original DAG by proxy, namely by proving that there is no inflation model on the inflation DAG that would reproduced the known probabilities on observed nodes.**

We illustrate this principle by first showing how inflation DAG considerations can be used to prove the infeasibility of particular distributions. Polynomial inequality constraints are derived later on in Sec. IV.

Example 1: **Perfect correlation cannot arise from the Triangle scenario.**

Let us ask, is it possible for the three original-scenario observable variables $\{A, B, C\}$ to be random but perfectly correlated? We call this the GHZ-type distribution, after an eponymous tripartite entangled quantum state greenberger1989going,30 [? ?]. So $P_{\text{GHZ}}(ABC)$ is given by

$$p_{\text{GHZ}}(abc) = \frac{[000] + [111]}{2} = \begin{cases} \frac{1}{2} & \text{if } \ a = b = c, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

{eq:ghzdi

[3] Rafael Chaves and E.W. are exploring the potential of entropic analysis based on fine-graining causal structures over non-root observable nodes. This generalizes the method of entropic inequalities, and might be capable of providing much stronger entropic infeasibility criteria.

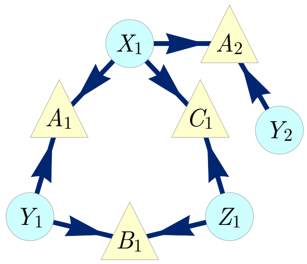

FIG. 4. A simple inflation of the Triangle scenario, also notably $\mathsf{AnSubDAG}_{(Fig.\ 3)}[A_1A_2B_1C_1]$. fig:simpleinflation

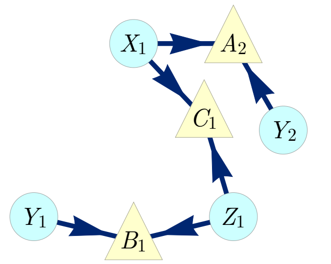

FIG. 5. An even simpler inflation of the Triangle scenario, also notably $\mathsf{AnSubDAG}_{(Fig.\ 4)}[A_2B_1C_1]$. fig:simplestinflation

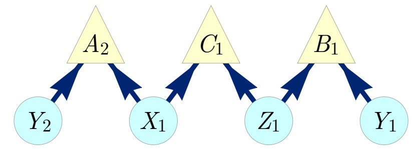

FIG. 6. Another representation of Fig. 5. Despite not containing the original scenario, this is a valid inflation per Eq. (1). fig:TriDa

We assume uniform binary variables for the sake of concreteness, but the argument is general. Let's temporarily (falsely) assume P_{GHZ} to be feasible for the triangle scenario. Since $\{A_2C_1\}$ is an injectable set we have $P(A_2C_1) = P(AC)$, and therefore P_{GHZ} implies that A_2 and C_1 are perfectly correlated. Similarly, since $\{B_1C_1\}$ is an injectable set we have $P(B_1C_1) = P(AC)$, and therefore B_1 and C_1 must be perfectly correlated, and by extension B_1 and A_2 are perfectly correlated as well. On the other hand, A_2 and B_1 must be statistically independent, as they do not share any common ancestor. The only way for two variables to be *both* perfectly correlated and independent is by being deterministic. This is not the case in P_{GHZ}, and thus we have certified that P_{GHZ} is not realizable from the Triangle causal structure, recovering the seminal result of [?] (steudel2010ancestors).

To be clear, one can employ *any* kind of causal infeasibility criteria on the inflation DAGs to constrain the distributions on the original DAG. For example, here's a quantitative version of the same proof against perfect correlation in the Triangle scenario, this time using entropic arguments: Trivially $H(A_2|B_1) \leq H(A_2|C_1) + H(C_1|B_1)$, because the amount of information required to learn A_2 from B_1 is surely less than the amount of information required to learn A_2 from C_1 and also to learn C_1 from B_1. On the other hand, the marginal independence of A_2 and B_1 implies $H(A_2|B_1) = H(A_2)$. P_{GHZ} is ruled out, because it would yield $H(A_2|C_1) = H(C_1|B_1) = 0$ while $H(A_2) = 1$. The entropic conclusion is $H(A_2) \leq H(A_2|C_1) + H(C_1|B_1)$, or equivalently, $I(A_2 : C_1) + I(B_1 : C_1) \leq H(C_1)$. The injectability of sets per Fig. 6 in turn yields

$$I(A : C) + I(B : C) \leq H(C), \tag{10}$$ {eq:monog

an inequality known as monogamy of correlations [?] (fritz2012bell). It forbids perfect correlation of all three variables. Because Eq. (10) is derivable from a non-broadcasting inflation DAG, it follows that monogamy of correlations holds even in the context of generalized probabilistic theories. This recovers a result of [? , Cor. 24] (pusey2014gdag). Indeed, Fig. 5 in Ref. [?] (pusey2014gdag) is essentially equivalent to Fig. 6 here.

Example 2: **The W-type distribution cannot arise from the Triangle scenario**

Another distribution inconsistent with the Triangle scenario is the W-type distribution, named after a quantum state appearing in Ref. [?] (3Qubits2Ways). To our knowledge, the infeasibility of the Triangle causal structure to explain the W-type distribution is a novel result here. The distribution $P_{\text{W}}(ABC)$ is given by

$$p_{\text{W}}(abc) = \frac{[100] + [010] + [001]}{3} = \begin{cases} \frac{1}{3} & \text{if } \ a + b + c = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$ {eq:wdist

To prove that P_{W} clashes with the Triangle causal structure, consider the inflation DAG of Fig. 3. The set $\{A_2C_1\}$ is injectable, so P_{W} implies that $C_1{=}0$ whenever $A_2{=}1$. Similarly, $B_1{=}0$ whenever $C_2{=}1$, and $A_1{=}0$ whenever $B_2{=}1$. The independence of A_2,B_2, and C_2 means that $p(A_2{=}B_2{=}C_2{=}1) = {}^1\!/_8$ according to P_{W}. But that would imply $p(A_1{=}B_1{=}C_1{=}0) \geq {}^1\!/_8$, which contradicts P_{W}, hence P_{W} cannot arise from the Triangle scenario.

The W-type distribution is difficult to witness as unrealizable using conventional causal inference techniques.

1. There are no conditional independence relations between the observable nodes of the Triangle scenario.
2. Shannon-type entropic inequalities cannot detect this distribution as not allowed by the Triangle scenario [? ? ?] (fritz2013marginal,c).
3. Moreover, *no* entropic inequality can witness the W-type distribution as unrealizable. [?] (weilenmann2016entropic) have constructed an inner approximation to the entropic cone of the Triangle causal structure, and the W-distribution lies inside this. In other words, a distribution with the same entropic profile as the W-type distribution *can* arise from the

Triangle scenario.

4. The newly-developed method of covariance matrix causal inference due to kela2016covariance [?], which gives tighter constraints than entropic inequalities for the Triangle scenario, also cannot detect inconsistency with the W-type distribution.

But the inflation DAG approach can, and does so very easily.

Example 3: **The PR-box cannot arise from the Bell scenario**

Consider the causal structure associated to the Bell [? ? ? ?] scenario [? (Fig. E#2), [? (Fig. 19), [? (Fig. 1), [? (Fig. 1), [? (Fig. 2b), [? (Fig. 2)], depicted here in Fig. 7. The observable variables are $\{A, B, X, Y\}$, and Λ is the latent common cause of A and B. An inflation of the Bell scenario is show in Fig. 8.

The PR-box distribution is famously inconsistent with the Bell scenario PROriginal,PRUnit [? ?]. Here we prove its unrealizability using inflation DAG logic. The distribution is given by $P_{\text{PR}}(ABXY) = P_{\text{PR}}(AB|XY)P(X)P(Y)$, where $P(X)$ and $P(Y)$ are arbitrary full-support distributions[4] over the binary variables X and Y, and

$$p_{\text{PR}}(ab|xy) = \begin{cases} \frac{1}{2} & \text{if } \mathsf{mod}_2[a+b] = x \cdot y, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

{eq:PRbox

We begin by recognizing that $\{A_1B_1X_1Y_1\}$, $\{A_1B_1X_2Y_1\}$, $\{A_1B_2X_1Y_2\}$, and $\{A_2B_2X_2Y_2\}$ are all injectable sets, and that X_1, X_2, Y_1, and Y_2 are all independent variables. No matter how biased the distributions $P(X)$ and $P(Y)$ are, however, surely the event $\{X_1, X_2, Y_1, Y_2\} = \{0, 1, 0, 1\}$ occurs sometimes. Whenever it does, P_{PR} specifies perfect correlation between A_1 and B_1, perfect correlation between A_1 and B_2, perfect correlation between A_2 and B_1, and perfect *anti*correlation between A_2 and B_2. Those four requirements are not mutually compatible: since perfect correlation is transitive, the first three properties entail perfect correlation between A_2 and B_2. Hence the PR-box distribution is disallowed by the Bell scenario.

FiXme Note: this?

The following section discusses how to procedurally derive polynomial infeasibility criteria pertaining to original DAG from properties of an inflation DAG.

IV. A PROCEDURE FOR DERIVING POLYNOMIAL INEQUALITIES

sec:ineqs

Our final goal in deriving polynomial inequalities is to derive constraints on the distribution of the observable variables in the original DAG. We accomplish this by proxy, namely by deriving constraints on the distributions over the variables in the pre-injectable sets of the inflation DAG. Just as distributions over the pre-injectable sets translate into distributions pertaining to the original DAG variables per Eq. (4), so do *constraints* on the distributions of pre-injectable sets translate into constraints on the original DAG variables.

We are aware of a variety of methods for identifying interesting constraints on the distributions over the variables in the pre-injectable sets. But here, we focus on the very simplest and weakest such constraint: the nonnegativity of probabilities together with the factorization Eq. (7). In fact, this is all that we have used in the previous section as well.

Our method is based on deriving inequalities for the so-called **marginal problem**. This relies on linear quantifier elimination, which is computationally feasible in at least some cases. A nonlinear method for deriving even tighter

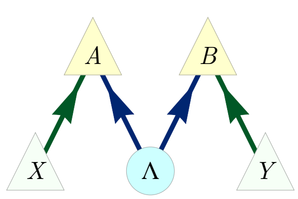

FIG. 7. The causal structure of the a bipartite Bell scenario. The local outcomes of Alice's and Bob's experimental probing is assumed to be a function of some latent common cause, in addition to their independent local experimental settings.

fig:NewBellDAG1

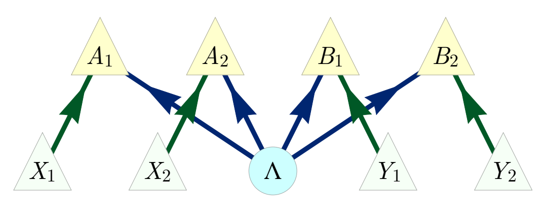

FIG. 8. An inflation DAG of the bipartite Bell scenario, where both local settings variables have been duplicated.

fig:BellD

[4] In the literature on the Bell scenario, these variables are known as "settings". Generally, We may think of endogenous observable variables as settings, coloring them light green in the DAG figures. Settings variables are natural candidates for conditioning on.

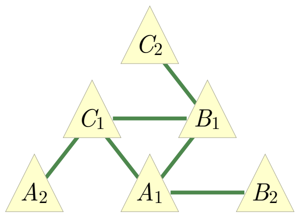

FIG. 9. The auxiliary injection graph corresponding to the inflation DAG in Fig. 3, wherein a pair of nodes are adjacent iff they are pairwise injectable.

fig:injection222

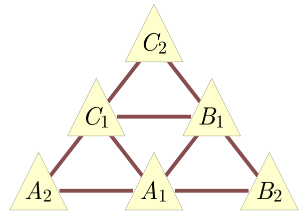

FIG. 10. An auxiliary ancestral dependance graph corresponding to the inflation DAG in Fig. 3, wherein a pair of nodes are adjacent iff they share a common ancestor.

fig:dependances222

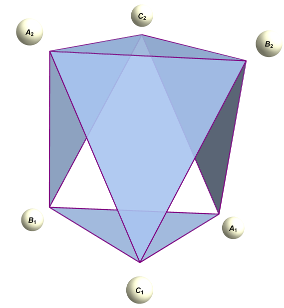

FIG. 11. The simplicial complex... Tobias - you please caption this?The 5 faces in this figure correspond to the pre-injectable sets.

fig:simpl

constraints is discussed in appendix Appendix A. The simpler method discussed here consists of two steps: 1) Identify the pre-injectable sets, and 2) Solve the marginal problem with respect to the pre-injectable sets.

njectable

Step 1: Identifying the Pre-Injectable Sets

To identify the pre-injectable sets, we first identify the injectable sets. To this end, it is useful to construct an auxiliary graph from the inflation DAG. Let the nodes of these auxiliary graphs be the observable nodes in the inflation DAG. The **injection graph**, then, is the undirected graph in which a pair of nodes A_i and B_j are adjacent if $\mathsf{An}(A_i B_j)$ is irredundant. The injectable sets are then precisely the cliques[5] in this graph, per Eq. (6).

Determining the pre-injectable sets from there can be done via constructing another graph that we call the **independence graph**. Its nodes are the injectable sets, and we connect two of these by an edge if their ancestral subgraphs are disjoint. Then by definition, the pre-injectable sets can be obtained as the cliques in this graph. Taking the union of all the injectable sets in such a clique results in a pre-injectable set. Since it is sufficient to only consider the maximal pre-injectable sets, one can eliminate all those pre-injectable sets that are contained in other ones, as a final step.

Let us also define the **ancestral dependence graph**, in which two nodes are adjacent if they share a common ancestor, and its complement the **ancestral independence graph**, in the ancestrally independent nodes are adjacent. To ascertain the factorization of a node set $\boldsymbol{U}$ into ancestrally-independent partitions one considers the subgraph on $\boldsymbol{U}$ of the ancestral dependence graph: the ancestrally-independent partitions are identically the distinct connected components of that subgraph. By examining the injection graph and the ancestral dependence graph, therefore, one is able to quickly determine all injectable sets and all ancestral independence relations.

Applying these prescriptions to the inflation DAG in Fig. 3 identifies the following ancestral independences maximal injectable and maximal pre-injectable sets as follows:

$$\underbrace{\begin{array}{c} A_2 \Lsh\emptyset\Rsh B_1 \\ A_2 \Lsh\emptyset\Rsh C_2 \\ B_2 \Lsh\emptyset\Rsh A_2 \\ B_2 \Lsh\emptyset\Rsh C_1 \\ C_2 \Lsh\emptyset\Rsh A_1 \\ C_2 \Lsh\emptyset\Rsh B_2 \end{array}}_{\substack{\text{pairwise}\\ \text{ancestral}\\ \text{independences}}} \qquad \underbrace{\begin{array}{c} \{A_2\} \Lsh\emptyset\Rsh \{B_1 B_2 C_2\} \\ \{B_2\} \Lsh\emptyset\Rsh \{A_2 C_1 C_2\} \\ \{C_2\} \Lsh\emptyset\Rsh \{A_1 A_2 B_2\} \\ \{A_2\} \Lsh\emptyset\Rsh \{B_2\} \Lsh\emptyset\Rsh \{C_2\} \end{array}}_{\substack{\text{maximal}\\ \text{ancestral}\\ \text{independences}}} \qquad \underbrace{\begin{array}{c} \{A_1 B_1\} \\ \{B_1 C_1\} \\ \{A_1 C_1\} \\ \{A_2 C_1\} \\ \{B_2 A_1\} \\ \{C_2 B_1\} \end{array}}_{\substack{\text{pairwise}\\ \text{injectable}\\ \text{sets}}} \qquad \underbrace{\begin{array}{c} \{A_1 B_2\} \\ \{B_1 C_2\} \\ \{A_2 C_1\} \\ \{A_1 B_1 C_1\} \end{array}}_{\substack{\text{maximal}\\ \text{injectable}\\ \text{sets}}} \qquad \underbrace{\begin{array}{c} \{A_1 B_1 C_1\} \\ \{A_1 B_2 C_2\} \\ \{A_2 B_1 C_2\} \\ \{A_2 B_2 C_1\} \\ \{A_2 B_2 C_2\} \end{array}}_{\substack{\text{maximal}\\ \text{pre-injectable}\\ \text{sets}}} \tag{13}$$

{eq:basic

[5] A clique is a set of nodes such that every single node is connected to every other.

such that the distributions on the pre-injectable sets relate to the original DAG distribution via

$$
\begin{aligned}
&P(A_1B_1C_1) = P(ABC) \\
&P(A_1B_2C_2) = P(C)P(AB) \\
&P(A_2B_1C_2) = P(A)P(BC) \\
&P(A_2B_2C_1) = P(B)P(AC) \\
&P(A_2B_2C_2) = P(A)P(B)P(C)
\end{aligned} \tag{14}
$$

{eq:prein

Having identified the pre-injectable sets (and how to use them), we next turn to deriving constraints on the distributions over the pre-injectable sets.

lsproblem

Step 2: Constraining the Distribution over Pre-Injectable Sets via the Marginal Problem

The most trivial constraint possible is the *existence of a joint probability distribution* over all the observable variables in the inflation DAG. Each of the five three-variable distributions in Eq. (14) is a different marginal distribution of the six-variable joint distribution $P(A_1A_2B_1B_2C_1C_2)$. Solving the marginal problem means finding inequalities on the marginal distributions such that the inequalities will satisfied only if there exists some joint distribution from which the distributions on the marginal sets can be recovered through marginalization. The marginal problem comes up in a variety of applications, and has been studied extensively; see [? fritz2013marginal] for further references.Tobias, say better? And add citations please! T: Ok, I'll get back to this on my second iteration

Comment about how solving marginals problem still gives polynomial inequalities.

From now on, one can apply *any* method for dealing with the marginal problem. For example if one is given a particular distribution P over the original scenario variables, then one can use Eq. (7) to compute the marginal distributions over the pre-injectable sets in an inflation DAG. Efficient linear programming can be used to to check whether there exists a joint distribution reproducing these marginals [? ? Korovin2012ImplementingCRA,Bobot2012SimplexSAT]. If such a joint distribution does not exist, then the given original-scenario distribution is witnessed as inconsistent with the original causal structure. If a joint distribution *does* exist then the problem remains open, and one can try again using a different inflation.

In order to derive inequalities that hold for *all* distributions that are feasible on the original DAG, say for a given number of outcomes per variable, we can formally solve the marginal problem via linear quantifier elimination, as we review in the following. This consists of eliminating unknowns from a set of linear inequalities and equalities.

The linear inequalities correspond to the nonnegativity of the joint distribution. Formally, the probability of any possible assignment of outcomes to to the observable variables is constrained to be nonnegative. For the six observable variables in Fig. 3, for example, the linear inequalities are given by $0 \leq P(A_1A_2B_1B_2C_1C_2)$. Note than a single inequality (or equality) pertaining to a probability *distribution* is actually shorthand for a whole set of inequalities (or equalities) pertaining to the probabilities of individual events. Taking the observable variables to be binary, for example, would mean that $0 \leq P(A_1A_2B_1B_2C_1C_2)$ would be shorthand for 64 distinct nonnegativity inequalities, namely

$$
\begin{aligned}
&0 \leq p(a_1a_2b_1b_2c_1c_2), \quad 0 \leq p(\overline{a}_1a_2b_1b_2c_1c_2), \quad 0 \leq p(a_1\overline{a}_2b_1b_2c_1c_2), \quad \ldots \quad , \\
&0 \leq p(\overline{a}_1\overline{a}_2b_1b_2c_1c_2), \quad 0 \leq p(\overline{a}_1a_2\overline{b}_1b_2c_1c_2), \quad \ldots
\end{aligned} \tag{15}
$$

[eqs]{eq:

etc. For a marginals problem based on the joint existence of n observable variables, each ranging over r possible outcomes, one would initialize the problem with r^n linear nonnegativity inequalities. Unlike probabilities pertaining to injectable sets, these joint probabilities are *not* fully specified by probabilities which are observed in the original scenario. We coin the term **gedankenprobability** to denote a probability pertaining to a not-pre-injectable set of inflation-DAG variables. The gedankenprobabilities evoke thought experiments, because knowing the original-scenario causal model determines the gedankenprobabilities.

The linear equations in the marginal problem are those equations which express each of the marginal probabilities as a sum over various different gedankenprobabilities, namely marginalization. Again using distribution shorthand, the five marginal distributions in Eq. (14) would correspond to the equations

$$
\begin{aligned}
P(A_1B_1C_1) &= \sum\nolimits_{A_2B_2C_2} P(A_1A_2B_1B_2C_1C_2),\\
P(A_1B_2C_2) &= \sum\nolimits_{A_2B_1C_1} P(A_1A_2B_1B_2C_1C_2),\\
P(A_2B_1C_2) &= \sum\nolimits_{A_1B_2C_1} P(A_1A_2B_1B_2C_1C_2),\\
P(A_2B_2C_1) &= \sum\nolimits_{A_1B_1C_2} P(A_1A_2B_1B_2C_1C_2),\\
P(A_2B_2C_2) &= \sum\nolimits_{A_1B_1C_1} P(A_1A_2B_1B_2C_1C_2).
\end{aligned}
\tag{16}
$$

To be clear, taking the observable variables to be binary would mean that each equality in Eqs. (16) is shorthand for 8 distinct marginalization equations. For example the equality $P(A_1B_1C_1) = \sum\limits_{A_2B_2C_2} P(A_1A_2B_1B_2C_1C_2)$ is shorthand for 8 equations of the form

$$
p(A_1{=}a, B_1{=}b, C_2{=}c) = \sum_{a'b'c'} p(a_1a'_2b_1b'_2c_1c'_2) \tag{17}
$$

for each of the 8 possible values of the tuple $\{abc\}$.

Solving the marginal problem means eliminating all the gedankenprobabilities such as $p(a_1a'_2b_1b'_2c_1c'_2)$ from the system of inequalities and equalities. Practically, this is accomplished by linear quantifier elimination. As an equivalent description of the same problem, but could likewise consider the problem of determining the facets of the marginal polytope, the vertices of which are known to be given by the deterministic points (compare Fine's theorem [?] in the Bell scenario case).

Geometrically, linear quantifier elimination is equivalent to projecting a high-dimensional polytope in halfspace representation (inequalities and equalities) into a lower-dimensional quotient space.

Polytope projection is a well-understood problem in computational optimization, and a surprising variety of algorithms are available for the task [? ? ?]. The oldest-known method for polytope projection, i.e. linear quantifier elimination, is an algorithm known as Fourier-Motzkin (FM) elimination [? ?], although Fourier-Černikov elimination variant [? ?], as well as Block Elimination and Vertex Enumeration [?], are also fairly popular. More advanced polytope projection algorithms, such as Equality Set Projection (ESP) and Parametric Linear Programming, have also recently become available [? ? ?].

Linear quantifier elimination routines are available in many software tools[6]. We have found custom-coding an linear elimination routine in *Mathematica*^{TM} to be most efficient, see Appendix B for further detail.

Linear quantifier elimination is already widely used in causal inference to derive entropic inequalities [? ? ?]. In that task, however, the quantifiers being eliminated are those entropies which refer to hidden variables. By contrast, the probabilities we consider here are exclusively in terms of observable variables right from the very start. The quantifiers we eliminate are the gedankenprobabilities, which are quite different from probabilities involving hidden variables.

Although linear quantifier elimination can be highly optimized, it can still prove computationally difficult. It is therefore sometimes useful to consider relaxations of the marginals problem. The *full* marginals problem is to find inequalities on the marginal distributions such that the inequalities are satisfied *if and only if* the given marginal distributions can be extended. It is much easier to generate necessary-but-insufficient inequalities, i.e. satisfied by all compatible marginal distributions but such that no-violation does not certify compatibility. We have identified a technique for rapidly generating such quantifier-free inequalities by restricting the search to inequalities of a very particular form. We found this alternative technique — trading generality for speed — to be extraordinarily practical. The type of inequalities that we consider are given by a certain class of tautologies in classical propositional logic, see Sec. VIII for further details.

As far as we can tell, our inequalities are not related to the nonlinear infeasibility criteria which have been derived specifically to constrain classical networks [? ? ?], nor to the nonlinear inequalities which account for interventions to a given causal structure [? ?].

---

[6] For example *MATLAB*^{TM}'s `MPT2`/`MPT3`, *Maxima*'s `fourier_elim`, *lrs*'s `fourier`, or *Maple*^{TM}'s (v17+) `LinearSolve` and `Projection`. The efficiency of most of these software tools, however, drops off markedly when the dimension of the final projection is much smaller than the initial space of the inequalities. FM elimination aided by Cernikov rules [? ?] is implemented in *qskeleton* [?]. ESP [? ? ?] is supported by `MPT2` but not `MPT3`, and by the (undocumented) option of projection in the *polytope* (v0.1.1 2015-10-26) python module.

V. EXAMPLES OF POLYNOMIAL INEQUALITIES FOR THE TRIANGLE SCENARIO

ributions

T: why not just refer to the list of ineqs here, state that we've done a successful quantifier elimination to demonstrate its practicality, and discuss one or two of the most interesting ineqs?

Here are some examples of causal infeasibility criteria for the Triangle scenario which we can derive by considering inflation DAGs.

The nontrivial polynomial inequality

$$p(a) + p(b) + p(c) \leq 1 + p(a)p(b) + p(ac) + p(bc) \tag{18}$$ {eq:tritr

is found to be a constraint on the Triangle scenario through by summing the follow two inequalities

$$\begin{aligned} &0 \leq 1 - p(a_2) - p(b_1) - p(c_1) + p(a_2b_1) + p(a_2c_1) + p(b_1c_1) - p(a_2b_1c_1) \quad \left[= p(\bar{a}_2\bar{b}_1\bar{c}_1)\right] \\ &0 \leq p(a_2b_1c_1) \end{aligned} \tag{19}$$ {eq:trisi

subject to the following transformations

$$\underbrace{p(a_2b_1) \rightarrow p(a_2)p(b_1)}_{\text{Factorization relation, per Eqs. (A-4).}} \qquad \text{and} \qquad \underbrace{\begin{array}{c} p(a_2) \rightarrow p(a)\,,\ p(b_1) \rightarrow p(b)\,,\ p(c_1) \rightarrow p(c), \\ p(a_2c_1) \rightarrow p(ac),\ p(b_1c_1) \rightarrow p(bc) \end{array}}_{\text{Mapping relations, per Eqs. (A-6)}}. \tag{20}$$

Indeed, this example has been chosen because Eq. (18) can be derived directly from the small ancestral subgraph of $\{A_2B_1C_1\}$, namely Fig. 6.

It may be constructive to rewrite this inequality in correlator form using outcomes in $\{-1, +1\}$ as

$$\langle AC\rangle - \langle BC\rangle \leq 1 - \langle A\rangle\langle B\rangle,$$

which indicates that like Eq. (10), it is also a sort of monogamy inequality: it is impossible for A and C to be strongly correlated while B and C are strongly anticorrelated.

A consequence of Eq. (18) is that the perfect correlation distribution per Eq. (9) is found to be unrealizable from the Triangle scenario. This conclusions follows by considering the special case of Eq. (18) where $a{\rightarrow}1$, $b{\rightarrow}1$, $c{\rightarrow}0$.

Slightly more involved but otherwise analogous considerations give rise to the inequality

$$p(a_2b_2c_2) \leq p(\bar{a}_1\bar{b}_1\bar{c}_1) + p(a_1b_2c_2) + p(a_2b_1c_2) + p(a_2b_2c_1) \tag{21}$$ {eq:trifa

which in turn yields

$$p(a)p(b)p(c) \leq p(\bar{a}\bar{b}\bar{c}) + p(c)p(ab) + p(b)p(ac) + p(a)p(bc) \tag{22}$$ {eq:Fritz

per

$$\underbrace{\begin{array}{l} p(a_1b_2c_2) \rightarrow p(c_2)p(a_1b_2) \\ p(a_2b_1c_2) \rightarrow p(a_2)p(b_1c_2) \\ p(a_2b_2c_1) \rightarrow p(b_2)p(a_2c_1) \\ p(a_2b_2c_2) \rightarrow p(a_2)p(b_2)p(c_2) \end{array}}_{\text{Factorization relations}} \qquad \text{and} \qquad \underbrace{\begin{array}{l} p(a_1b_2) \rightarrow p(ab) \\ p(a_2c_1) \rightarrow p(ac) \\ p(b_1c_2) \rightarrow p(bc) \\ p(\bar{a}_1\bar{b}_1\bar{c}_1) \rightarrow p(\bar{a}\bar{b}\bar{c}) \end{array}}_{\text{Nontrivial mapping relations}}. \tag{23}$$

A consequence of Eq. (22) is that the W-type distribution per Eq. (11) is found to be unrealizable from the Triangle scenario, by considering the special case of Eq. (22) where $a{\rightarrow}1$, $b{\rightarrow}1$, $c{\rightarrow}1$. Eq. (22) requires the use of a broadcasting inflation, and therefore does not hold in the context of general probability theories.

VI. INEQUALITIES FOR THE BIPARTITE BELL SCENARIO

To further illustrate our inflation-DAG approach, we demonstrate how to recover all Bell inequalities [? ? ?] Brunner2013Bell,bell1966 via our method. To keep things simple we only discuss the case of a bipartite Bell scenario with two possible "settings" here, but the cases of more settings and/or more parties are totally analogous.

Consider the causal structure associated to the Bell/CHSH [? ? ? ?] experiment [? (Fig. E#2), ? (Fig. 19), ? (Fig. 1), ? (Fig. 1), ? (Fig. 2b), ? (Fig. 2)], depicted here in Fig. 7. The observable variables are A, B, X, Y, and Λ is the latent common cause of A and B.

In the Bell scenario DAG, one usually works with the conditional distribution $P(AB|XY)$, which is an array of distributions over A and B indexed by the possible values of X and Y, instead of with the original distribution $P(ABXY)$. The maximal pre-injectable sets then are

$$\begin{aligned}
&\{A_1B_1, X_1X_2Y_1Y_2\}\\
&\{A_1B_2, X_1X_2Y_2Y_2\}\\
&\{A_2B_1, X_1X_2Y_2Y_2\}\\
&\{A_2B_2, X_1X_2Y_2Y_2\},
\end{aligned}$$

where we have put commas in order to clarify that every maximal pre-injectable set contains *all* "settings" variables. Using These pre-injectables result in the given marginal distributions to take on the form

$$\begin{aligned}
P_{A_1B_1X_1X_2Y_1Y_2}(abx_1x_2y_1y_2) &= P_{ABXY}(abx_1y_1)P_X(x_2)P_Y(y_2),\\
P_{A_1B_2X_1X_2Y_1Y_2}(abx_1x_2y_1y_2) &= P_{ABXY}(abx_1y_2)P_X(x_2)P_Y(y_1),\\
P_{A_2B_1X_1X_2Y_1Y_2}(abx_1x_2y_1y_2) &= P_{ABXY}(abx_2y_1)P_X(x_1)P_Y(y_2),\\
P_{A_2B_2X_1X_2Y_1Y_2}(abx_1x_2y_1y_2) &= P_{ABXY}(abx_2y_2)P_X(x_1)P_Y(y_1),\\
P_{X_1X_2Y_1Y_2}(x_1x_2y_1y_2) &= P_X(x_1)P_X(x_2)P_Y(y_1)P_Y(y_2).
\end{aligned}$$

Although the last equations follows from each of the others together with the Markov condition $P_{XY} = P_X P_Y$ for the original distribution, it nevertheless turns out to be useful to write down explicitly: putting $x_1 = y_1 = 0$ and $x_2 = y_2 = 1$ and dividing the first equation by the last one results in

$$P_{ABXY}(ab|00) = \sum_{a',b'} P_{A_1A_2B_1B_2X_1X_2Y_1Y_2}(aa'bb'|0101),$$

and similarly also $P_{ABXY}(ab|01)$, $P_{ABXY}(ab|10)$ and $P_{ABXY}(ab|11)$ can be written as marginals of a conditional distribution. By Fine's Theorem [?], this implies the existence of a hidden-variable model. Conversely, if a hidden-variable model exists, then the existence of the inflation model implies the existence of a solution to the marginal problem. In conclusion, we therefore find in the case of the inflation DAG Fig. 8, our method provides a necessary and sufficient condition for feasibility with the Bell causal structure. In particular, the marginal problem on this inflation DAG reproduces *all* Bell inequalities.

VII. GEDANKENPROBABILITIES AND THE QUANTUM NO-BROADCASTING THEOREM

sicallity

It is worth emphasizing that broadcasting gedankenprobabilities are strictly classical constructs. If the latent node in the Bell scenario in Fig. 7 is allowed to be a quantum resource $\mathcal{H}^{d_A \otimes d_B}$, for example, then broadcasting gedankendistributions such as $P(A|x, A|\overline{x}, ...)$ or $P(A_1, A_2, ...)$ are **physically prohibited** if the quantum state is suitably entangled.

More precisely, quantum states are governed by a no-broadcasting theorem [? ?]: If half the state is sent to Alice and she performs some measurement on it, she fundamentally perturbs the state by measuring it. Post-measurement, that half of the state cannot be "re-sent" to Alice, that she might re-measure it using a different measurement setting. As a consequence of the no-broadcasting theorem, in the inflation DAG picture a quantum state which was initially available to a single party cannot be distributed both to Alice-copy-#1 and Alice-copy-#2 in the way a classical hidden variable could be. More generally, there is an analogous no-broadcasting theorem in the regime of epistemically-restricted general probabilistic theories (GPTs) [? ? ? ?], so that this impossibility also holds in many theories other than quantum theory.

This means that considerations on inflation DAGs cannot be used to derive quantum causal infeasibility criteria whenever a gedankenprobability presupposes the ability to broadcast a latent node's system. Broadcasting and non-broadcasting sets of variables are distinguished per Eq. (8).

Not every inflation requires broadcasting, however, and hence not every gedankenprobability is physically prohibited by quantum theory. Fig. 6 is an example of a non-broadcasting inflation. Constraints derived from non-broadcasting inflations are valid even in the GPT paradigm. Consequently the inequality in Eq. (18), which was derived from

Fig. 6, is therefore a causal infeasibility criterion which holds for the Triangle scenario even when the latent nodes are allowed to be quantum states. This confirms our numerical computations, which indicated that (eq:tritrivial1 18) does not have any quantum violations. The same is true for monogamy of correlations, per Eq. (10). Since the GHZ-type distribution Eq. (9) violates both of these inequalities, it is forbidden even per the relaxed GPT Triangle scenario, as was pointed out earlier.

It might also be possible to derive quantum causal infeasibility criteria if one appropriately modifies Step 1 to generate a different initial set of nonnegativity inequalities. This new set should capture the nonnegativity of only quantum-physically-meaningful marginal probability distributions. From this perspective, a broadcasting inflation DAG is an abstract logical concept, as opposed to a hypothetical physical construct. Indeed, the distributions in a quantum broadcasting inflation DAG can be characterized in terms of the logical broadcasting maps of Coecke2011 [?]. Note that $p(a_1, a_2)$ and other broadcasting-implicit gedankenprobabilities can be *negative* pursuant to a logical broadcasting map, and hence Step 1 in Sec. III would need to be modified.

An analysis along these lines has already been carried out successfully by Chaves2015infoquantum [?] in the derivation of entropic inequalities for non-classical causal structures. Although Chaves2015infoquantum [?] do not invoke inflation DAGs, they do employ conditional structure, which therefore gives rise to broadcasting sets. Chaves2015infoquantum [?] take pains to avoid including broadcasting gedankenentropies in any of their initial entropic inequalities, precisely as we would want to do in constructing our initial probability inequalities. Unlike entropic inequalities, the derivation of probability inequalities has not yet been achieved for non-classical causal structures other than Bell scenarios.

Our current inflation DAG method can be employed to derive causal infeasibility criteria for general causal structures, thus generalizing Bell inequalities somewhat. From a quantum foundations perspective, however, generalizing Tsirelson inequalities Tsirelson1980,Brunner2013Bell [? ?]—the ultimate constraints on what quantum theory makes possible—is even more desirable. Deriving additional quantum causal infeasibility criteria for general causal structure is therefore a priority for future research. T: move this to conclusions?

VIII. TAUTOLOGIES OF THE SUPPLEMENTED EXCLUDED MIDDLE

sec:TSEM

The process of linear quantifier elimination, while orders of magnitude more computationally amenable than its nonlinear variant, is nevertheless nontrivial. When the number of observable random variables in the inflation DAG is too large, it quickly happens that even the linear quantifier elimination algorithms may be too slow for practical application. To this end we have developed a strategy that identifies only one particular class of causal infeasibility criteria, but does so nearly instantly.

In this approach we construct polynomial inequalities corresponding to some tautology of the excluded middle (TEM). In classical logic, the TEM refers to that self-evident truism that every proposition is either true or false, and hence

$$\mathsf{True} \iff \mathsf{Or}\big[A{=}a, A{\neq}a\big] \;. \tag{24}$$

Note that we treat statistical events such as random variables yielding particular outcomes as identically logical propositions. The TEM can also be supplemented with certain "givens", which we take to be known-true propositions on the left-hand-side. These "givens" can then be interspersed throughout the right-rand-side while still yielding a valid tautology. Thus

$$\mathsf{And}\big[\mathbf{a}, \mathbf{b}\big] \implies \mathsf{Or}\begin{bmatrix}\mathsf{And}\big[\mathbf{a}, c\big]\,, \\ \mathsf{And}\big[\mathbf{b}, \bar{c}\big]\end{bmatrix} \tag{25}$$ {eq:TSEMd

is an example of a tautology of the *supplemented* excluded middle (T*S*EM). For pedagogical clarity we color and bold the "given" outcomes on both sides of the tautology.

Every TSEM can be converted into a linear inequality by virtue of two connections between classical logical and probability:

1. As the antecedent always implies the consequent, the probability of the antecedent is necessarily less-than-or-equal-to the probability of the consequent. If $j \implies k$ then $p(j) \leq p(k)$.
2. The probability of a disjunction of events is less-than-or-equal-to the sum of the probabilities of the individual events, i.e. $p(j \vee k) = p(j) + p(k) - p(j, k) \leq p(j) + p(k)$.

The inequality which would correspond to the TSEM in Eq. (25) is

$$p(ab) \leq p(ac) + p(b\bar{c})\,. \tag{26}$$

TSEM inequalities can be used as precursors for polynomial inequalities by applying them to inflation DAGs. For example the TSEM inequality

$$p(a_2 b_1) \leq p(a_2 b_1) + p(\bar{b}_1 c_1) \tag{27}$$

applied to our inflation of the Triangle scenario would imply the causal infeasibility criterion

$$p(a)p(c) \leq p(ab) + p(\bar{b}c) \tag{28}$$

pursuant to Fig. 6.

In fact almost every causal infeasibility criterion in this paper is a TSEM inequality! The Triangle scenario inequality Eq. (21) corresponds to the logical TSEM

$$\mathsf{And}\left[\mathbf{a_2}, \mathbf{b_2}, \mathbf{c_2}\right] \implies \mathsf{Or}\begin{bmatrix} \mathsf{And}\left[\bar{a}_1, \bar{b}_1, \bar{c}_1\right], \\ \mathsf{And}\left[a_1, \mathbf{b_2}, \mathbf{c_2}\right], \\ \mathsf{And}\left[\mathbf{a_2}, b_1, \mathbf{c_2}\right], \\ \mathsf{And}\left[\mathbf{a_2}, \mathbf{b_2}, c_1\right] \end{bmatrix} \tag{29}$$

and then Bell scenario inequality in ?? follows from the logical TSEM

$$\mathsf{And}\left[\mathbf{a_1}, \mathbf{b_1}, \mathbf{x_1}, \mathbf{\bar{x}_2}, \mathbf{y_1}, \mathbf{\bar{y}_2}\right] \implies \mathsf{Or}\begin{bmatrix} \mathsf{And}\left[\bar{a}_2, \mathbf{b_1}, \mathbf{x_1}, \mathbf{\bar{x}_2}, \mathbf{y_1}, \mathbf{\bar{y}_2}\right] \\ \mathsf{And}\left[\mathbf{a_1}, \bar{b}_2, \mathbf{x_1}, \mathbf{\bar{x}_2}, \mathbf{y_1}, \mathbf{\bar{y}_2}\right] \\ \mathsf{And}\left[a_2, b_2, \mathbf{x_1}, \mathbf{\bar{x}_2}, \mathbf{y_1}, \mathbf{\bar{y}_2}\right] \end{bmatrix} \tag{30}$$

etc.

Generating TSEM inequalities is computationally much easier than performing quantifier elimination. A practical alternative to Sec. III, then, is the four-step process

1. Generate TSEM inequalities on the inflation DAG.
2. Factor the probabilities (same as Step 2 in Sec. III).
3. Map as many probabilities to the original DAG as possible (same as Step 3 in Sec. III).
4. Discard any remaining inequalities which involve unmappable gedankenprobabilities.

Restricting one's search to TSEM inequalities makes deriving causal infeasibility criteria extremely tractable even for large DAGs, but a consequence, however, is that only one class of possible inequalities is being considered.

In order to make a TSEM inequality as compelling as possible one should ensure that every event on the left hand side is also referenced at least once on the right hand side. If one find a TSEM inequality lacking this property then one should tighten the inequalities by deleting any events which occur only on the left hand side. This preserves the tautology while increasing the probability of the left hand side.

We note that the connection between classical propositional logic and linear inequalities has been used previously in the task of causal inference. We reiterate that inequalities resulting from propositional logic, however, are a subset of the inequalities that result from linear quantifier elimination. Consequently, linear quantifier elimination is the preferable tool for deriving inequalities whenever the elimination is computationally tractable. Noteworthy examples of works deriving causal infeasibility criteria via classical logic are [?] and [?]: Eq. (19) here corresponds to Eq. (2-4) in Ref. [?], and ?? here corresponds to Eq. (30) in Ref. [?].

IX. CONCLUSIONS

Our main contribution is a new way of deriving causal infeasibility criteria, namely the inflation DAG approach. An inflation DAG naturally implies nonlinear hybrid inequalities, i.e. containing gedankenprobabilities, which implicitly constrain the set of distribution consistent with the original causal structure. If desirable, one can further eliminate the gedankenprobabilities via quantifier elimination. Polynomial inequalities can be obtained through linear elimination techniques.

These inequalities are necessary but not sufficient conditions on a joint distribution for it to be explained by the causal structure. We've seen that they can be stronger than entropic inequalities sometimes, yet just how strong they are is still unclear. A distribution might satisfy all our polynomial inequalities and yet not be realizable from the causal structure. Our methods yields tight causal infeasibility criteria for Bell scenarios, but those scenarios are exceptional in that the sets of realizable distributions form a convex polytope.

The most elementary of all causal infeasibility criteria are the conditional independence (CI) relations. Our method explicitly incorporates all marginal independence relations implied by a causal structure. We have found that generic

CI relations also appear to be implied by our polynomial inequalities. In future research we hope to clarify the process through which CI relations are manifested as properties of the inflation DAG.

A single causal structure has unlimited potential inflations. Selecting the right inflation from which to derive polynomial inequalities is an interesting challenge. To this end, it would be desirable to understand how particular features of the original causal structure are exposed when different nodes in the DAG are duplicated. By isolating which features are exposed in each inflation we could conceivably quantify the causal inference strength of each inflation. In so doing, we might find that inflated DAGs beyond a certain level of variable duplication need not be considered. The multiplicity beyond which further inflation is irrelevant is presumably related to the maximum degree of those polynomials which tightly characterize a causal scenario. Presently, however, it is not clear how to upper bound either number.

Our method turns the quantum no-broadcasting theorem NoCloningQuantum1996,NoCloningGeneral2006 [? ?] on its head by emphasizing that classical hidden variables *can*, in-principle, be cloned. This classical cloning possibility motivates the inflation DAG method, and underpins the implied causal infeasibility criteria. We have speculated about generalizing our method to obtain causal infeasibility criteria that constitute necessary constraints even for *quantum* causal scenarios, a common desideratum in recent works fritz2012bell,pusey2014gdag,Chaves2015infoquantum,ChavesNoSignalling,BeyondBellIII [? ? ? ? ?]. It would be enlightening to understand the extent to which our (classical) polynomial inequalities are violated in quantum theory. A variety of techniques exist for estimating the amount by which a linear inequality NPA2008Long,I3322NPA1 [? ?] is violated in quantum theory, but finding the quantum violation of a *polynomial* inequality is a more challenging task NPAReview [?].

The difference between classical ontic-state duplication and quantum no-broadcasting makes the inequalities that result from our consideration to be especially suited for distinguishing the set of quantum-realizable distributions from its subset of classically-realizable distributions. Causal infeasibility criteria that are sensitive to the classical-quantum distinction are precisely the sort of generalizations of the Bell inequalities which are sought after, in order to study the quantum features of generalized causal scenarios. Entropic inequalities have been lacking in this regard fritz2012bell,pusey2014gda [? ? ?], and the inflation DAG considerations proposed here constitute a significant alternative strategy.

ACKNOWLEDGMENTS

Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Economic Development and Innovation.

Appendix A: Nonlinear Quantifier Elimination Approach

imination

In the main text we considered deriving constraints on the pre-injectable sets via the marginal problem. There are further implications in the inflation DAG, however, which can be leveraged to further constrain the marginal distributions consisted with the inflation DAG.

Conditional Independence Relations

The most familiar such further constraints are the conditional independence relations relating observable variables in the inflation DAG. Conditional independence relations are inferred by d-separation; if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are d-separated in the (inflation) DAG by $\boldsymbol{Z}$, then we infer $\boldsymbol{X}\perp\boldsymbol{Y}|\boldsymbol{Z}$. The d-separation criterion is explained at length in Refs. pearl2009causality,stud [? ? ? ?], we present it only informally here.

We first define the notion of **exclusive ancestors**, because the usual notion of ancestors is inclusive, i.e. $\mathsf{An}(\boldsymbol{Z})$ contains $\boldsymbol{Z}$. Let the exclusive ancestors of $\boldsymbol{Z}$ so that $\mathsf{An}(\backslash\boldsymbol{Z}) \coloneqq \mathsf{An}(\boldsymbol{Z}) \setminus \boldsymbol{Z}$. Next we define **conditioning a DAG** over $\boldsymbol{Z}$, i.e. $G_{\backslash\boldsymbol{Z}}$, which is formed by deleting any outgoing edges from $\boldsymbol{Z}$ in G. Conditional DAGs enable the definition of **conditional ancestors**, i.e.

$$\mathsf{An}_G(\boldsymbol{X}|\boldsymbol{Z}) \coloneqq \begin{cases} \mathsf{An}_{G_{\backslash\boldsymbol{Z}}}(\boldsymbol{X}) & \text{if } \mathsf{An}_{G_{\backslash\boldsymbol{Z}}}(\boldsymbol{X}) and \mathsf{An}_G(\backslash\boldsymbol{Z}) are disjoint \\ \mathsf{An}_{G_{\backslash\boldsymbol{Z}}}(\boldsymbol{X}) \bigcup \mathsf{An}_G(\backslash\boldsymbol{Z}) & \text{if } \mathsf{An}_{G_{\backslash\boldsymbol{Z}}}(\boldsymbol{X}) and \mathsf{An}_G(\backslash\boldsymbol{Z}) overlap. \end{cases} \tag{A-1}$$

Finally, $\boldsymbol{X}$ and $\boldsymbol{Y}$ are d-separated by $\boldsymbol{Z}$ iff $\mathsf{An}_G(\boldsymbol{X}|\boldsymbol{Z})$ and $\mathsf{An}_G(\boldsymbol{Y}|\boldsymbol{Z})$ are disjoint.

Every conditional independence relation can be translated into a nonlinear constraint on probabilities, as $\boldsymbol{X}\perp\boldsymbol{Y}|\boldsymbol{Z}$ implies $p(\boldsymbol{xy}|\boldsymbol{z}) = p(\boldsymbol{x}|\boldsymbol{z})p(\boldsymbol{y}|\boldsymbol{z})$ for all $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$. As we generally prefer to work with unconditional probabilities, we summarize as follows: If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are d-separated by $\boldsymbol{Z}$ then $p(\boldsymbol{xyz})p(\boldsymbol{z}) = p(\boldsymbol{xz})p(\boldsymbol{yz})$ for all $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$. Such nonlinear constraints can be incorporated as restrictions on distributions consistent with the inflation DAG, supplementing the more trivial marginal-problem type constraints, and hence implicitly further constraining the injectable sets. Nonlinear quantifier elimination can then be used to make the implied constraints explicit.

Many modern computer algebra systems have functions capable of tackling nonlinear quantifier elimination symbolically[7]. Currently, however, it is not practical to perform nonlinear quantifier elimination on large polynomial systems with many quantifiers. Nevertheless, the nonlinear constraints can be easily accounted for numerically. Upon substituting numeric values for all the injectable probabilities the former quantifier elimination problem is converted to a universal quantifier existence problem: Do there exist gedankenprobabilities that satisfy the full set of linear and nonlinear constraints? Most computer algebra systems can resolve such *satisfiability* questions quite rapidly[8].

TABLE I. A comparison of different approaches constraining the pre-injectable sets. The primary divide is quantifier elimination versus quantifier existence, with approaches being further subdivided into linear and nonlinear variants.

Approach	Also Known As	Difficulty	End Results
Nonlinear Quantifier Elimination	Resolving Partial Existential Closure	Very Hard	Necessary-but-not-Sufficient Polynomial Inequalities
Nonlinear Quantifier Satisfiability	Universal Existential Closure, Nonlinear Programming	Easy	Certify the infeasibility of a specific distribution
Linear Quantifier Elimination	Polytope Projection	Moderate	Necessary-but-not-Sufficient Polynomial Inequalities
Linear Quantifier Satisfiability	Linear Programming, Universal Existential Closure	Very Easy	Certify the infeasibility of a specific distribution
Alternate: Logical Tautologies	Combinatorial / Set-Theoretic Implications Partial Marginal Problem	Easy	Necessary-but-not-Sufficient Polynomial Inequalities

It is also possible to use a mixed strategy of linear and nonlinear quantifier elimination, such as ChavesPolynomial [?] advocates. The explicit results of Ref. ChavesPolynomial [?] are therefore consequences of any inflation DAG, achieved by applying a mixed quantifier elimination strategy.

Coinciding Marginal Distributions

In principle one can exploit any d-separation criteria from the inflation DAG in order to transform or supplement the inequalities. If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are d-separated by Z, then $p(\boldsymbol{xyz}) = \frac{p(\boldsymbol{xz})p(\boldsymbol{yz})}{p(\boldsymbol{z})}$ pearl2009causality,spirtes2011causation,studeny2005proba [? ? ? ?].

At the moment this is just a text dump from the earlier manuscript. Needs extensive cleanup!!!

These are the kinds of equalities and inequalities we can make use of. Move to last? These are sort of optional - we don't actually use them... Just as ancestral independence of nodes implies marginal independence of variables, so too d-seperation of nodes implies conditional independence. If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are don't share a common ancestor, then $\forall_{xy}\; p(\boldsymbol{xy}) = p(\boldsymbol{x})p(\boldsymbol{y})$. If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are d-separated by Z, then $\forall_{xyz}\; p(\boldsymbol{xyz})p(\boldsymbol{z}) = p(\boldsymbol{xz})p(\boldsymbol{yz})$ pearl2009causality,spirtes2011 [? ? ? ?]. For example, in Fig. 3

$$\begin{matrix} A_2 \Lsh\emptyset\Rsh \{B_1, C_2\} \\ B_2 \Lsh\emptyset\Rsh \{A_2, C_1\} \\ C_2 \Lsh\emptyset\Rsh \{A_1, B_2\} \end{matrix} \qquad \text{implies that} \qquad \begin{matrix} P(A_1A_2B_2C_2) = P(C_2)\times P(A_1A_2B_2) \\ P(A_2B_1B_2C_2) = P(A_2)\times P(B_1B_2C_2) \\ P(A_2B_2C_1C_2) = P(B_2)\times P(A_2C_1C_2) \\ P(A_2B_2C_2) = P(A_2)\times P(B_2)\times P(C_2) \end{matrix} \tag{A-2}$$ [eqs]{eqs

To be clear, each equality in Eqs. (A-2) implies many more equalities at the level of individual probabilities, such as $p(a_2b_1b_2c_2) = p(a_2)p(b_1b_2c_2)$, $p(a_2b_1b_2) = p(a_2)p(b_1b_2)$, $p(a_2b_1) = p(a_2)p(b_1)$, etc.

———————————- Deriving polynomial inequalities can be done in four steps.

1. Construct linear non-negativity inequalities.
2. Convert the inequalities to nonlinear ones via conditional independence relations.
3. Use injectable sets to connect the inequalities to the original DAG.
4. Perform quantifier elimination.

rateineqs **Step 1**: **Generate an initial set of linear inequalities**

TODO: Change to EQUATIONS and inequalities, so that the inequality-only picture follows from the original one... UNDERWAY. EDITS IN PROGRESS HERE.

---

[7] For example *Mathematica*^{TM}'s `Resolve` command, *Redlog*'s `rlposqe`, or *Maple*^{TM}'s `RepresentingQuantifierFreeFormula`, etc.

[8] For example *Mathematica*^{TM} `` Reduce`ExistsRealQ `` function. Specialized satisfiability software such as SMT-LIB's `check-sat` BarFT-SMTLIB [?] are particularly apt for this purpose. One can also exploit the fact than any nonlinear optimizer will return an error when a set of constraints cannot be satisfied. Nonlinear optimizers include *Maple*^{TM}'s `NLPSolve`, *Mathematica*^{TM}'s `NMinimize`, and dozens of free and commercial optimizers for *AMPL* and/or *GAMS*

In order to derive inequalities on the original DAG, we begin by composing a set of linear conditions pertaining to the observable variables in the inflated DAG. Let the initial set be the nonnegativity of probability for all possible assignments to all observable variables. For simplicity, we take all observable variables to be binary, but the derivation can easily be adjusted to account for any number of outcomes. Taking the six observable variables in Fig. 3 to be binary, however, lead to 64 distinct nonnegativity conditions, corresponding to $0 \leq P(A_1A_2B_1B_2C_1C_2)$. The starting nonnegativity inequalities therefore look like

$$\begin{aligned}
&0 \leq p(a_1a_2b_1b_2c_1c_2)\\
&0 \leq [p(\bar{a}_1a_2b_1b_2c_1c_2) =]\, p(a_2b_1b_2c_1c_2) - p(a_1a_2b_1b_2c_1c_2)\\
&0 \leq [p(a_1\bar{a}_2b_1b_2c_1c_2) =]\, p(a_1b_1b_2c_1c_2) - p(a_1a_2b_1b_2c_1c_2)\\
&\vdots\\
&0 \leq [p(\bar{a}_1\bar{a}_2b_1b_2c_1c_2) =]\, p(b_1b_2c_1c_2) - p(a_1b_1b_2c_1c_2) - p(a_2b_1b_2c_1c_2) + p(a_1a_2b_1b_2c_1c_2)\\
&0 \leq \left[p(\bar{a}_1a_2\bar{b}_1b_2c_1c_2) =\right] p(a_2b_2c_1c_2) - p(a_1a_2b_2c_1c_2) - p(a_2b_1b_2c_1c_2) + p(a_1a_2b_1b_2c_1c_2)\\
&\vdots
\end{aligned} \tag{A-3}$$

etc. These inequalities are found by iterating over the general definition of $p(\boldsymbol{x}\overline{\boldsymbol{y}}) = p(\boldsymbol{x}) - p(\boldsymbol{xy})$ applied to all possible joint probabilities.

step:fac **Step 2**: **Infer factorization of probabilities from structural independence**

These "trivial" linear inequalities on the inflated DAG can be made into (weak) nontrivial polynomial inequalities by accounting for the marginal independence of certain random variable subsets. Independence of various distributions implies factorization of joint probabilities; the factorization of probabilities is the first step in transforming our set of otherwise-linear conditions.

We inspect the inflated DAG to find pairwise ancestral independences. In Fig. 3, for example, there are six pairs of ancestrally independent individual nodes, which consequently imply many marginal independences.

$$\begin{array}{l} A_2 \Lsh\emptyset\Rsh \{B_1, C_2\}\\ B_2 \Lsh\emptyset\Rsh \{A_2, C_1\}\\ C_2 \Lsh\emptyset\Rsh \{A_1, B_2\}\end{array} \quad \text{implies that} \quad \begin{array}{l} P(A_1A_2B_2C_2) = P(C_2) \times P(A_1A_2B_2)\\ P(A_2B_1B_2C_2) = P(A_2) \times P(B_1B_2C_2)\\ P(A_2B_2C_1C_2) = P(B_2) \times P(A_2C_1C_2)\\ P(A_2B_2C_2) = P(A_2) \times P(B_2) \times P(C_2)\end{array} \tag{A-4}$$ [eqs]{eq:

etc. Interestingly, the original DAG in Fig. 1 does not imply any observable marginal independence relations, nor any observable conditional independence relations at all, for that matter.

By accounting for marginal independence relations, the set of linear inequalities is transformed into a set of polynomial inequalities. Eqs. (A-4) imply factorization of many different probabilities, such as $p(a_2b_1b_2c_2) = p(a_2)p(b_1b_2c_2)$, $p(a_2b_1b_2) = p(a_2)p(b_1b_2)$, $p(a_2b_1) = p(a_2)p(b_1)$, etc. In principle one can exploit any d-separation criteria from the inflation DAG in order to transform or supplement the inequalities. If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are d-separated by Z, then $p(\boldsymbol{xyz}) = \frac{p(\boldsymbol{xz})p(\boldsymbol{yz})}{p(\boldsymbol{z})}$ pearl2009causality,spirtes2011causation,studeny2005probabilistic,koller2009probabilistic [? ? ? ?]. For simplicity, however, we have restricted our attention here to exclusively ancestral independences.

p:findmap **Step 3**: **Map marginal distributions in the inflated DAG to those in the original structure**

Here we account for injectable sets. As any subset of an injectable set is also an injectable set, we enumerate here only *maximal* injectable sets. The maximal injectable sets in Fig. 3 are

$$\begin{array}{l} \{A_1B_2\},\\ \{B_1C_2\},\\ \{A_2C_1\},\\ \text{and } \{A_1B_1C_1\},\end{array} \quad \text{which in turn imply that} \quad \begin{array}{l} P(A_1B_2) = P(AB),\\ P(B_1C_2) = P(BC),\\ P(A_2C_1) = P(AC),\\ \text{and } P(A_1B_1C_1) = P(ABC).\end{array} \tag{A-5}$$ [eqs]{eq:

The injectable sets allow us to transform our set of inequalities once more, this time into inequalities which have bearing on the original causal structure. We call this transformation of probabilities the **injection map**; the injection map pursuant to Eqs. (A-5) takes

$$\underbrace{\begin{array}{l} \{p(a_1) = p(a_2)\} \to p(a)\\ \{p(b_1) = p(b_2)\} \to p(b)\\ \{p(c_1) = p(c_2)\} \to p(c)\end{array}}_{\text{By the } \textit{definition} \text{ of inflation.}} \qquad \underbrace{\begin{array}{l} \{p(a_1b_1) = p(a_1b_2)\} \to p(ab)\\ \{p(a_1c_1) = p(a_2c_1)\} \to p(ac)\\ \{p(b_1c_1) = p(b_1c_2)\} \to p(bc)\\ p(a_1b_1c_1) \to p(abc)\end{array}}_{\text{Via the injectable sets.}} \tag{A-6}$$ [eqs]{eq:

After applying the injection map we are left we a system of **hybrid inequalities** of sorts, simultaneously containing two radically different kinds of probabilities. Some probabilities now pertain to the original DAG, but they appears alongside many probabilities which *do not inject* to the original DAG. Probabilities which cannot be related to the original causal structure include $\{p(a_1b_1c_2), p(a_1b_2c_1), p(a_2b_1c_1)\}$, and any joint probability which references more than one instance of a duplicated variable, such as $\{p(a_1a_2), p(a_1a_2b_1), ...\}$. The probabilities pertaining to the inflated DAG which have no parallel in the original DAG are precisely probabilities regarding non-injectable sets of variables.

We name these non-injectable probabilities **gedankenprobabilities**, as they could be measured in-principle if one were to physically construct the inflated causal structure[9]. As we are really only concerned with the original DAG, however, these "unmeasured" joint distributions are effectively just thought experiments. The in-principle existence of gedankenprobabilities, however, is critical to inferring causal infeasibility criteria for the original DAG.

As any superset of a non-injectable set is also not injectable, we enumerate here only *minimal* non-injectable sets. The *minimal* non-injectable sets in Fig. 3 are

$$\{A_1A_2\}, \quad \{B_1B_2\}, \quad \{C_1C_2\}, \quad \{A_1B_1C_2\}, \quad \{A_1B_2C_1\}, \quad \{A_2B_1C_1\}. \tag{A-7}$$ {eq:corei

Eq. (A-7) implies that our set of hybrid inequalities will be riddled with 34 different gedankenprobabilities.

imination **Step 4**: **Quantifier elimination of the gedankenprobabilities**

We can infer implications for the original random variable from the system of hybrid inequalities obtained after Step 3. This inference task is essentially a form of quantifier elimination, where the quantifiers to be eliminated are the gedankenprobabilities. Thus, the final step toward obtaining the desired causal infeasibility criteria is to eliminate the gedankenprobabilities from our system of hybrid polynomial inequalities. This quantifier elimination problem is well defined mathematically, although it is a challenging problem when the quantifiers are related nonlinearly.

Many modern computer algebra systems have functions capable of tackling this sort of problem fully symbolically[10]. Currently, however, it is not practical to perform nonlinear quantifier elimination on large polynomial systems with many quantifiers. We consider, therefore, two other strategies for making effective use of the hybrid inequalities.

Firstly, one may substitute numeric values for all the injectable probabilities appearing in the polynomial inequality set. Upon doing so, the quantifier elimination problem is converted to a quantifier existence problem: Do there exist gedankenprobabilities that satisfy the resulting system of polynomial inequalities? Most computer algebra systems can resolve such *satisfiability* questions quite rapidly[11].

Note that real-world data with uncertainties can also be incorporated into these satisfiability questions. Instead of asserting that a particular probability is equal to a given *value*, one can incorporate new inequalities which constrain the experimentally-known probabilities to lie in given *intervals*. Assigning probabilities to intervals as opposed to numeric values results in further free parameters in the system, but the problem nevertheless remains one of *universal* existential closure, and can be efficiently tested.

Appendix B: Further Polytope Projection Algorithms

lgorithms

The Equality Set Projection (ESP) algorithm [? ?] (jones2004equality,JonesThesis2005) is ideal for handling inflation DAGs, because its computational complexity scales only according to the facet count of the final projection. Our use of larger-and-larger inflation DAGs to obtain causal infeasibility criteria on the same underlying original DAG means that while the complexity of the starting polytope is unbounded, the complexity of the projection is finite. Practically, this suggests that the ESP algorithm could parse the implications due to a very large inflation DAGs efficiently. Formally, ESP should require minimal computational overhead to consider a larger inflation DAG relative to considering a much smaller inflation DAG, when the *implications* of the small and large inflations are similar. By contrast, the computation complexity of Fourier-Motzkin (FM) elimination algorithm scales with the number of quantifiers being eliminated. The number of gedankenprobabilities requiring elimination is exponentially related to the number of variables in the inflation DAG. The FM algorithm, therefore, is utterly impractical very for large inflation DAGs.

Another positive feature of the ESP algorithm is that it commences outputting quantifier-free inequalities immediately, and terminates upon deriving the complete set of inequalities. By contrast, FM works by eliminating one quantifier at a time. Terminating the ESP algorithm before it reaches completion would result in an incomplete list of inequalities.

---

[9] The inflated causal structure is always hypothetically constructable, classically, hence the though-experiment terminology. A gedankenprobability is also a well-defined hypothetical in quantum theory whenever the variables comprise a non-broadcasting set.

[10] For example *Mathematica*^{TM}'s `Resolve` command, *Redlog*'s `rlposqe`, or *Maple*^{TM}'s `RepresentingQuantifierFreeFormula`, etc.

[11] For example *Mathematica*^{TM} `Reduce`ExistsRealQ` function. Specialized satisfiability software such as SMT-LIB's `check-sat` [? ] (BarFT-SMTLIB) are particularly apt for this purpose. One can also exploit the fact than any nonlinear optimizer will return an error when a set of constraints cannot be satisfied. Nonlinear optimizers include *Maple*$^{TM}$'s `NLPSolve`, *Mathematica*$^{TM}$'s `NMinimize`, and dozens of free and commercial optimizers for *AMPL* and/or *GAMS*

Even an incomplete list is valuable, though, since the causal infeasibility criteria we are deriving are anyways necessary but not sufficient.

Vertex projection (VP) algorithms are another computational tool which may be used to assist in linear quantifier elimination [? Avis2000lrs]. VP works by first enumerating the vertices of the initial polytope (H-rep to V-rep), projecting the vertices, and then converting back to inequalities (V-rep to H-rep). For generic high-dimensional polytopes, the operation of converting from a representation in terms of halfspaces to one in terms of extremal-vertices representations can be computationally costly (high-d H-rep to V-rep). Starting from a vertex representation in a high dimensional space, however, one can immediately determine the vertex representation of the polytope's projection in a lower dimensional space. The projection is along the coordinate axes, so one just "discards" the coordinate of the eliminated quantifier. To obtain the inequalities which characterize the projected polytope one then applies a convex hull algorithm to the projected vertices (low-d V-rep to H-rep).

For probability distributions, however, the extremal vertices are precisely the deterministic possibilities. Since the extremal vertices of the initial polytope are easily enumerated, it is possible to avoid the high-d V-rep to H-rep step entirely. There is a one-to-one correspondence between the inflation-DAG's initial generating inequalities and its initial extreme observable probability distributions. We used this V-rep to H-rep technique to project the initial polytope implied by Fig. 3 (Step 1) to an intermediate 23-dimensional polytope, where each of the 23 remaining can be mapped the original DAG. Only then did we apply the transformations of factorization and mappings (Steps. 2,3) to convert those linear inequalities to polynomial inequalities pertaining to the original DAG. We found that the V-rep to H-rep technique, using *lrs* [lrs], was orders-of-magnitude faster than FM elimination at obtaining the same result.

Yet another technique is also possible. Suppose the initial polytope is given by $\left\{\vec{x},\vec{y}\,|\,\hat{A}.\vec{x}+\hat{B}.\vec{y}\geq \boldsymbol{c}\right\}$, where y are the quantifiers. If we can find any completely nonnegative vector $\boldsymbol{w}$ such that $\boldsymbol{w}.\hat{B}=\vec{0}$ then we automatically establish the quantifier-free inequality $\boldsymbol{w}.\hat{A}.\vec{x}\geq \boldsymbol{w}.\boldsymbol{c}$. Solving for "random" nonnegative vectors $\boldsymbol{w}$ is easy; solving for all possible solutions is rather more difficult. [? BalasProjectionCone] refined this method so that each extremal construction of $\boldsymbol{w}$ corresponds to an irredundant inequality in the H-rep description of the projected polytope. Nevertheless, even without utilizing the full projection cone, this technique can be used to rapidly obtain a few quantifier-free inequalities.

Appendix C: Optimized Algorithm for Recognizing Redundant Inequalities

edundancy

When performing Fourier-Motzkin linear quantifier elimination one must periodically filter out redundant inequalities from the set of linear inequalities. Equivalently, the means identifying redundant halfspace constraints in the description of the polytope. An individual constraint in a set is redundant if it is implied by the other constraints.

An individual linear inequality is redundant if and only if it is a *positive* linear combination of the others [Thm. 5.8 in [? fordan1999projection]][12]. This is related to the V-rep characterization of polyhedral cones: If a cone is defined such that $W_{\hat{M}}\coloneqq\left\{\vec{x}\,|\exists_{\boldsymbol{v}\geq\boldsymbol{0}}:\,\hat{M}.\boldsymbol{v}=\vec{x}\right\}$ then $\vec{b}\in W_{\hat{M}}$ if and only if the linear system of equations $\hat{M}.\boldsymbol{v}=\vec{b}$ has a solution such that all the elements of $\boldsymbol{v}$ are nonnegative. Thus, the computational tool required is one which accepts as input the matrix $\hat{M}$ and the column vector $\vec{b}$ and returns $\vec{b}\in W_{\hat{M}}$ as True or False.

Below, we present two possible *Mathematica*^{TM} implementations which assess if a given column $\vec{b}$ can be expressed as a positive linear combination of the columns of $\hat{M}$. The former function is easy to understand, but the latter utilizes

[12] The "if" is obvious. The "only if" is a consequence of Farka's lemma [? fordan1999projection].

efficient low-level code and *Mathematica™*'s internal error-handling to rapidly recognize infeasible linear programs.

```
PositiveLinearSolveTest[M_?MatrixQ, b_]:=With[{vars = Thread[Subscript[x, Dimensions[M][[2]]]]},
     Resolve[Exists[Evaluate[vars], AllTrue[vars, NonNegative],
     And@@Thread[M.vars == Flatten[b]]]]];
```

or

```
PositiveLinearSolveTest[M_?MatrixQ, b_]/;Dimensions[b]==={Length[M], 1}:=
Module[{rowcount, columncount, fakeobjective, zeroescolumn},
 {rowcount, columncount} = Dimensions[M];
 fakeobjective = SparseArray[{}, {columncount}, 0.0]; zeroescolumn = SparseArray[{}, {rowcount, 1}];
 Internal`HandlerBlock[{Message, Switch[#1, Hold[Message[LinearProgramming::lpsnf, ___], _], Throw[False]]&},
     Quiet[Catch[
       LinearProgramming[fakeobjective, M, Join[b, zeroescolumn, 2], Method → Simplex]; True
     ], {LinearProgramming::lpsnf}]]];
```

To illustrate examples of a when a positive solution to the linear system exists and when it does not, consider the following two examples:.

$$\texttt{PositiveLinearSolveTest}[\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & -1\end{pmatrix}, \begin{pmatrix}1\\ -2\end{pmatrix}] \;==\; \texttt{False}$$

$$\texttt{PositiveLinearSolveTest}[\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & -2\end{pmatrix}, \begin{pmatrix}1\\ -1\end{pmatrix}] \;==\; \texttt{True}$$

If $\hat{A}$ is the matrix who's rows are nonnegativity inequalities, then the following test determines if row n is redundant.

```
RedundantRowQ[A_?MatrixQ, n_Integer]:=PositiveLinearSolveTest@@Reverse[Transpose/@TakeDrop[A, {n}]].
```

Note that a `True` response from `RedundantRowQ` indicates that the row n is redundant.

Appendix D: Recognizing observationally equivalent DAGs

One expects that an edge $A \to B$ can be added to DAG G while leaving G observationally invariant if the new connection does not introduce any new information about observable variables to B. We can formalize this notion in the language of sufficient statistics. To do so, however, a few background definitions are in order.

Perfectly Predictable: The random variable X is perfectly predictable from a set of variables $\boldsymbol{Z}$, hereafter $\boldsymbol{Z} \vDash X$, if X can be completely inferred from knowledge of $\boldsymbol{Z}$ alone. In a deterministic DAG, for example, every non-root node is perfectly predictable given its parents, $\mathsf{pa}[X] \vDash X$. Indeed, in a deterministic DAG the node X is perfectly predictable from $\boldsymbol{Z}$ if X is a deterministic descendant of $\boldsymbol{Z}$. Operationally, X is a deterministic descendant of Z if the intersection of [the ancestors of X] with [the non-ancestors of Z] is a subset of [the descendants of Z]. Happily though, perfectly predictability can be extrapolated from a causal structure with minimal effort: $\boldsymbol{Z} \vDash X$ if every directed path to X from any root node is blocked by $\boldsymbol{Z}$.

Markov Blanket: The Markov Blanket for a set of nodes $\boldsymbol{V}$, hereafter $\mathsf{MB}[\boldsymbol{V}]$, is the set of all of $\boldsymbol{V}$'s children, parents, and co-parents. The Markov Blanket is so defined because the nodes in $\boldsymbol{V}$ are conditionally independent of *everything* given $\mathsf{MB}[\boldsymbol{V}]$. If the random variables in the Markov Blanket $\mathsf{MB}[\boldsymbol{V}]$ are known, then information about nodes inside $\boldsymbol{V}$ has no bearing on nodes outside the Markov Blanket and vice versa.

Markov Partition: New! I made this up Nov 24. Useful do you think? A set of variables $\boldsymbol{Z}$ is a Markov Partition for a pair of random variables X and Y, hereafter $X \dashv \boldsymbol{Z} \vdash Y$, if the pair are conditionally independent of eachother given *any superset* of $\boldsymbol{Z}$. Operationally, this means that X and Y are d-separated by every superset of $\boldsymbol{Z}$. Equivalently, $X \dashv \boldsymbol{Z} \vdash Y$ if $\mathsf{MB}[\boldsymbol{V}] \subseteq \boldsymbol{Z}$ and $X \in \boldsymbol{V}$ while $Y \notin \boldsymbol{V}$, or if $\mathsf{MB}[\boldsymbol{V}] \subseteq \boldsymbol{Z}$ and $Y \in \boldsymbol{V}$ while $X \notin \boldsymbol{V}$. Happily though, Markov Partitions can be extrapolated from a causal structure with minimal effort: $X \dashv \boldsymbol{Z} \vdash Y$ if and only if X and Y would be in *disconnected components* under the deletion of all edges initiation from $\boldsymbol{Z}$.

Sufficient Statistic: A set of nodes $\boldsymbol{Z}$ is a sufficient statistic for A relative to X, hereafter $\boldsymbol{Z} \vdash A|X$, if and only if all inferences about X which can be made given knowledge of A are also inferable *without* knowing A but with knowing $\boldsymbol{Z}$ instead. In other words, learning A can never teach anything new about X if $\boldsymbol{Z}$ is already known. If $X = A$, then the *only way* $\boldsymbol{Z}$ can stand in for A when making inferences about A is if A is perfectly predicable given $\boldsymbol{Z}$, i.e. $\boldsymbol{Z} \vdash A|A \iff \boldsymbol{Z} \vDash A$. If $A \neq B$ then there are four and only four?) ways that $\boldsymbol{Z} \vdash A|X$ can be implied by a DAG: If $\boldsymbol{Z} \vDash A$, if $\boldsymbol{Z} \vDash X$, if $\mathsf{MB}[\boldsymbol{V}] \subseteq \boldsymbol{Z}$ and $A \in \boldsymbol{V}$ while $X \notin \boldsymbol{V}$, or if $\mathsf{MB}[\boldsymbol{V}] \subseteq \boldsymbol{Z}$ and $X \in \boldsymbol{V}$ while $A \notin \boldsymbol{V}$. Alternatively: If $A \neq B$ then there are THREE ways that $\boldsymbol{Z} \vdash A|X$ can be implied by a DAG: If $\boldsymbol{Z} \vDash A$, if $\boldsymbol{Z} \vDash X$, and if $A \dashv \boldsymbol{Z} \vdash X$.

dgeadding **Theorem 1.** *An edge $A \to B$ can be added to G without observational impact if $\mathsf{pa}[B]$ are a sufficient statistic for A relative to all observable nodes, i.e. $\forall_{observable\ X} : \mathsf{pa}[B] \vdash A|X$.*
In particular, the edge $A \to B$ can always be added whenever $\mathsf{pa}[B] \vDash A$, including, but not limited to, the instance $\mathsf{pa}[A] \subseteq \mathsf{pa}[B]$.
Furthermore, the edge $\Lambda \to B$ can be also always be added whenever Λ is latent and $\mathsf{MB}[\Lambda] \subseteq \mathsf{pa}[B]$.

We can also define an analogous condition for when an edge can be removed from a DAG without impacting it observationally.

edropping **Corollary 1.1.** *An edge $A \to B$ can be dropped from G to form G' such that G and G' are observationally equivalent if ~~and only if~~ the edge $A \to B$ can be added (back) to G' while leaving G' observationally invariant per Theorem 1.*

On the subject of adding observationally-invariant edges, it is important to recognize when latent nodes can be introduced (or dropped) without observational impact.

entadding **Theorem 2.** *A (root) latent node Λ can be removed from G without observational impact if Λ has only one child node and no co-parents (Λ is "equivalent to local randomness"), or if Λ's children are also all children of another single latent node (Λ is "covered-for by another latent node"). Conversely, a new root latent node Λ can be introduced along with various outgoing edges, without observational impact, if Λ would be equivalent to local randomness or covered-for by another latent node.*

Naturally, two causal structures are observationally equivalent if one can be transformed into the other without observational impact, via Theorems 1 and 2. Some examples of observationally equivalent scenarios, and the steps which interconvert them, are given in Fig. 12.

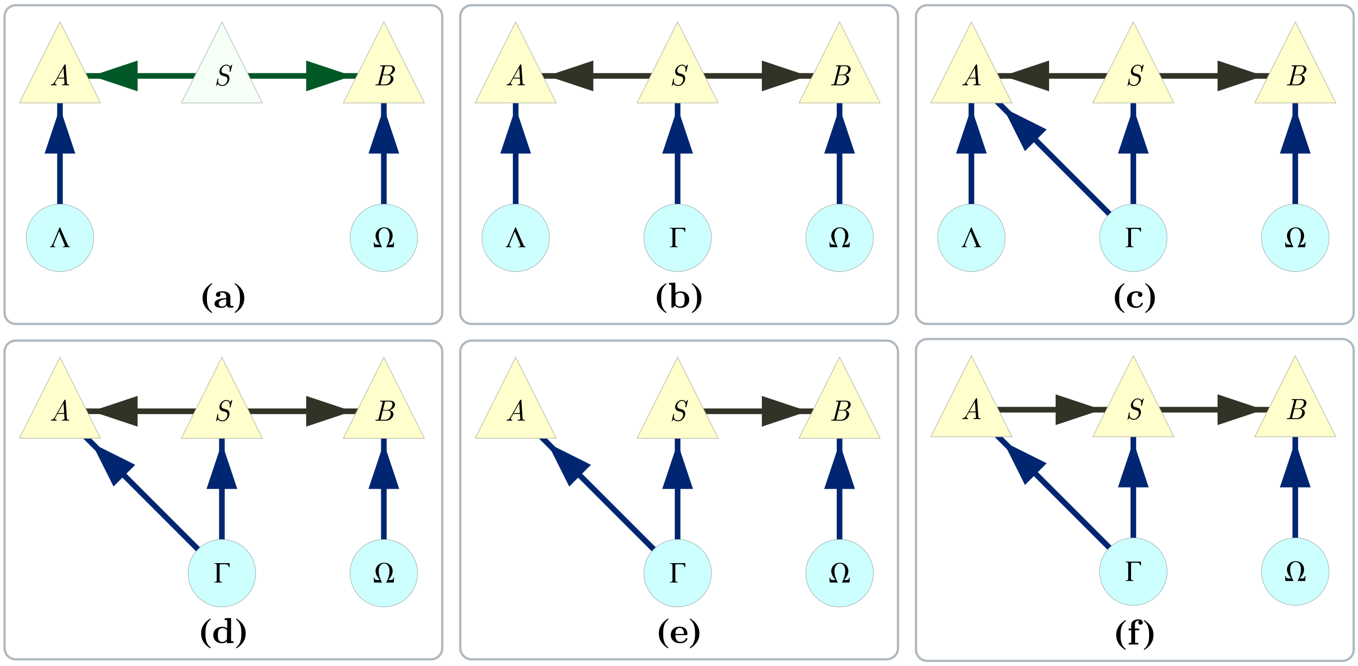

FIG. 12. A set of observational equivalent causal structures. The reasons the changes are observational invariant are as follows:
(a)$\sim$(b) because Γ is useless in (b), and as such Γ can be dropped from (b) per Theorem 2.
(b)$\sim$(c) because $\mathsf{MB}\left[\Gamma\right] \subseteq \mathsf{pa}\left[A\right]$ in (b), and as such $\Gamma \to A$ can be added to (b) per Theorem 1.
(c)$\sim$(d) because Λ is redundant to Γ in (c), and as such Λ can be dropped from (c) per Theorem 2.
(d)$\sim$(e) because $\mathsf{pa}\left[S\right] \subseteq \mathsf{pa}\left[A\right]$ in (e), and as such $S \to A$ can be added to (e) per Theorem 1.
(e)$\sim$(f) because $\mathsf{pa}\left[A\right] \subseteq \mathsf{pa}\left[S\right]$ in (e), and as such $A \to S$ can be added to (e) per Theorem 1. fig:equiv

Appendix E: Tobias's Original 7 Inequalities

"I present several inequalities... together with a method of proof which has a combinatorial flavour. No quantum violations of any of these inequalities has been found to date."

Theorem 3. *The following inequalities hold for all classical correlations in the triangle scenario:*

(a) $p(a)p(c) \leq p(ab) + p(\bar{b}c)$

(b) $p(ab\bar{c})p(a\bar{b}c)p(\bar{a}bc) \leq p(abc) + p(\bar{a}\bar{b})p(ab\bar{c})p(a) + p(\bar{a}\bar{c})p(a\bar{b}c)p(c) + p(\bar{b}\bar{c})p(\bar{a}bc)p(b)$

(c) $p(ab\bar{c})p(a\bar{b}c)p(\bar{a}bc) \leq p(abc)^2 + 2\big(p(\bar{a}\bar{b})p(ab\bar{c}) + p(\bar{a}\bar{c})p(a\bar{b}c) + p(\bar{b}\bar{c})p(\bar{a}bc)\big)$

(d) $p(abc)^2 p(\bar{a}\bar{b}\bar{c}) \leq p(ab\bar{c})p(a\bar{b}c)p(\bar{a}bc) + \big(2p(abc) + p(\bar{a}\bar{b}\bar{c})\big)\big(1 - p(abc) - p(\bar{a}\bar{b}\bar{c})\big)$

(e) $p(abc)^2 p(\bar{a}\bar{b}\bar{c}) \leq p(abc)^3 + \big(2p(abc) + p(\bar{a}\bar{b}\bar{c})\big)\big(1 - p(abc) - p(\bar{a}\bar{b}\bar{c})\big)$

(f) $p(a)p(b)p(c) \leq p(\bar{a}\bar{b}\bar{c}) + p(ab)p(c) + p(ac)p(b) + p(bc)p(a)$

(g) $p(a)p(b)p(c) \leq p(\bar{a}\bar{b}\bar{c})^2 + 2\Big(p(ab)p(c) + p(ac)p(b) + p(bc)p(a)\Big)$

It is quite likely that some of these inequalities are dominated by the others, but I do not know for sure whether any of them are actually redundant."

Note that Eq. (22) implies inequalities (a), (b), and (f). I haven't checked the others yet. ~EW

Appendix F: Some of Elie's Triangle Scenario Inequalities

The following inequalities follow from Fig. 3 using linear quantifier elimination. The list is incomplete; longer inequalities have been excised in order to fit the page. A denser but complete list is embedded as the final page of this PDF. All the inequalities below are nontrivial, and are also unique under permutation of the parties or inverting outcomes. In expectation-value formation we imagine the two possible outcomes of each random variables to be +1 and −1.

$$0 \le 1 + \langle AB\rangle + \langle AC\rangle + \langle B\rangle\langle C\rangle$$
$$0 \le 2 - 2\langle AC\rangle + \langle A\rangle\langle B\rangle\langle C\rangle - \langle C\rangle\langle AB\rangle$$
$$0 \le 2 - \langle ABC\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle + \langle A\rangle\langle BC\rangle$$
$$0 \le 2 + \langle ABC\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle - \langle A\rangle\langle BC\rangle$$
$$0 \le 3 + \langle A\rangle + \langle B\rangle + \langle C\rangle + 3\langle AB\rangle - \langle AC\rangle - \langle B\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle$$
$$0 \le 3 + \langle A\rangle + \langle B\rangle - \langle C\rangle + 3\langle AB\rangle + \langle AC\rangle + \langle B\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle - \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle$$
$$0 \le 3 + \langle C\rangle - 2\langle AB\rangle - 2\langle BC\rangle + \langle A\rangle\langle B\rangle + \langle A\rangle\langle B\rangle\langle C\rangle - \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle$$
$$0 \le 3 + \langle B\rangle + \langle AB\rangle - 2\langle BC\rangle - \langle A\rangle\langle B\rangle + \langle A\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle$$
$$0 \le 3 + \langle B\rangle + \langle AB\rangle - 2\langle BC\rangle + \langle A\rangle\langle B\rangle - \langle A\rangle\langle C\rangle - \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle$$
$$0 \le 3 - \langle A\rangle + \langle B\rangle + \langle C\rangle - \langle ABC\rangle + \langle A\rangle\langle B\rangle + \langle A\rangle\langle C\rangle - \langle B\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle + \langle A\rangle\langle BC\rangle$$
$$0 \le 3 + \langle A\rangle + \langle B\rangle - \langle C\rangle - 2\langle AB\rangle + \langle ABC\rangle + \langle A\rangle\langle B\rangle + \langle A\rangle\langle C\rangle + \langle B\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle - \langle A\rangle\langle BC\rangle$$
$$0 \le 4 + 2\langle C\rangle - 2\langle AB\rangle - 2\langle AC\rangle - \langle ABC\rangle + 2\langle A\rangle\langle B\rangle + 2\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle + \langle A\rangle\langle BC\rangle$$
$$0 \le 4 - 2\langle B\rangle - 2\langle AB\rangle - 3\langle BC\rangle + \langle ABC\rangle + \langle B\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle$$
$$0 \le 4 - 2\langle C\rangle - 2\langle AB\rangle - 2\langle AC\rangle - 3\langle BC\rangle + \langle ABC\rangle + 2\langle A\rangle\langle B\rangle + \langle B\rangle\langle C\rangle - \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle$$
$$0 \le 4 + 2\langle AB\rangle - 2\langle AC\rangle + \langle BC\rangle + \langle ABC\rangle + 2\langle A\rangle\langle B\rangle + 2\langle A\rangle\langle C\rangle - \langle B\rangle\langle C\rangle - \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle$$
$$0 \le 4 + 2\langle AB\rangle - 2\langle AC\rangle + \langle BC\rangle + \langle ABC\rangle - 2\langle A\rangle\langle B\rangle + 2\langle A\rangle\langle C\rangle - \langle B\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle$$
$$0 \le 4 - 2\langle AB\rangle + 3\langle BC\rangle + \langle ABC\rangle + 2\langle A\rangle\langle B\rangle + \langle B\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle - \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle$$
$$0 \le 4 - 2\langle B\rangle - 2\langle AB\rangle + \langle ABC\rangle - 2\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle - \langle A\rangle\langle BC\rangle$$
$$0 \le 4 - 2\langle A\rangle - 2\langle AB\rangle - 2\langle AC\rangle + \langle ABC\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle - \langle A\rangle\langle BC\rangle$$
$$0 \le 4 - 2\langle C\rangle - 2\langle AB\rangle - 2\langle AC\rangle - 2\langle BC\rangle + \langle ABC\rangle + 2\langle A\rangle\langle B\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle - \langle A\rangle\langle BC\rangle$$
$$0 \le 4 - 2\langle AB\rangle - 2\langle AC\rangle - 2\langle BC\rangle + \langle ABC\rangle + 2\langle A\rangle\langle B\rangle + 2\langle A\rangle\langle C\rangle + 2\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle + \langle A\rangle\langle BC\rangle$$
$$0 \le 5 + \langle A\rangle + \langle B\rangle + \langle C\rangle + 3\langle AB\rangle + \langle AC\rangle - 4\langle BC\rangle - 2\langle A\rangle\langle B\rangle + \langle B\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle$$
$$0 \le 5 + \langle A\rangle + \langle B\rangle + \langle C\rangle + 3\langle AB\rangle - \langle AC\rangle - 4\langle BC\rangle + 2\langle A\rangle\langle B\rangle - 2\langle A\rangle\langle C\rangle + \langle B\rangle\langle C\rangle - \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle$$
$$0 \le 5 + 3\langle A\rangle + \langle B\rangle + \langle C\rangle + \langle AB\rangle + 3\langle AC\rangle + \langle BC\rangle - \langle ABC\rangle + 2\langle A\rangle\langle B\rangle + 2\langle A\rangle\langle B\rangle\langle C\rangle - 2\langle C\rangle\langle AB\rangle$$
$$0 \le 5 + 3\langle A\rangle + \langle B\rangle + \langle C\rangle + \langle AB\rangle + 3\langle AC\rangle + \langle BC\rangle - \langle ABC\rangle + 2\langle A\rangle\langle B\rangle - 2\langle A\rangle\langle B\rangle\langle C\rangle + 2\langle C\rangle\langle AB\rangle$$
$$0 \le 6 - 3\langle AB\rangle - 4\langle AC\rangle + \langle A\rangle\langle B\rangle + 2\langle A\rangle\langle C\rangle + 2\langle B\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle - \langle C\rangle\langle AB\rangle - 2\langle B\rangle\langle AC\rangle - 2\langle A\rangle\langle BC\rangle$$
$$0 \le 6 + 2\langle B\rangle + 3\langle AB\rangle - 4\langle AC\rangle + \langle A\rangle\langle B\rangle + 2\langle A\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle - 2\langle B\rangle\langle AC\rangle - 2\langle A\rangle\langle BC\rangle$$
$$0 \le 6 - 2\langle A\rangle + 2\langle B\rangle - 3\langle AB\rangle - 5\langle AC\rangle + \langle A\rangle\langle B\rangle + \langle A\rangle\langle C\rangle + 2\langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle - 2\langle A\rangle\langle BC\rangle$$
$$0 \le 6 - 3\langle AB\rangle + \langle AC\rangle - 2\langle BC\rangle + \langle A\rangle\langle B\rangle + \langle A\rangle\langle C\rangle - 4\langle B\rangle\langle C\rangle + 2\langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle - 2\langle A\rangle\langle BC\rangle$$
$$0 \le 6 + \langle AB\rangle - 3\langle AC\rangle + 2\langle BC\rangle + \langle A\rangle\langle B\rangle + \langle A\rangle\langle C\rangle - 4\langle B\rangle\langle C\rangle - 2\langle A\rangle\langle B\rangle\langle C\rangle + \langle C\rangle\langle AB\rangle - \langle B\rangle\langle AC\rangle - 2\langle A\rangle\langle BC\rangle$$
$$0 \le 6 + 2\langle C\rangle + 3\langle AC\rangle - 5\langle BC\rangle - 2\langle A\rangle\langle B\rangle + \langle A\rangle\langle C\rangle + \langle B\rangle\langle C\rangle - 2\langle A\rangle\langle B\rangle\langle C\rangle + 2\langle C\rangle\langle AB\rangle + \langle B\rangle\langle AC\rangle + \langle A\rangle\langle BC\rangle$$
$$0 \le 6 - 3\langle AB\rangle - 2\langle AC\rangle - 2\langle BC\rangle - 2\langle ABC\rangle + \langle A\rangle\langle B\rangle + 4\langle B\rangle\langle C\rangle + \langle A\rangle\langle B\rangle\langle C\rangle - \langle C\rangle\langle AB\rangle - 2\langle B\rangle\langle AC\rangle$$

(#1): $0 \le 1 + \langle B\rangle \langle C\rangle + \langle AB\rangle + \langle AC\rangle$

(#2): $0 \le 2 + \langle A\rangle \langle B\rangle \langle C\rangle - \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle$

(#3): $0 \le 2 + \langle C\rangle \langle AB\rangle + \langle B\rangle \langle AC\rangle - \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#4): $0 \le 2 + \langle C\rangle \langle AB\rangle + \langle B\rangle \langle AC\rangle + \langle A\rangle \langle BC\rangle - \langle ABC\rangle$

(#5): $0 \le 4 - 2 \langle A\rangle - 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle - \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#6): $0 \le 4 - 2 \langle B\rangle - 2 \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle + \langle B\rangle \langle AC\rangle - \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#7): $0 \le 3 + \langle A\rangle \langle B\rangle + \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - 2 \langle AB\rangle - \langle C\rangle \langle AB\rangle - \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle$

(#8): $0 \le 3 + \langle B\rangle - \langle A\rangle \langle B\rangle + \langle A\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + \langle C\rangle \langle AB\rangle - \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle$

(#9): $0 \le 3 + \langle B\rangle + \langle A\rangle \langle B\rangle - \langle A\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + \langle C\rangle \langle AB\rangle + \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle$

(#10): $0 \le 3 + \langle A\rangle + \langle B\rangle + \langle C\rangle - \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + 3 \langle AB\rangle + \langle C\rangle \langle AB\rangle - \langle AC\rangle - \langle B\rangle \langle AC\rangle$

(#11): $0 \le 3 + \langle A\rangle + \langle B\rangle - \langle C\rangle + \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + 3 \langle AB\rangle - \langle C\rangle \langle AB\rangle + \langle AC\rangle - \langle B\rangle \langle AC\rangle$

(#12): $0 \le 4 - 2 \langle B\rangle + \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - \langle B\rangle \langle AC\rangle - 3 \langle BC\rangle + \langle ABC\rangle$

(#13): $0 \le 5 + 3 \langle A\rangle + \langle B\rangle + 2 \langle A\rangle \langle B\rangle + \langle C\rangle + 2 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle - 2 \langle C\rangle \langle AB\rangle + 3 \langle AC\rangle + \langle BC\rangle - \langle ABC\rangle$

(#14): $0 \le 5 + 3 \langle A\rangle + \langle B\rangle + 2 \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + 3 \langle AC\rangle + \langle BC\rangle - \langle ABC\rangle$

(#15): $0 \le 4 + 2 \langle A\rangle \langle B\rangle + \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - 2 \langle AB\rangle - \langle C\rangle \langle AB\rangle - \langle B\rangle \langle AC\rangle + 3 \langle BC\rangle + \langle ABC\rangle$

(#16): $0 \le 4 + 2 \langle A\rangle \langle B\rangle - 2 \langle C\rangle - 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#17): $0 \le 4 + 2 \langle A\rangle \langle B\rangle + 2 \langle C\rangle + 2 \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle + \langle A\rangle \langle BC\rangle - \langle ABC\rangle$

(#18): $0 \le 5 + \langle A\rangle + \langle B\rangle - 2 \langle A\rangle \langle B\rangle + \langle C\rangle + \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + 3 \langle AB\rangle + \langle C\rangle \langle AB\rangle + \langle AC\rangle - \langle B\rangle \langle AC\rangle - 4 \langle BC\rangle$

(#19): $0 \le 6 + \langle A\rangle \langle B\rangle + 4 \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - 3 \langle AB\rangle - \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 2 \langle ABC\rangle$

(#20): $0 \le 6 + 2 \langle B\rangle + \langle A\rangle \langle B\rangle + 2 \langle A\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + 3 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 4 \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - 2 \langle A\rangle \langle BC\rangle$

(#21): $0 \le 4 - 2 \langle A\rangle \langle B\rangle + 2 \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle + \langle BC\rangle + \langle ABC\rangle$

(#22): $0 \le 4 + 2 \langle A\rangle \langle B\rangle - 2 \langle C\rangle + \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle - 3 \langle BC\rangle + \langle ABC\rangle$

(#23): $0 \le 4 + 2 \langle A\rangle \langle B\rangle + 2 \langle A\rangle \langle C\rangle + 2 \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle + \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#24): $0 \le 4 + 2 \langle A\rangle \langle B\rangle + 2 \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle - \langle B\rangle \langle AC\rangle + \langle BC\rangle + \langle ABC\rangle$

(#25): $0 \le 5 + \langle A\rangle + \langle B\rangle + \langle C\rangle + \langle A\rangle \langle C\rangle - 2 \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + \langle B\rangle \langle AC\rangle - \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#26): $0 \le 6 + \langle A\rangle \langle B\rangle + 2 \langle A\rangle \langle C\rangle + 2 \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - 3 \langle AB\rangle - \langle C\rangle \langle AB\rangle - 4 \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - 2 \langle A\rangle \langle BC\rangle$

(#27): $0 \le 3 - \langle A\rangle + \langle B\rangle + \langle A\rangle \langle B\rangle + \langle C\rangle + \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + \langle C\rangle \langle AB\rangle + \langle B\rangle \langle AC\rangle + \langle A\rangle \langle BC\rangle - \langle ABC\rangle$

(#28): $0 \le 5 + \langle A\rangle + \langle B\rangle + 2 \langle A\rangle \langle B\rangle + \langle C\rangle + \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - \langle AB\rangle + \langle C\rangle \langle AB\rangle + \langle AC\rangle - \langle B\rangle \langle AC\rangle - 2 \langle A\rangle \langle BC\rangle - 2 \langle ABC\rangle$

(#29): $0 \le 5 + \langle A\rangle + \langle B\rangle + \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle C\rangle \langle AB\rangle + \langle AC\rangle + \langle B\rangle \langle AC\rangle - \langle BC\rangle + \langle A\rangle \langle BC\rangle - 2 \langle ABC\rangle$

(#30): $0 \le 6 + \langle A\rangle \langle B\rangle + \langle A\rangle \langle C\rangle - 4 \langle B\rangle \langle C\rangle + 2 \langle A\rangle \langle B\rangle \langle C\rangle - 3 \langle AB\rangle + \langle C\rangle \langle AB\rangle + \langle AC\rangle - \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle$

(#31): $0 \le 6 + \langle A\rangle \langle B\rangle + \langle A\rangle \langle C\rangle - 4 \langle B\rangle \langle C\rangle - 2 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + \langle C\rangle \langle AB\rangle - 3 \langle AC\rangle - \langle B\rangle \langle AC\rangle + 2 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle$

(#32): $0 \le 6 - 2 \langle A\rangle + 2 \langle B\rangle + \langle A\rangle \langle B\rangle + \langle A\rangle \langle C\rangle + 2 \langle A\rangle \langle B\rangle \langle C\rangle - 3 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 5 \langle AC\rangle - \langle B\rangle \langle AC\rangle - 2 \langle A\rangle \langle BC\rangle$

(#33): $0 \le 6 - 2 \langle A\rangle \langle B\rangle + 2 \langle C\rangle + \langle A\rangle \langle C\rangle + \langle B\rangle \langle C\rangle - 2 \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle C\rangle \langle AB\rangle + 3 \langle AC\rangle + \langle B\rangle \langle AC\rangle - 5 \langle BC\rangle + \langle A\rangle \langle BC\rangle$

(#34): $0 \le 3 + \langle A\rangle + \langle B\rangle + \langle A\rangle \langle B\rangle - \langle C\rangle + \langle A\rangle \langle C\rangle + \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle + \langle B\rangle \langle AC\rangle - \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#35): $0 \le 5 + \langle A\rangle + \langle B\rangle + 2 \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle A\rangle \langle C\rangle + \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + 3 \langle AB\rangle + \langle C\rangle \langle AB\rangle - \langle AC\rangle + \langle B\rangle \langle AC\rangle - 4 \langle BC\rangle$

(#36): $0 \le 8 + 2 \langle A\rangle - 2 \langle B\rangle - \langle A\rangle \langle B\rangle + 2 \langle B\rangle \langle C\rangle + 3 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle - 6 \langle AC\rangle + \langle B\rangle \langle AC\rangle - 3 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#37): $0 \le 8 - 2 \langle B\rangle + 2 \langle A\rangle \langle B\rangle - 2 \langle C\rangle - \langle B\rangle \langle C\rangle - 3 \langle A\rangle \langle B\rangle \langle C\rangle + 3 \langle C\rangle \langle AB\rangle - 6 \langle AC\rangle + \langle B\rangle \langle AC\rangle + \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#38): $0 \le 8 + 2 \langle A\rangle + \langle A\rangle \langle C\rangle + 2 \langle B\rangle \langle C\rangle + 3 \langle A\rangle \langle B\rangle \langle C\rangle + 6 \langle AB\rangle - \langle C\rangle \langle AB\rangle + \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 3 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#39): $0 \le 8 - 2 \langle A\rangle \langle B\rangle + 2 \langle C\rangle + \langle A\rangle \langle C\rangle - 3 \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + 3 \langle C\rangle \langle AB\rangle + \langle AC\rangle + 2 \langle B\rangle \langle AC\rangle - 6 \langle BC\rangle - \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#40): $0 \le 5 - \langle A\rangle + \langle B\rangle - 2 \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle A\rangle \langle C\rangle + 2 \langle B\rangle \langle C\rangle + \langle AB\rangle - 2 \langle C\rangle \langle AB\rangle + \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#41): $0 \le 5 - \langle A\rangle + \langle B\rangle + 2 \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle A\rangle \langle C\rangle - 2 \langle B\rangle \langle C\rangle + \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + \langle AC\rangle + 2 \langle B\rangle \langle AC\rangle - \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#42): $0 \le 7 + \langle A\rangle + \langle B\rangle - \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + 3 \langle C\rangle \langle AB\rangle - \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - 3 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle - 3 \langle ABC\rangle$

(#43): $0 \le 5 + \langle A\rangle + \langle B\rangle - \langle A\rangle \langle B\rangle + \langle C\rangle + 2 \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle - \langle C\rangle \langle AB\rangle + \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#44): $0 \le 5 + \langle A\rangle - \langle B\rangle + 2 \langle A\rangle \langle B\rangle + \langle C\rangle + \langle A\rangle \langle C\rangle + 2 \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + \langle B\rangle \langle AC\rangle + \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#45): $0 \le 5 + \langle A\rangle + \langle B\rangle - 2 \langle A\rangle \langle B\rangle + \langle C\rangle - \langle A\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle - \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle + \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#46): $0 \le 6 - 2 \langle A\rangle + 2 \langle A\rangle \langle B\rangle - \langle A\rangle \langle C\rangle - 4 \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + \langle AC\rangle + \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle$

(#47): $0 \le 6 - 4 \langle A\rangle \langle B\rangle - 2 \langle C\rangle + 2 \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + 2 \langle B\rangle \langle AC\rangle + \langle BC\rangle + \langle A\rangle \langle BC\rangle$

(#48): $0 \le 7 + \langle A\rangle + \langle B\rangle + 2 \langle A\rangle \langle B\rangle + \langle C\rangle + 4 \langle A\rangle \langle C\rangle - 2 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + \langle C\rangle \langle AB\rangle - \langle AC\rangle + \langle B\rangle \langle AC\rangle + 3 \langle BC\rangle - \langle A\rangle \langle BC\rangle - 2 \langle ABC\rangle$

(#49): $0 \le 5 + \langle A\rangle + \langle B\rangle + \langle C\rangle - \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle - 2 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle + \langle A\rangle \langle BC\rangle - \langle ABC\rangle$

(#50): $0 \le 5 + \langle A\rangle + \langle B\rangle + \langle A\rangle \langle B\rangle - \langle C\rangle - 2 \langle A\rangle \langle C\rangle - 2 \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle - \langle C\rangle \langle AB\rangle + \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle + \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#51): $0 \le 3 + \langle A\rangle + \langle B\rangle + \langle A\rangle \langle B\rangle + \langle C\rangle - \langle A\rangle \langle C\rangle + \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle + \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#52): $0 \le 5 + \langle A\rangle - \langle B\rangle - 2 \langle A\rangle \langle B\rangle + \langle C\rangle - \langle A\rangle \langle C\rangle + \langle B\rangle \langle C\rangle + \langle AB\rangle + \langle C\rangle \langle AB\rangle + 2 \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle - 2 \langle ABC\rangle$

(#53): $0 \le 5 + \langle A\rangle - \langle B\rangle + \langle A\rangle \langle B\rangle + \langle C\rangle - \langle A\rangle \langle C\rangle + 2 \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + 2 \langle B\rangle \langle AC\rangle + \langle BC\rangle + \langle A\rangle \langle BC\rangle - 2 \langle ABC\rangle$

(#54): $0 \le 5 + \langle A\rangle + \langle B\rangle - \langle A\rangle \langle B\rangle - \langle C\rangle + 2 \langle A\rangle \langle C\rangle + \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + \langle AC\rangle + \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + 2 \langle ABC\rangle$

(#55): $0 \le 5 - \langle A\rangle + \langle B\rangle + \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + \langle AC\rangle + \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + 2 \langle ABC\rangle$

(#56): $0 \le 7 + \langle A\rangle + \langle B\rangle - 2 \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle - 2 \langle C\rangle \langle AB\rangle - 3 \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - 3 \langle A\rangle \langle BC\rangle + 3 \langle ABC\rangle$

(#57): $0 \le 7 + \langle A\rangle + \langle B\rangle + \langle A\rangle \langle B\rangle - \langle C\rangle - 2 \langle A\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + 3 \langle C\rangle \langle AB\rangle + 3 \langle AC\rangle + 2 \langle B\rangle \langle AC\rangle + \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + 3 \langle ABC\rangle$

(#58): $0 \le 5 + \langle A\rangle - \langle B\rangle - 2 \langle A\rangle \langle B\rangle + \langle C\rangle - \langle A\rangle \langle C\rangle + 2 \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle + \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#59): $0 \le 8 + 4 \langle A\rangle \langle B\rangle + 2 \langle A\rangle \langle C\rangle - 2 \langle B\rangle \langle C\rangle + 2 \langle A\rangle \langle B\rangle \langle C\rangle - 4 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle - \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 3 \langle A\rangle \langle BC\rangle - 3 \langle ABC\rangle$

(#60): $0 \le 3 + \langle A\rangle + \langle B\rangle + \langle A\rangle \langle B\rangle + \langle C\rangle + \langle A\rangle \langle C\rangle + \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle + 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle + 2 \langle BC\rangle + \langle A\rangle \langle BC\rangle - \langle ABC\rangle$

(#61): $0 \le 5 - \langle A\rangle + \langle B\rangle + 2 \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + \langle AC\rangle + 2 \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#62): $0 \le 5 + \langle A\rangle + \langle B\rangle - 2 \langle A\rangle \langle B\rangle + \langle C\rangle - \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle - 2 \langle A\rangle \langle B\rangle \langle C\rangle - \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle + 2 \langle BC\rangle - \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#63): $0 \le 7 + \langle A\rangle + \langle B\rangle + \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle A\rangle \langle C\rangle + 2 \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + 3 \langle C\rangle \langle AB\rangle + \langle AC\rangle + 2 \langle B\rangle \langle AC\rangle - 3 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + 3 \langle ABC\rangle$

(#64): $0 \le 9 - \langle A\rangle + \langle B\rangle + 2 \langle A\rangle \langle B\rangle + \langle C\rangle + 2 \langle A\rangle \langle C\rangle - 2 \langle B\rangle \langle C\rangle - 2 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + 3 \langle C\rangle \langle AB\rangle + \langle AC\rangle + 3 \langle B\rangle \langle AC\rangle - \langle BC\rangle + 3 \langle A\rangle \langle BC\rangle - 4 \langle ABC\rangle$