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, Germany*

(Dated: April 28, 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 a probability distribution over some random variables, the problem of **causal inference** is to determine a plausible set of causal relations between these variables that could have generated the given distribution. This type of problem comes up in a wide variety of statistics applications, from sussing out biological pathways to enabling machine learning pearl2009causality,spirtes2011causation,studeny2005probabilistic,koller2009probabilistic [? ? ? ?]. A related problem is to start with a given set of causal relations, and then to determine the set of all probability distribution that these relations can generate on random variables that live on the nodes of this causal structure. Taking both of these starting points together, one can also start with *both* a given distribution and a candidate causal structure and ask whether the two are compatible, in the sense that the causal structure could in principle have generated the given distribution.

In the simplest setting, a causal structure hypothesis consists of a directed acyclic graph (DAG) which only contains the observed variables as nodes. In this case, obtaining a verdict on the compatibility of a given distribution with the causal structure is simple: the compatibility holds if and only if the distribution is Markov with respect to the DAG. There exist techniques to determine the compatible causal structures from the distribution only, i.e. without having an *a priori* hypothesis pearl2009causality [?].

A significantly more difficult case is when one considers a causal structure hypothesis which consists of a DAG with **latent** or **hidden** nodes, so that the set of observable variables is only a subset of the nodes of the DAG. This case occurs e.g. in situations where one needs to deal with the possible presence of unobserved **confounders**, as in experiment design. While the conditional independence relations implied by the causal structure on the observable nodes are still necessary conditions for compatibility of a distribution with the causal structure hypothesis, these equations are generally no longer sufficient. Finding necessary and sufficient conditions is a difficult open problem. Determining plausible candidate causal structures from the given distribution is even more difficult, but this can be reduced to the previous problem e.g. by enumerating all DAGs (say with a fixed number of latent nodes) and checking one at a time for compatibility with the distribution. Due to the possibility of making this reduction, and because the compatibility problem is already difficult enough, we will focus on the case when a DAG (with latent nodes) is given as a description of the causal structure. We then consider both the problem of deciding whether a given distribution can arise from this causal structure, as well as that of describing the entire set of these distributions for a fixed number of values per variable.

Historically, the insufficiency of the causal independence relations for causal inference in the presence of latent variables has first been observed by Bell in the context of the hidden variable problem of quantum physics bell1964einstein [?]. Roughly speaking, Bell derived an inequality that any distribution consistent with the spacetime causal structure must satisfy, and found this inequality to be violated by distributions generated from suitably entangled quantum states[1]. Subsequently, Pearl has derived another inequality, the **instrumental inequality** pearl1995instrumental [?], which witnesses the impossibility to model certain three-variable distribution by the *instrumental scenario* causal structure hypothesis. More recently, steudel2010ancestors [?] have derived inequalities which formalize an extended version of Reichenbach's principle by witnessing the impossibility to model a given distribution on n variables in terms of a causal structure without a common (hidden) ancestor for any subset of at least c nodes, for any $n, c \in \mathbb{N}$. Motivated by Bell's inequality, many other Bell-type inequalities have been used in quantum physics Brunner2013Bell [?], but the foundational role of causal structure has only recently been appreciated WoodSpekkens,fritz2012bell,pusey2014gdag,BeyondBellII [? ? ? ?].

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. As far as we know, our method is the first systematic tool for doing causal inference with latent variables that goes beyond observable conditional independence relations and does not assume any bounds on the number of values of each latent variable. While our method can be used to systematically generate necessary conditions for compatibility with a given causal structure hypothesis, we do not know whether these inequalities are also sufficient, and we currently have conflicting evidence on this question. On the one hand, our method rederives all Bell-type inequalities (Sec. VI), but on the other hand we have not yet been able to obtain Pearl's instrumental inequality from our method.

Our 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.

[1] This inconsistency has subsequently become known as *quantum nonlocality* Brunner2013Bell [?]. Although the term suggests the existence of nonlocal interactions, in the sense that the actual causal structure may be different from the hypothesized one, this interpretation is at odds with the fact that no nonlocal interactions have been observed in nature, implying that their presence would require fine-tuning WoodSpekkens [?]. A less problematic alternative conclusion from Bell's theorem is the impossibility to model quantum physics in terms of the usual notions of "classical" probability theory.

II. CAUSAL MODELS AND CAUSAL INFERENCE

finitions

A **causal model** consists of a pair of objects: a **causal structure** and a set of **causal parameters**. We define each in turn.

The causal structure specifies a directed acyclic graph (DAG). Recall that a DAG G consists of a set of nodes and directed edges (i.e., ordered pairs of nodes), which we denote by $\mathsf{Nodes}[G]$ and $\mathsf{Edges}[G]$ respectively. Each node corresponds to a random variable and a directed edge between two nodes corresponds to there being a direct causal influence from one variable to the other. 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. $\mathsf{Pa}_G(X) = \{ Y \mid Y \to X \}$. 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. $\mathsf{Ch}_G(X) = \{ Y \mid X \to Y \}$. If $\boldsymbol{U}$ is a set of nodes, then we put $\mathsf{Pa}_G(\boldsymbol{U}) := \bigcup_{X \in \boldsymbol{U}} \mathsf{Pa}_G(X)$ and $\mathsf{Ch}_G(\boldsymbol{U}) := \bigcup_{X \in \boldsymbol{U}} \mathsf{Ch}_G(X)$. The **ancestors** of a set of nodes $\boldsymbol{U}$, denoted $\mathsf{An}_G(\boldsymbol{U})$, are defined as those nodes which have a directed *path* to some node in $\boldsymbol{U}$, including the nodes in $\boldsymbol{U}$ themselves. Equivalently (dropping the G subscript), $\mathsf{An}(\boldsymbol{U}) := \bigcup_{n \in \mathbb{N}} \mathsf{Pa}^n(\boldsymbol{U})$, where $\mathsf{Pa}^n(\boldsymbol{U})$ is inductively defined via $\mathsf{Pa}^0(\boldsymbol{U}) := \boldsymbol{U}$ and $\mathsf{Pa}^{n+1}(\boldsymbol{U}) := \mathsf{Pa}(\mathsf{Pa}^n(\boldsymbol{U}))$. The causal structure also incorporates a distinction between the nodes of the DAG, namely, between those that are observed, denoted $\mathsf{ObservedNodes}[G]$ and those that are latent, denoted $\mathsf{LatentNodes}[G]$.

The set of causal parameters specify, for each node, the conditional probability distribution over the values of the random variable associated to that node, given the values of the variables associated to the node's parents. (In the case of root nodes, the parents are the null set and the conditional probability distribution is simply a probability distribution.) We will denote a conditional probability distribution over a variable Y given a variable X by $P_{Y|X}$, while the particular conditional probability of the variable X taking the value x given that the variable Y takes the values y is denoted $P_{Y|X}(y|x)$. Therefore, a given set of causal parameters, denoted F_G, has the form

$$F_G \equiv \{ P_{A|\mathsf{Pa}_G(A)} : A \in \mathsf{Nodes}[G] \}. \tag{1}$$

A DAG G and a set of causal parameters F_G defines a causal model, which we denote by $M = (G, F_G)$. In this case, $\mathsf{DAG}[M]$ and $\mathsf{Params}[M]$ denote G and F_G respectively.

As is standard, a causal model M specifies a joint distribution over all variables in the DAG, namely,

$$P(\mathsf{Nodes}[G]) = \Pi_{A \in \mathsf{Nodes}[G]} P_{A|\mathsf{Pa}_G(A)}, \tag{2}$$

(typically called the Markov condition), and the joint distribution over the observed variables is obtained from the joint distribution over all variables by marginalization over the latent variables

$$P(\mathsf{ObservedNodes}[G]) = \sum_{\{X : X \in \mathsf{LatentNodes}[G]\}} P(\mathsf{Nodes}[G]). \tag{3}$$

In addition to the notion of a causal model, it is useful to have a notion of a **causal hypothesis**, which we define to be a set of causal models. We shall consider only causal hypotheses for which the models in the set share the same DAG, and we denote such a causal hypothesis by H_G where G is the associated DAG.

The most simple sort of causal hypothesis is one that includes the full set of causal models associated to a DAG G, which we denote by H_G^{full}. Our results will primarily concern this sort of hypothesis.

With these notions in hand, we can finally specify the sort of causal inference problem for which one can usefully apply our technique: a decision problem takes two inputs, namely, a specification of observational data and a specification of a causal hypothesis, and outputs whether or not they are consistent with one another.

The most informative sort of observational data is a specification of the joint distributions over all of the observed variables. However, one can imagine that the observational data specifies merely an interval in the space of possible joint distributions. And there are other ways in which the observational data may be less informative than a specification of the joint distribution. For instance, one might merely have a description of the conditional independence relations that hold among the observed variables. Or one might merely have a specification of the marginals of the joint distribution for certain subsets of the observed variables[2]

On the side of the causal hypothesis, it is most common to consider the hypothesis of a particular DAG, with no restriction on the set of parameters that supplement it. One can also, however, consider causal hypotheses that are

---

[2] In this last case, the causal inference problem corresponds to a version of the marginals problem, but where the joint distribution is constrained by the causal hypothesis. Such a causally-constrained version of the marginals problem will be central to our inflation DAG technique.

more fine-grained, for instance, by incorporating constraints on how certain nodes in the DAG functionally depend on their parents. An example is the assumption of an additive noise model for certain causal influences: if an observed variable Y has an observed variable X as a parent and also a latent variable U, then the noise is deemed additive if $Y = \alpha X + \beta U$ for some scalars α and β [provide references]. Clearly, the conditional probability distribution $P_{Y|XU}$ that can be achieved in such an additive noise model are a subset of valid conditional distributions. More general constraints on the conditional distributions have also been explored [provide references]. Our inflation DAG technique will require consideration of a novel and unusual sort of causal hypothesis, wherein the conditional probability distributions associated to distinct nodes are sometimes required to be equivalent.

In broad strokes, the inflation DAG technique is a way of mapping a given causal inference problem onto a new causal inference problem where the observational and causal inputs of the latter problem are determined by the observational and causal inputs of the original problem.

III. INFLATION: A TOOL FOR CAUSAL INFERENCE

We now introduce the notion of **an inflation of a causal model**. If the original causal model is associated to a DAG G, then a nontrivial inflation of this model is associated to a different DAG, G'. We refer to G' as an inflation of G. There are many possible choices of G' for a given G (specified below), hence many possible inflations of a given DAG. We denote the set of these by $\mathsf{Inflations}[G]$. The choice of an element $G' \in \mathsf{Inflations}[G]$ is the only freedom in the inflation of a causal model. Once a choice is made, the set of parameters of the inflation model is fixed uniquely by the the set of parameters of the original model. We denote the function on models associated with such an inflation by $\mathsf{Inflation}[G \to G']M$ (specified below).

We begin by defining the condition under which a DAG G' is an inflation of a DAG G. This requires some preliminary definitions.

The **subgraph** of G induced by restricting attention to the set of nodes $\boldsymbol{V} \subseteq \mathsf{Nodes}[G]$ will be denoted $\mathsf{SubDAG}_G[\boldsymbol{V}]$. 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** of $\boldsymbol{V}$, denoted $\mathsf{AnSubDAG}_G[\boldsymbol{V}]$, which is the minimal subgraph containing the full ancestry of $\boldsymbol{V}$, $\mathsf{AnSubDAG}_G[\boldsymbol{V}] \coloneqq \mathsf{SubDAG}_G[\mathsf{An}_G(\boldsymbol{V})]$.

Inflation involves a sort of copying operation on nodes of the DAG. Specifically, every node of G' can be understood as a copy of a node of G. If A denotes a node in G that has copies in G', then we denote these copies by $A_1, \ldots, A_k$, and the variable that indexes the copies is termed the **copy-index**. When two objects (e.g. nodes, sets of nodes, DAGs, etc...) are the same up to copy-indices, then we use $\sim$ to indicate this. For instance, we have $A_i \sim A_j \sim A$. This copying operation must also preserve the causal structure of the DAG, in a manner that is formalized by the following definition.

Definition 1. *The DAG G' is said to be an inflation of the DAG G if and only if for every node A_i in G', the ancestral subgraph of A_i in G' is equivalent, under removal of the copy-index, to the ancestral subgraph of A in G,*

$$G' \in \mathsf{Inflations}[G] \quad \textit{iff} \quad \forall A_i \in \mathsf{Nodes}[G'] : \ \mathsf{AnSubDAG}_{G'}[A_i] \sim \mathsf{AnSubDAG}_G[A]. \tag{4}$$ {eq:defin

To illustrate the notion of the inflation of a DAG, we consider the DAG of Fig. 1, which is called the *Triangle scenario* (for obvious reasons) and which has been studied by many authors [? (Fig. E#8), ? (Fig. 18b), ? (Fig. 3), ? (Fig. 6a), ? (Fig. 1a), ? (Fig. 8), ? (Fig. 1b), ? (Fig. 4b)] Different inflations of the Triangle scenario are depicted in Figs. 2 to 6.

We now turn to specifying the function $\mathsf{Inflation}[\,]_{G\to G'}$, that is to specifying how the set of parameters of a causal model transform under inflation. One we've said how DAG inflated, then we need TWO things: How models (and hypothesis) inflate, and also how observational data inflates. Best to word intro to subject in a way that suggests we'll be getting around to both topics sequentially? PHYSICS needed here. Why mathematical abstraction? Explain that due to physical construct-ability, therefore we expect that models on the original DAG can be used to build a model on the inflation DAG...

Definition 2. *Consider causal models M and M' where $\mathsf{DAG}[M] = G$ and $\mathsf{DAG}[M'] = G'$ and such that G' is an inflation of G. The causal model M' is the $G \to G'$ inflation of M, that is, $M' = \mathsf{Inflation}[M]$, if and only if for every node A_i in G', the manner in which A_i depends causally on its parents within G' must be the same as the manner in which A depends causally on its parents within G. Noting that $A_i \sim A$ and that $\mathsf{Pa}_{G'}(A_i) \sim \mathsf{Pa}_G(A)$ (given Eq. (4)), one can formalize this condition as:*

$$\forall A_i \in \mathsf{Nodes}[G'] : \ P_{A_i|\mathsf{Pa}_{G'}(A_i)} = P_{A|\mathsf{Pa}_G(A)}, \tag{5}$$ {eq:funcd

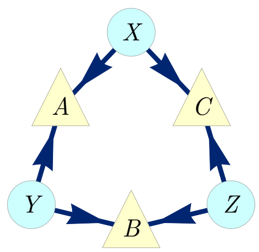

FIG. 1. The causal structure of the Triangle scenario.

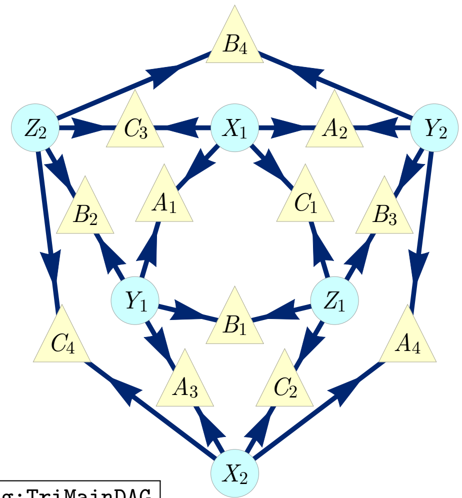

fig:TriMainDAG

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

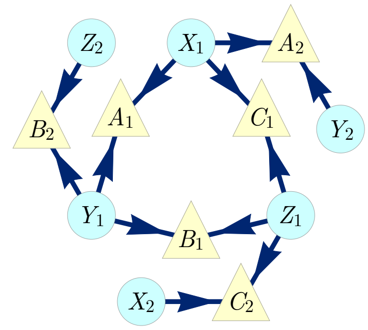

FIG. 3. Another inflation of the Triangle scenario consisting, also notably $\mathsf{AnSubDAG}_{(\text{Fig. 2})}[A_1A_2B_1B_2C_1C_2]$. fig:Tri22

To sum up then, inflation is a mapping to a new causal model wherein each given variable in the original DAG may have counterparts in the inflation DAG and where the causal dependencies in the inflation DAG are given by the corresponding causal dependencies in the original DAG. Note that the operation of modifying a DAG and equipping the modified version with causal dependencies that mirror those of the original also appears in the *do calculus* of [?]. pearl2009causalit FiXme Note: non-Shannon

Because causal hypotheses are sets of causal models, and the inflation map acts on causal models, it has an induced action on causal hypotheses. Note that $\mathsf{Inflation}[H_G^{\mathsf{full}}] \neq H_{G'}^{\mathsf{full}}$ because even if there is no restriction on the set of parameters supplementing G, inflation imposes a restriction on the set of parameter supplementing G'. Specifically, if nodes A_i and A_j on G' are copies of a single node A on G,then the set of parameters on G' are constrained to satisfy $P_{A_i|\mathsf{Pa}_{G'}(A_i)} = P_{A_j|\mathsf{Pa}_{G'}(A_j)}$.

To see how inflation is relevant for causal inference, we must explain how the distributions that can be achieved in the inflation model are constrained by those in the original model. In what follows, we assume that $G' \in \mathsf{Inflations}[G]$ and that $C' = \mathsf{Inflation}[C]$.

Note, first of all, that for any sets of nodes $\boldsymbol{U} \in \mathsf{Nodes}[G']$ and $\tilde{\boldsymbol{U}} \in \mathsf{Nodes}[G]$,

$$\text{if} \quad \mathsf{AnSubDAG}_{G'}[\boldsymbol{U}] \sim \mathsf{AnSubDAG}_{G}[\tilde{\boldsymbol{U}}] \quad \text{then} \quad P_{C'}(\boldsymbol{U}) = P_C(\tilde{\boldsymbol{U}}). \tag{6}$$ {eq:coinc

This follows from the fact that the probability distributions over $\boldsymbol{U}$ and $\tilde{\boldsymbol{U}}$ depend only on their ancestral subgraphs and the parameters defined thereon, which by the definition of inflation are the same for $\boldsymbol{U}$ and for $\tilde{\boldsymbol{U}}$.

It is useful to have a name for a set of nodes in the inflation DAG, $\boldsymbol{U} \subseteq \mathsf{Nodes}[G']$, such that one can find a

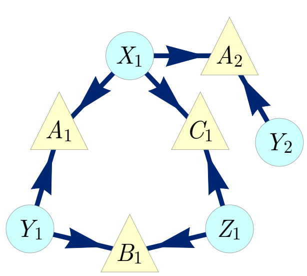

FIG. 4. A simple inflation of the Triangle scenario, also notably $\mathsf{AnSubDAG}_{(Fig.\ 3)}[A_1A_2B_1C_1]$.

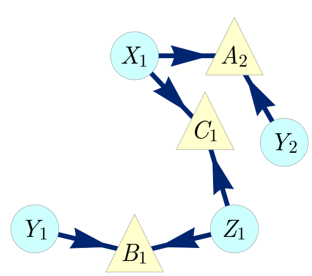

FIG. 5. An even simpler inflation of the Triangle scenario, also notably $\mathsf{AnSubDAG}_{(Fig.\ 4)}[A_2B_1C_1]$. fig:simpleinflation

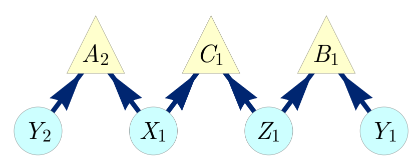

FIG. 6. Another representation of Fig. 5. Despite not containing the original scenario, this is a valid inflation per Eq. (4). fig:simplestinflation fig:TriDa

corresponding set in the original DAG, $\tilde{\boldsymbol{U}} \subseteq \mathsf{Nodes}[G]$, which has an equivalent ancestral subgraph. We say that such subsets of the nodes of G' are injectable into G and we call them the **injectable sets**,

$$\boldsymbol{U} \in \mathsf{InjectableSets}[G'] \quad \text{iff} \quad \exists \tilde{\boldsymbol{U}} \subseteq \mathsf{Nodes}[G] : \mathsf{AnSubDAG}_{G'}[\boldsymbol{U}] \sim \mathsf{AnSubDAG}_{G}[\tilde{\boldsymbol{U}}]. \tag{7}$$

{eq:defin

In Fig. 3, for example, the set $\{A_1B_1C_1\}$ is injectable because its ancestral subgraph is equivalent up to copy-indices to the ancestral subgraph of $\{ABC\}$ in the original DAG (which is just the full DAG), and the set $\{A_2C_1\}$ is injectable because its ancestral subgraph is equivalent to that of $\{AC\}$ in the original DAG.

Note that it is clear that a set of nodes in the inflation DAG can only be injectable if it contains at most one copy of any node from the original DAG. Similarly, it can only be injectable if its ancestral subgraph also contains at most one copy of any node from the original DAG. Thus, in Fig. 3, $\{A_1A_2C_1\}$ is not injectable because it contains two copies of A, and $\{A_2B_1C_1\}$ is not injectable because its ancestral subgraph contains two copies of Y.

The sets of nodes of G' that will be of primary interest to us are not the injectable sets per se, but sets satisfying a slightly weaker constraint. To define the latter sorts of sets, which we term *pre-injectable*, we must first introduce some additional terminology.

We refer to a pair of nodes which do not share any common ancestor as being **ancestrally independent**, for which we invent the notation $X \Lsh\!\emptyset\!\Rsh Y$. Generalizing to sets, $\boldsymbol{U} \Lsh\!\emptyset\!\Rsh \boldsymbol{V}$ indicates that no node in $\boldsymbol{U}$ shares a common ancestor with any node in $\boldsymbol{V}$,

$$\boldsymbol{U} \Lsh\!\emptyset\!\Rsh \boldsymbol{V} \quad \text{iff} \quad \mathsf{An}(\boldsymbol{U}) \cap \mathsf{An}(\boldsymbol{V}) = \emptyset. \tag{8}$$

Ancestral independence is equivalent to d-separation by the empty set [? ? ? ?].

pearl2009causality,spirtes2011causation,studeny2005probabil

A set of nodes $\boldsymbol{U}$ in the inflation DAG G' will be called **pre-injectable** whenever it is a union of injectable sets with disjoint ancestries.

$$\boldsymbol{U} \in \mathsf{PreInjectableSets}[G'] \quad \text{iff} \quad \exists \{\boldsymbol{U}_i \in \mathsf{InjectableSets}[G']\} \quad \text{s.t.} \quad \boldsymbol{U} = \bigcup_i \boldsymbol{U}_i \quad \text{and} \quad \forall i \neq j : \boldsymbol{U}_i \Lsh\!\emptyset\!\Rsh \boldsymbol{U}_j. \tag{9}$$

{eq:defpr

Note that every injectable set is a trivial example of a pre-injectable set.

Because ancestral independence in the DAG implies statistical independence for any probability distribution consistent with the DAG, it follows that if $\boldsymbol{U}$ is a pre-injectable set and $\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_n$ are the ancestrally independent components of $\boldsymbol{U}$, then

$$P_{C'}(\boldsymbol{U}) = P_{C'}(\boldsymbol{U}_1) P_{C'}(\boldsymbol{U}_2) \cdots P_{C'}(\boldsymbol{U}_n). \tag{10}$$

Furthermore, because each injectable set of variables $\boldsymbol{U}_i$ satisfies Eq. (6), it follows that joint distributions on pre-injectable sets in the inflation model can be expressed in terms of distributions defined on the original causal model,

eq:coincidingdistrodef

$$P_{C'}(\boldsymbol{U}) = P_{C}(\tilde{\boldsymbol{U}}_1) P_{C}(\tilde{\boldsymbol{U}}_2) \cdots P_{C}(\tilde{\boldsymbol{U}}_n). \tag{11}$$

{eq:prein

This latter property implies that the observable probability distribution over any pre-injectable set in the inflation model is fully specified by the ~~parameters~~ observable probability distribution ~~in the original causal model~~ per the original observable data. Indeed, this relation is what motivates us to consider the pre-injectable sets.

Consider a causal inference problem where the observational input is a probability distribution over a set $\boldsymbol{O}$ of observed variables, denoted $P(\boldsymbol{O})$ and the causal hypothesis is $H_{G,S}$ where $\mathsf{ObservedNodes}[G] = \boldsymbol{O}$. Suppose that one seeks only to determine whether the causal hypothesis is consistent with the observational input. This holds if and only if

$$\exists C \in H_{G,S} : P_C(\mathsf{ObservedNodes}[G]) = P(\boldsymbol{O}). \tag{12}$$

Every inflation $G \to G'$ defines a new causal inference problem for which the decision regarding consistency is the same as for the original problem.

The causal hypothesis of the new causal inference problem is simply the image under a $G \to G'$ inflation of the causal hypothesis of the original causal inference problem, and one seeks to evaluate its consistency with observational data that is determined by the observational data of the original in the following way.

Let $\boldsymbol{O}$ denote the image of $\tilde{\boldsymbol{O}}$ under the $G \to G'$ inflation, and consider the subsets of $\boldsymbol{O}$ that are pre-injectable relative to the $G \to G'$ inflation, denoted $\mathsf{PreInjectableSets}[\boldsymbol{O}]$. Recall that for every set of nodes $\boldsymbol{U} \in \mathsf{PreInjectableSets}[\boldsymbol{O}]$ there is a partition $\boldsymbol{U} = \bigcup_i \boldsymbol{U}_i$ where the $\boldsymbol{U}_i$ are subsets of $\boldsymbol{O}$ that are injectable relative to the $G \to G'$ inflation and that are ancestrally independent. Recall also that, by the definition of a $G \to G'$ inflation, for any set of nodes

$\boldsymbol{U}_i \subseteq \boldsymbol{O}$ that is injectable, one can find a set of nodes $\tilde{\boldsymbol{U}}_i \subseteq \tilde{\boldsymbol{O}}$ such that $\forall i : \boldsymbol{U}_i \sim \tilde{\boldsymbol{U}}_i$. The observational data of the new causal inference problem is that $\forall \boldsymbol{U} \in \mathsf{PreInjectableSets}[\boldsymbol{O}] : P(\boldsymbol{U}) = P(\tilde{\boldsymbol{U}}_1)P(\tilde{\boldsymbol{U}}_2)\dots P(\tilde{\boldsymbol{U}}_n)$, where $P(\tilde{\boldsymbol{U}}_i)$ is the marginal of $P(\tilde{\boldsymbol{O}})$ on $\tilde{\boldsymbol{U}}_i$.

Therefore the observational input to the new causal inference problem is not the distribution on $\boldsymbol{O}$, but rather a specification of the marginals of this distribution on each of the pre-injectable sets of $\boldsymbol{O}$ under the $G \to G'$ inflation.

Summarizing, we have:

Lemma 3. *The causal hypothesis $H_{G,S}$ is consistent with the observational data $P(\tilde{\boldsymbol{O}})$ if and only if the image of this causal hypothesis under a $G \to G'$ inflation, denoted $H_{G',S'}$, is consistent with the observational data that* mainlemma $\forall \boldsymbol{U} \in \mathsf{PreInjectableSets}[\boldsymbol{O}] : P(\boldsymbol{U}) = P(\tilde{\boldsymbol{U}}_1)P(\tilde{\boldsymbol{U}}_2)\dots P(\tilde{\boldsymbol{U}}_n)$.

It follows that any consistency conditions that one can derive for the new causal inference problem immediately yield consistency conditions for the original causal inference problem. Indeed, any standard tool of causal inference can be applied to the new problem and conditions derived therefrom provide novel conditions for the original problem. Any causal inference tool, therefore, can potentially have its power augmented by combining it with inflation.

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

Consider the following causal inference problem. The observational data is a joint distribution over three binary-outcome variables, P^{obs}_{ABC}, where the marginal on each variable is uniform and the three are perfectly correlated,

$$P^{\text{obs}}_{ABC} = \frac{[000]+[111]}{2}, \quad \text{i.e.} \quad P^{\text{obs}}_{ABC}(abc) = \begin{cases} \frac{1}{2} & \text{if } a{=}b{=}c, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$ {eq:ghzdi

The causal hypothesis is the set of causal models associated to the DAG of the triangle scenario, depicted in Fig. 3. The problem is to decide whether the observational data is consistent with this causal hypothesis.

To solve this problem, we consider an inflated causal inference problem, namely the problem induced by the particular inflation depicted in Fig. fig:TriDagSubA2B1C1 6.

The sets $\{A_2C_1\}$ and $\{B_1C_1\}$ are injectable (hence trivially pre-injectable). The set $\{A_2B_1\}$ is also pre-injectable because the singleton sets $\{A_2\}$ and $\{B_1\}$ are each injectable and $A_2 \,{\Lsh}\!\emptyset\!{\Rsh}\, B_1$. Therefore, the inflated observational data for our new causal inference problem includes

$$\begin{aligned} P^{\text{obs-inf}}_{A_2C_1} &= P^{\text{obs}}_{AC} = \frac{[00]+[11]}{2} \\ P^{\text{obs-inf}}_{B_1C_1} &= P^{\text{obs}}_{BC} = \frac{[00]+[11]}{2} \\ P^{\text{obs-inf}}_{A_2B_1} &= P^{\text{obs}}_{A} \otimes P^{\text{obs}}_{B} = \frac{[0]+[1]}{2} \otimes \frac{[0]+[1]}{2}. \end{aligned} \tag{14}$$ [eqs]{eq:

The question is whether such constraints are consistent with the causal hypothesis that the DAG is that of Fig. 6, where the parameters supplementing the DAG are unrestricted but for that the latent variables Y_1 and Y_2 are identically distributed.

It is not difficult to see that the answer to the question is negative. And this verdict can be rendered without even appealing to the form of the inflation DAG (Note, however, that the inflation DAG has played a role in defining the new causal inference problem insofar as it has dictated the form that the inflated observational data should take). It suffices to make use of results concerning the marginals problem.

There is no joint distribution $P_{A_2B_1C_1}$ having such marginal as given in Eqs. (14). The only joint distribution that exhibits perfect correlation between A_2 and C_1 and between B_1 and C_1 also exhibits perfect correlation between A_2 and B_1.

We have therefore certified that perfect correlations among A, B and C is not consistent with the Triangle causal structure, recovering the seminal result of steudel2010ancestors [?].

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

Consider another causal inference problem concerning the triangle scenario, namely, that of determining whether the triangle DAG can explain a joint distribution P^{W}_{ABC} of the form

$$P^{\text{W}}_{ABC} = \frac{[100]+[010]+[001]}{3}, \quad \text{i.e.} \quad P^{\text{obs}}_{ABC}(abc) = \begin{cases} \frac{1}{3} & \text{if } a{+}b{+}c{=}1, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$ {eq:wdist

We call this the W-type distribution[3]. To our knowledge, the fact that the Triangle causal structure is inconsistent with the W-type distribution has not been demonstrated previously.

To prove the inconsistency of P^{W} with the Triangle causal structure, we consider the inflation DAG of Fig. 3. The sets $\{A_2C_1\}$, $\{B_2A_1\}$, $\{A_2B_1\}$ and $\{A_1B_1C_1\}$ are injectable. The set $\{A_2B_2C_2\}$ is pre-injectable because the singleton sets $\{A_2\}$, $\{B_2\}$ and $\{C_2\}$ are injectable and *all* ancestrally independent. It follows that we must consider the inflated observational data

$$P^{\text{W-inf}}_{A_2C_1} = P^{\mathrm{W}}_{AC} = \frac{[10]+[01]+[00]}{3} \tag{16} \quad \boxed{\texttt{\{W1\}}}$$

$$P^{\text{W-inf}}_{B_2A_1} = P^{\mathrm{W}}_{BA} = \frac{[10]+[01]+[00]}{3} \tag{17} \quad \boxed{\texttt{\{W2\}}}$$

$$P^{\text{W-inf}}_{C_2B_1} = P^{\mathrm{W}}_{CB} = \frac{[10]+[01]+[00]}{3} \tag{18} \quad \boxed{\texttt{\{W3\}}}$$

$$P^{\text{W-inf}}_{A_1B_1C_1} = P^{\mathrm{W}}_{ABC} = \frac{[100]+[010]+[001]}{3} \tag{19} \quad \boxed{\texttt{\{W4\}}}$$

$$P^{\text{W-inf}}_{A_2B_2C_2} = P^{\mathrm{W}}_{A} \otimes P^{\mathrm{W}}_{B} \otimes P^{\mathrm{W}}_{C} = \frac{[1]+2\cdot[0]}{3} \otimes \frac{[1]+2\cdot[0]}{3} \otimes \frac{[1]+2\cdot[0]}{3}\,. \tag{20} \quad \boxed{\texttt{\{W5\}}}$$

Eq. (16) implies that $C_1{=}0$ whenever $A_2{=}1$. Similarly, $A_1{=}0$ whenever $B_2{=}1$ and $B_1{=}0$ whenever $C_2{=}1$. Eq. (20) implies that A_2,B_2, and C_2 are uncorrelated and uniformly distributed, which means that sometimes they all take the value 1. Summarizing, we have

$$\begin{aligned}
& A_2{=}1 \implies C_1{=}0 \\
& B_2{=}1 \implies A_1{=}0 \\
& C_2{=}1 \implies B_1{=}0 \\
\text{and sometimes}\quad & A_2{=}1, B_2{=}1, C_2{=}1.
\end{aligned}$$

But these constraints clearly imply that

$$\text{sometimes}\quad A_1{=}0, B_1{=}0, C_1{=}0,$$

which contradicts Eq. (19).

Again, we have reached our verdict here simply by noting that the distributions defined in Eqs. (16)-(20) cannot arise as the marginals of a single distribution on $A_1, B_1, C_1, A_2, B_2, C_2$. Specifically, we have leveraged an approach to the marginal problem inspired by Hardy's version of Bell's theorem [? ?], see Sec. VIII for further discussion of Hardy-type paradoxes and their applications.

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 [? ? ?].
3. Moreover, *no* entropic inequality can witness the W-type distribution as unrealizable. [?] 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 [?], which gives tighter constraints than entropic inequalities for the Triangle scenario, also cannot detect inconsistency with the W-type distribution.

But the inflation technique 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.

We consider the distribution $P^{\mathrm{obs}}_{ABXY} = P^{\mathrm{PR}}_{AB|XY} \otimes P^{\mathrm{free}}_{X} \otimes P^{\mathrm{free}}_{Y}$, where P^{free}_{X} and P^{free}_{Y} are arbitrary full-support

[3] Because its correlations are reminiscent of those one obtains for the quantum state appearing in Ref. [?], and which is called the W state.

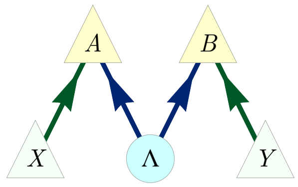

FIG. 7. The causal structure of the a bipartite Bell scenario. The local outcomes of Alice's and Bob's experimental probing is assumed to be a function of some latent common cause, in addition to their independent local experimental settings. fig:NewBellDAG1

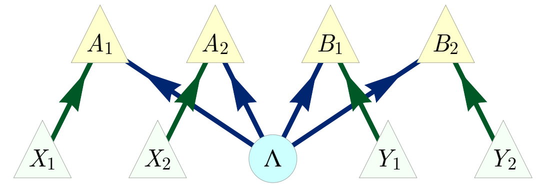

FIG. 8. An inflation DAG of the bipartite Bell scenario, where both local settings variables have been duplicated. fig:BellD

distributions[4] over the binary variables X and Y, and

$$P^{\text{PR}}_{AB|XY}(ab|xy) = \begin{cases} \frac{1}{2} & \text{if } \mathsf{mod}_2[a{+}b]{=}x{\cdot}y, \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$ {eq:PRbox

The correlations described by this distribution are known as PR-box correlations after Popescu and Rohrlich, and they are well-known to be inconsistent with the the Bell scenario [? ?] (PROriginal,PROnit). Here we prove this inconsistency using the inflation technique.

We use the inflation of the Bell DAG shown in Fig. 8.

We begin by recognizing that $\{A_1B_1X_1Y_1\}$, $\{A_2B_1X_2Y_1\}$, $\{A_1B_2X_1Y_2\}$, and $\{A_2B_2X_2Y_2\}$ are all injectable sets, and that $\{X_1X_2Y_1Y_2\}$ is a pre-injectable set because the singleton sets $\{X_1\}$, $\{X_2\}$, $\{Y_1\}$, and $\{Y_2\}$ are all injectable and ancestrally independent of one another. It follows that the observable data inflates according to

$$\begin{aligned} P^{\text{obs-inf}}_{A_1B_1X_1Y_1} &= P^{\text{obs}}_{ABXY} \\ P^{\text{obs-inf}}_{A_2B_1X_2Y_1} &= P^{\text{obs}}_{ABXY} \\ P^{\text{obs-inf}}_{A_1B_2X_1Y_2} &= P^{\text{obs}}_{ABXY} \\ P^{\text{obs-inf}}_{A_2B_2X_2Y_2} &= P^{\text{obs}}_{ABXY} \\ P^{\text{obs-inf}}_{X_1X_2Y_1Y_2} &= P^{\text{free}}_{X} \otimes P^{\text{free}}_{X} \otimes P^{\text{free}}_{Y} \otimes P^{\text{free}}_{Y} \end{aligned} \quad \text{and hence} \quad \begin{aligned} P^{\text{obs-inf}}_{A_1B_1|X_1Y_1} &= P^{\text{PR}}_{AB|XY} \\ P^{\text{obs-inf}}_{A_2B_1|X_2Y_1} &= P^{\text{PR}}_{AB|XY} \\ P^{\text{obs-inf}}_{A_1B_2|X_1Y_2} &= P^{\text{PR}}_{AB|XY} \\ P^{\text{obs-inf}}_{A_2B_2|X_2Y_2} &= P^{\text{PR}}_{AB|XY} \end{aligned}. \tag{22}$$ [eqs]{eq:

Given the assumption that the distributions P^{free}_{X} and P^{free}_{Y} are full support, it follows from Eqs. (22) that sometimes $\{X_1, X_2, Y_1, Y_2\} = \{0, 1, 0, 1\}$. For these values, Eqs. (22) specifies perfect correlation between A_1 and B_1. Similarly, it the inflated observational data also specified 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, however: since perfect correlation is transitive, the first three properties entail perfect correlation between A_2 and B_2.

The mathematical structure of this proof parallels that of standard proofs of the inconsistency of PR-box correlations with the Bell structure. Standard proofs focus on a set of variables $\{A_0, A_1, B_0, B_1\}$ where A_0 is the value of A when $X = 0$, and similarly for the others. The fact that there must be a joint distribution over these variables can be inferred from the structure of the Bell DAG and the fact that one can assume without loss of generality that the causal dependencies are deterministic (a result known as Fine's theorem [?] (FineTheorem)). It is then sufficient to note that the marginals given by the PR-box correlations do not arise from any joint distribution. Nonetheless, our proof is conceptually distinct insofar as the variables to which we apply the marginals problem are not conditioned on a setting value. And we do not require Fine's theorem.

Many causal inference techniques can be used to derive sufficient conditions for the inconsistency of the causal hypothesis $H_{G',S'}$ with the observational data $\forall \boldsymbol{U} \in \mathsf{PreInjectableSets}[\boldsymbol{O}] : P(\boldsymbol{U}) = P(\tilde{\boldsymbol{U}}_1)P(\tilde{\boldsymbol{U}}_2)\dots P(\tilde{\boldsymbol{U}}_n)$. We will call sets of such conditions ~~inconsistency witnesses~~ **infeasibility criteria**. I like the name inconsistency witness. Shall we switch everything over to that?

---

[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.

IV. EXAMPLE APPLICATIONS OF THE INFLATION TECHNIQUE

ributions

Here are some examples of causal infeasibility criteria for the Triangle scenario which we can derive by considering the inflation DAG of Fig. 6.

For technical convenience, let us assume that all observed variables take values in $\{-1,+1\}$. By virtue of the existence of a joint distribution of A_2, B_1 and C_1, we can conclude pitowsky_boole_1994,Pitowsky1989,kellerer_marginal_1964,leggett [? ? ? ? ?],

$$\langle A_2C_1\rangle + \langle B_2C_1\rangle - \langle A_2B_1\rangle \leq 1. \tag{23}$$ {eq:polym

A consequence of the manner in which observational data is inflated per Fig. 6, i.e. the non-distribution-specific relations which appear as part of Eqs. (14), we can therefore conclude that if an observable distribution is compatible with the Triangle scenario DAG it must satisfy

$$\langle AC\rangle + \langle BC\rangle \leq 1 + \langle A\rangle\langle B\rangle, \tag{24}$$ {eq:polym

which we think of as a sort of monogamy inequality: it is impossible for C to be strongly correlated with both A and B, unless A and B are both strongly biased.

Alternatively, we could also assume variables with any number of outcomes and start with the inequality fritz2013marginal [?]

$$I(A_2 : C_1) + I(C_1 : B_1) - I(A_2 : B_1) \leq H(B_1), \tag{25}$$

which also simply follows from the existence of a joint distribution of all variables. Again, implicit per Eqs. (14) is that that the third term vanishes regardless of what the observable distribution might be, purely as a consequence of *how* observable distributions are inflated pursuant to Fig. 6. We therefore derive

$$I(A : C) + I(C : B) \leq H(B), \tag{26}$$ {eq:monog

which is the original entropic monogamy inequality derived for the Triangle scenario in fritz2012bell [?]. Our rederivation in terms of inflation is essentially the proof of pusey2014gdag [?].

Slightly more involved but otherwise analogous considerations can be applied to the inflation DAG of Fig. 3, where in particular we have the pre-injectable sets $\{A_1B_1C_1\}$, $\{A_1B_2C_2\}$, $\{A_2B_1C_2\}$, $\{A_2B_2C_1\}$ and $\{A_2B_2C_2\}$, resulting in the factorization relations

$$\begin{aligned}
P_{A_1B_1C_1} &= P_{ABC},\\
P_{A_1B_2C_2} &= P_{AB}\otimes P_C,\\
P_{B_1C_2A_2} &= P_{BC}\otimes P_A,\\
P_{A_2C_1B_2} &= P_{AC}\otimes P_B,\\
P_{A_2B_2C_2} &= P_A\otimes P_B\otimes P_C.
\end{aligned} \tag{27}$$ {eq:tri22

In this case, let us assume binary variables with values in $\{0,1\}$. Now we can again use the existence of a joint distribution over all six observable variables, which implies e.g. the inequality

$$P_{A_2B_2C_2}(111) \leq P_{A_1B_1C_1}(000) + P_{A_1B_2C_2}(111) + P_{B_1C_2A_2}(111) + P_{A_2C_1B_2}(111), \tag{28}$$ {eq:Fritz

which one can show to be valid for every joint distribution of all six variables simply by checking that it holds on every deterministic assignments of values, from which the general case follows by linearity. Applying the above factorization relations turns this into the polynomial inequality

$$P_A(1)P_B(1)P_C(1) \leq P_{ABC}(000) + P_{AB}(11)P_C(1) + P_{BC}(11)P_A(1) + P_{AC}(11)P_B(1), \tag{29}$$ {eq:Fritz

which is yet another necessary condition for compatibility with the Triangle scenario causal structure, and now takes genuine three-way correlations into account. A consequence of this inequality is again that the W-type distribution per Eq. (15) is found to be inconsistent with the Triangle scenario, since the right-hand side vanishes but the left-hand side does not.

Appendix D provides a list of further polynomial inequalities that we have derived for the Triangle scenario using the method developed in the following section.

V. DERIVING POLYNOMIAL INEQUALITIES SYSTEMATICALLY

sec:ineqs

We have defined causal inference as a decision problem, namely testing the consistency of some observational data with some some causal hypothesis. We've shown that this decision can be negatively answered by proxy, namely by demonstrating inconsistency of *inflated* data with an *inflated* hypothesis. The inflation technique can also be used to derive infeasibility criteria, however, by **deriving constraints on the pre-injectable sets**. Any such constraint is also an implicit consequence of the original hypothesis, and hence a relevant infeasibility criterion.

The "big" problem, therefore, is rather straightforward: We seek to derive causal infeasibility criteria from the inflation hypothesis on the injectable sets. This task, however, is just a special instance of generic causal inference: Given some causal hypothesis, what can we say about how it constrains possible observable marginal distribution? Any technique for deriving causal infeasibility criteria is therefore relevant when using the inflation technique. Interestingly, weak constraints from the inflation hypothesis translate into strong constraints pursuant to the original hypothesis.

In the discussion that follows we continue to assume that the original hypothesis is nothing more than supposing the causal structure to be given by the original DAG. Furthermore we presume that the joint distribution over all original observable variables is accessible. Moreover, we limit our attention to deriving polynomial inequalities in terms of probabilities. The potential of using inflation as tool for deriving entropic inequalities is considered separately in Appendix C.

In what follows we consider three different strategies for constraining possible marginal distributions from the inflation hypothesis.

- The full nonlinear strategy attempts to leverage many different kinds of constraints which are implicit in the inflation hypothesis. This strategy yields the strongest infeasibility criteria, but relies on computationally-difficult nonlinear quantifier elimination.
- An intermediate strategy asks only if the various marginal distributions are compatible with *any* joint distribution, without regard to the specific causal structure of the inflation DAG whatsoever. Solving the marginal problem amounts to a special linear quantifier elimination computation, one which can be computed efficiently using convex hull algorithms. The resulting infeasibility criteria are nevertheless still polynomial inequalities.
- Another strategy is based on probabilistic Hardy-type paradoxes, which we connect to the hypergraph transversal problem. This strategy requires the least computational effort, but is limited in that it only yields polynomial inequalities of a very particular form.

In the narrative below the marginal problem is discussed first; the nonlinear strategy is presented as supplementing the marginal problem with additional constraints. The most computationally efficient strategy is presented as a relaxation of the marginal problem, and is discussed separate from the other two strategies, namely in Sec. VIII.

Preliminary to every strategy, however, is the identification of the pre-injectable sets.

njectable

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 $\mathsf{An}(A_iB_j)$ is irredundant. The injectable sets are then precisely the cliques[5] in this graph, per Eq. (7).

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.

Applying these prescriptions to the inflation DAG in Fig. 3 identifies ancestral independencies and (pre-)injectable

[5] A clique is a subset of nodes such that every node in the subset is connected to every other node in the subset.

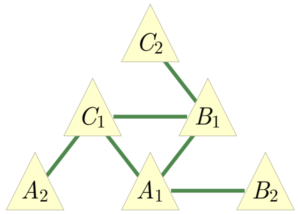

FIG. 9. The auxiliary injection graph corresponding to the inflation DAG in Fig. 3, wherein a pair of nodes are adjacent iff they are pairwise injectable. fig:injection222

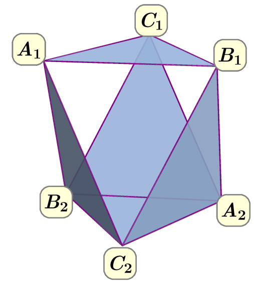

FIG. 10. The simplicial complex... Tobias - you please caption this?The 5 faces in this figure correspond to the 5 maximal pre-injectable sets per Fig. 3, namely $\{A_1B_1C_1\}$, $\{A_1B_2C_2\}$, $\{A_2B_1C_2\}$, $\{A_2B_2C_1\}$ and $\{A_2B_2C_2\}$. fig:simpl

sets as follows:

$$\underbrace{\begin{array}{c} A_2 \Lsh\!\emptyset\!\Rsh B_1 \\ A_2 \Lsh\!\emptyset\!\Rsh C_2 \\ B_2 \Lsh\!\emptyset\!\Rsh A_2 \\ B_2 \Lsh\!\emptyset\!\Rsh C_1 \\ C_2 \Lsh\!\emptyset\!\Rsh A_1 \\ C_2 \Lsh\!\emptyset\!\Rsh B_2 \end{array}}_{\text{pairwise ancestral independencies}} \qquad \underbrace{\begin{array}{c} \{A_2\} \Lsh\!\emptyset\!\Rsh \{B_1B_2C_2\} \\ \{B_2\} \Lsh\!\emptyset\!\Rsh \{A_2C_1C_2\} \\ \{C_2\} \Lsh\!\emptyset\!\Rsh \{A_1A_2B_2\} \\ \{A_2\} \Lsh\!\emptyset\!\Rsh \{B_2\} \Lsh\!\emptyset\!\Rsh \{C_2\} \end{array}}_{\text{maximal ancestral independencies}} \qquad \underbrace{\begin{array}{c} \{A_1B_1\} \\ \{B_1C_1\} \\ \{A_1C_1\} \\ \{A_2C_1\} \\ \{B_2A_1\} \\ \{C_2B_1\} \end{array}}_{\text{pairwise injectable sets}} \qquad \underbrace{\begin{array}{c} \{A_1B_2\} \\ \{B_1C_2\} \\ \{A_2C_1\} \\ \{A_1B_1C_1\} \end{array}}_{\text{maximal injectable sets}} \qquad \underbrace{\begin{array}{c} \{A_1B_1C_1\} \\ \{A_1B_2C_2\} \\ \{A_2B_1C_2\} \\ \{A_2B_2C_1\} \\ \{A_2B_2C_2\} \end{array}}_{\text{maximal pre-injectable sets}} \tag{30}$$

{eq:basic

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

$$\forall_{abc}: \begin{cases} P_{A_1B_1C_1}(abc) = P_{ABC}(abc) \\ P_{A_1B_2C_2}(abc) = P_C(c)P_{AB}(ab) \\ P_{A_2B_1C_2}(abc) = P_A(a)P_{BC}(bc) \\ P_{A_2B_2C_1}(abc) = P_B(b)P_{AC}(ac) \\ P_{A_2B_2C_2}(abc) = P_A(a)P_B(b)P_C(c) \end{cases} \tag{31}$$

{eq:prein

Eq. (31) is an equivalent restatement of Eq. (27). Having identified the pre-injectable sets (and how to use them), we next consider various ways to invoke constraints on the distributions over the pre-injectable sets.

lsproblem

Constraining Possible Distributions over Pre-Injectable Sets via the Marginal Problem

The most trivial constraint on possible marginal probabilities, regardless of causal structure, is simply the *existence of some joint probability distribution* from which the marginal distributions can be recovered through marginalization, i.e. the possible marginal distributions must be **marginally compatible**. This isn't really causal inference – as no hypothesis is considered – but rather more of a preliminary sanity check. If the marginal distributions are not marginally compatible, then the answer to "Can these marginal distributions be explained by this particular causal hypothesis?" is automatically "No".

The necessary and sufficient conditions for marginal compatibility are easy enough to state. There must exist some joint distribution (a collection of nonnegative joint probabilities) such that the marginal distributions can be recovered (each marginal distribution is a sum over various joint probabilities). **Solving the marginal problem** means resolving these "exists" statements into quantifier-free inequalities such that satisfaction of all such inequalities is necessary and sufficient for marginal compatibility. An efficient algorithm to solve the marginal problem is given in Appendix A. The marginal problem comes up in a variety of applications, and has been studied extensively; see [?] fritz2013margina for further references.

As an example, here's how the marginal problem would be phrased as a partial-existential-closure problem with respect to the five three-variable marginal distributions corresponding to the pre-injectable sets in Eq. (31). For simplicity, we assume that all observable variables are binary[6].

[6] If the observables are not binary, then the resulting binary-outcome inequalities are necessary for marginal compatibility, in that they should hold for any course-graining of the observational data into two classes, but the binary-outcome inequalities are no longer sufficient.

In order for the five pre-injectable sets in Eq. (31) to be marginally compatible there must exist 64 nonnegative joint probabilities, i.e. satisfying

$$\forall_{a_1a_2b_1b_2c_1c_2} : \; 0 \le P_{A_1A_2B_1B_2C_1C_2}(a_1a_2b_1b_2c_1c_2), \tag{32}$$

{eq:nonne

constrained further by marginal distributions given by

$$\begin{aligned}
\forall_{a_1b_1c_1} &: \; P_{A_1B_1C_1}(a_1b_1c_1) = \sum\nolimits_{a_2b_2b_2} P_{A_1A_2B_1B_2C_1C_2}(a_1a_2b_1b_2c_1c_2), \\
\forall_{a_1b_2c_2} &: \; P_{A_2B_2C_2}(a_1b_2c_2) = \sum\nolimits_{a_2b_1c_1} P_{A_1A_2B_1B_2C_1C_2}(a_1a_2b_1b_2c_1c_2), \\
\forall_{a_2b_1c_2} &: \; P_{A_2B_1C_2}(a_2b_1c_2) = \sum\nolimits_{a_1b_2c_1} P_{A_1A_2B_1B_2C_1C_2}(a_1a_2b_1b_2c_1c_2), \\
\forall_{a_2b_2c_1} &: \; P_{A_2B_2C_1}(a_2b_2c_1) = \sum\nolimits_{a_1b_1c_2} P_{A_1A_2B_1B_2C_1C_2}(a_1a_2b_1b_2c_1c_2), \\
\forall_{a_2b_2c_2} &: \; P_{A_2B_2C_2}(a_2b_2c_2) = \sum\nolimits_{a_1b_1c_1} P_{A_1A_2B_1B_2C_1C_2}(a_1a_2b_1b_2c_1c_2).
\end{aligned} \tag{33}$$

[eqs]{eq:

The marginal compatibility constraints in this case are, therefore, 64 inequalities and 40 equalities. Solving the marginal problem means eliminating 64 terms from those inequalities and equalities, namely any $p_{A_1A_2B_1B_2C_1C_2}(\_\_\_)$.

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 not-pre-injectable joint probabilities, which are quite different from probabilities involving hidden variables.

When solving the marginal problem is too difficult, one may consider solving a relaxation of it, instead. One extremely computationally amenable relaxation of the marginal problem is to enumerate probabilistic Hardy-type paradoxes. This is discussed later on in Sec. VIII.

Constraining Possible Distributions over Pre-Injectable Sets via Conditional Independence Relations

The marginal problem asks about the existence of *any* joint distribution which recovers the marginal distributions. In causal inference, however, there are plenty of other constraints on the sorts of joint distributions which are consistent with some causal hypothesis. The minimal constraint embedded in any causal hypothesis is the idea of causal structure. Thus it is natural to supplement the marginal problem with additional constraints, motivated by causal structure, on the hypothetical inflation-DAG observable joint distribution.

The most familiar causally-motivated constraints on a joint distribution are **conditional independence relations**, in particular, observable conditional independence relations. 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 the conditional independence $\boldsymbol{X}\perp\boldsymbol{Y}|\boldsymbol{Z}$. The d-separation criterion is explained at length in [? ? ? ?], so we elect not to review it here.

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 rewrite this as follows: If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are d-separated by $\boldsymbol{Z}$, then $p_{\boldsymbol{XYZ}}(\boldsymbol{xyz})p_{\boldsymbol{Z}}(\boldsymbol{z}) = p_{\boldsymbol{XZ}}(\boldsymbol{xz})p_{\boldsymbol{YZ}}(\boldsymbol{yz})$ for all $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$. Such nonlinear constraints can be incorporated as further restrictions on the sorts of joint distributions consistent with the inflation DAG, supplementing the basic nonnegativity of probability constraints of the marginal problem.

For example, in Fig. 3 we find that A_1 and C_2 are d-separated by $\{A_2B_2\}$, and so one might incorporate the family of nonlinear equalities $p_{A_1A_2B_2C_2}(a_1a_2b_2c_2)p_{A_2B_2}(a_2b_2) = p_{A_1A_2B_2}(a_1a_2b_2)p_{A_2B_2C_2}(a_2b_2c_2)$ for all a_1, a_2, b_2 and c_2. Note that every semi-marginal probability introduced in some nonlinear condition independence equation should also be *defined* by marginalization, i.e. as a sum of various joint probabilities, so that one need not eliminate further quantifiers. The relevant substitutions of this example would be

$$\begin{aligned}
\forall_{a_2b_2} &: & P_{A_2B_2}(a_2b_2) &\to & &\sum\nolimits_{a_1b_1c_1c_2} P_{A_1A_2B_1B_2C_1C_2}(a_1a_2b_1b_2c_1c_2), \\
\forall_{a_1a_2b_2c_2} &: & P_{A_1A_2B_2C_2}(a_1a_2b_2c_2) &\to & &\sum\nolimits_{b_1c_1} P_{A_1A_2B_1B_2C_1C_2}(a_1a_2b_1b_2c_1c_2), \\
\forall_{a_1a_2b_2} &: & P_{A_1A_2B_2}(a_1a_2b_2) &\to & &\sum\nolimits_{b_1c_1c_2} P_{A_1A_2B_1B_2C_1C_2}(a_1a_2b_1b_2c_1c_2)
\end{aligned} \tag{34}$$

etc.

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

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

systems with many quantifiers. It may be possible to exploit results on the particular algebraic-geometric structure of these particular systems [?] (garcia_bayesian_2005). But also without using quantifier elimination, the nonlinear constraints can be easily accounted for numerically. Upon substituting numerical values for all the injectable probabilities, the former quantifier elimination problem is converted to a universal quantifier existence problem: Do there exist quantifier that satisfy the full set of linear and nonlinear numeric (as opposed to symbolic) constraints? Most computer algebra systems can resolve such *satisfiability* questions quite rapidly[8].

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.

Constraining Possible Distributions over Pre-Injectable Sets via Coinciding Marginals

The inflation hypothesis is more than just causal structure, however, even if the original hypothesis did not constrain possible causal models beyond the original DAG. By restricting to exclusively *inflation* models, however, we require $P(A_i|\mathsf{Pa}_{G'}(A_i)) = P(A_j|\mathsf{Pa}_{G'}(A_j))$, per Eq. (5). Consequently, the distributions over different injectable sets must occasionally coincide, i.e. $P(\boldsymbol{X}) = P(\boldsymbol{Y})$ whenever both $\boldsymbol{X}$ and $\boldsymbol{Y}$ are injectable, and $\tilde{\boldsymbol{X}} = \tilde{\boldsymbol{Y}}$. Sometimes sets of random variables which are *not* injectable, however, can also be shown to have necessarily coinciding distributions. For example, we can verify that $P(A_1A_2B_1) = P(A_1A_2B_2)$ follows from Fig. 3, even though $\{A_1A_2B_1\}$ and $\{A_1A_2B_2\}$ are not injectable sets.

Equations such as $\forall_{a_1a_2b}: \; p_{A_1A_2B_1}(a_1a_2b) = p_{A_1A_2B_2}(a_1a_2b)$ are also intrinsic parts of the inflation hypothesis, and may incorporated into either linear or nonlinear quantifier eliminations in order to derive stronger infeasibility criteria. The details of how to recognize coinciding distributions beyond the obvious coincidences implied by injectable or pre-injectable sets are discuss in Appendix B.

Move this paragraph up, Tobias? EW 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 [? ? ?] (TavakoliStarNetworks,RossetNetworks,Tavakol…), nor to the nonlinear inequalities which account for interventions to a given causal structure [? ?] (kang2007polynomialconstraints,steeg2011relaxation).

VI. BELL SCENARIOS AND INFLATION

scenarios

To further illustrate our inflation-DAG approach, we demonstrate how to recover all Bell inequalities [? ? ?] (Brunner2013Bell,bell1966l…) 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 [? ? ? ?] (bell1964einstein,Brunner2013Bell,bell1966LHVs,CHSHOriginal) experiment [? (Fig. E#2), ? (Fig. 19), ? (Fig. 1), ? (Fig. 1), ? (Fig. 2b), ? (Fig. 2)] (BeyondBellII, wolfe2015nonconvexity, steeg2011relaxation, WoodSpekkens, chaves2014novel), 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 $P(AB|XY)$ instead of with the original distribution $P(ABXY)$. The conditional distribution is an *array* of distributions over A and B, indexed by the possible values of X and Y. The maximal pre-injectable sets then are

$$\begin{aligned} &\{A_1B_1, X_1X_2Y_1Y_2\}\\ &\{A_1B_2, X_1X_2Y_2Y_2\}\\ &\{A_2B_1, X_1X_2Y_2Y_2\}\\ &\{A_2B_2, X_1X_2Y_2Y_2\}, \end{aligned} \tag{35}$$

where we have put commas in order to clarify that every maximal pre-injectable set contains *all* "settings" variables. These pre-injectable sets are specified by the original observable distribution via

$$\forall_{abx_1x_2y_1y_2}: \begin{cases} P_{A_1B_1X_1X_2Y_1Y_2}(abx_1x_2y_1y_2) = P_{ABXY}(abx_1y_1)P_X(x_2)P_Y(y_2),\\ P_{A_1B_2X_1X_2Y_1Y_2}(abx_1x_2y_1y_2) = P_{ABXY}(abx_1y_2)P_X(x_2)P_Y(y_1),\\ P_{A_2B_1X_1X_2Y_1Y_2}(abx_1x_2y_1y_2) = P_{ABXY}(abx_2y_1)P_X(x_1)P_Y(y_2),\\ P_{A_2B_2X_1X_2Y_1Y_2}(abx_1x_2y_1y_2) = P_{ABXY}(abx_2y_2)P_X(x_1)P_Y(y_1),\\ P_{X_1X_2Y_1Y_2}(x_1x_2y_1y_2) = P_X(x_1)P_X(x_2)P_Y(y_1)P_Y(y_2). \end{cases} \tag{36}$$

---

[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.

By dividing the first four inequalities by the latter we obtain

$$\forall_{abx_1x_2y_1y_2}: \begin{cases} P_{A_1B_1|X_1X_2Y_1Y_2}(ab|x_1x_2y_1y_2) = P_{AB|XY}(ab|x_1y_1) \\ P_{A_1B_2|X_1X_2Y_1Y_2}(ab|x_1x_2y_1y_2) = P_{AB|XY}(ab|x_1y_2) \\ P_{A_2B_1|X_1X_2Y_1Y_2}(ab|x_1x_2y_1y_2) = P_{AB|XY}(ab|x_2y_1) \\ P_{A_2B_2|X_1X_2Y_1Y_2}(ab|x_1x_2y_1y_2) = P_{AB|XY}(ab|x_2y_2). \end{cases} \tag{37}$$

If we then impose marginal compatibility according the the marginal problem we find that the (minimal!) consequence of the inflation hypothesis is

$$P_{A_1B_1|X_1X_2Y_1Y_2}(ab|x_1x_2y_1y_2) = \sum\nolimits_{a_2,b_2} P_{A_1A_2B_1B_2X_1X_2Y_1Y_2}(a_1a_2b_1b_2|x_1x_2y_1y_2) \tag{38}$$

etc. For the inflated observation data to consistent with the inflation hypothesis, therefore, we would require

$$\forall_{ab}: \begin{cases} P_{AB|XY}(ab|00) = \sum_{a',b'} P_{A_1A_2B_1B_2X_1X_2Y_1Y_2}(aa'bb'|0101) \\ P_{AB|XY}(ab|10) = \sum_{a',b'} P_{A_1A_2B_1B_2X_1X_2Y_1Y_2}(a'abb'|0101) \\ P_{AB|XY}(ab|01) = \sum_{a',b'} P_{A_1A_2B_1B_2X_1X_2Y_1Y_2}(aa'b'b|0101) \\ P_{AB|XY}(ab|11) = \sum_{a',b'} P_{A_1A_2B_1B_2X_1X_2Y_1Y_2}(a'ab'b|0101) \end{cases} \tag{39}$$

{eq:final

The existence of an *array* of distributions, i.e. Eq. (39), is equivalent to the existence of a **hidden variable model**, as noted in Fine's Theorem FineTheorem [?]. Thus, an inflation model exists if and only if a hidden-variable model exists for the original observable variables.

In conclusion, we therefore find in the case of the inflation DAG Fig. 8, the inflation method yields necessary and sufficient infeasibility criteria for the Bell causal structure, i.e. the Bell/CHSH inequalities, just by requiring marginal compatibility of the pre-injectable sets. More generally, we can use Fine's theorem to show that applying the marginal problem to suitable inflation DAGs can reproduce *all* Bell inequalities, for any standard Bell scenario, no matter how many parties or settings or possible outcomes.

VII. QUANTUM CAUSAL INFERENCE AND THE NO-BROADCASTING THEOREM

sicallity

In the causal inference problems with latent nodes that we have considered so far, the latent nodes correspond to unobserved random variables. This describes things that come up in *classical* physics (and things outside of physics). In *quantum* physics, however, the latent nodes may instead carry *quantum systems*. Whenever this is allowed, we say that the DAG represents a **quantum causal structure**. Some quantum causal structures are famously capable of generating distributions over the observable variables that would not be possible classically.

The set of quantumly realizable distributions is superficially quite similar to the classical subset pusey2014gdag,fritz2012bell [? ?]. For example, classical and quantum distributions alike respect all conditional independence relations implied by the common underlying causal structure pusey2014gdag [?]. It is an interesting problem to find quantum distributions that are not realizable classically, or to show that there are no such distributions on a given DAG.

However, this is by no means an easy task. For example, recent work has found that quantum causal structure also implies many of the entropic inequalities that hold classically pusey2014gdag,Chaves2015infoquantum,ChavesNoSignalling [? ? ?]. To date, no quantum distribution has been found 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 root variables ("settings") leads to a different kind of entropic inequality, and these have proven somewhat quantum-sensitive braunstein1988entropic,SchumacherInequality,chave [? ? ?]. Such inequalities are still limited, however, in that they only apply in the presence of observable root nodes[9], and they still fail to witness certain distributions as classically infeasible chaves2014novel,fritz2012bell [? ?].

We hope that polynomial inequalities derived from broadcasting inflation DAGs will provide an additional tool for witnessing certain quantum distributions as non-classical. For example due to the results of Sec. VI, it seems conceivable that these inequalities will be much stronger and provide much tighter constraints than entropic criteria.

It is worth pondering how it is possible that some of the inequalities that can be derived via inflation—such as Bell inequalities—have quantum violations, i.e. why one cannot expect them to be valid for all quantum distributions as well. The reason for this is that duplicating an outgoing edge in a DAG during inflation amounts to **broadcasting**

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

the value of the random variable. For example while the information about X in Fig. 1 was "sent" to A and C, the information about X_1 is sent to A_1 *and* A_2 and C in the inflation Fig. 4. Since quantum theory satisfies a no-broadcasting theorem [? ?], one cannot expect such broadcasting to be possible quantumly. More generally, there is an analogous no-broadcasting theorem in the regime of epistemically restricted general probabilistic theories (GPTs) [? ? ? ?], so that the same statement applies in many theories other than quantum theory. As a consequence, a quantum or general probabilistic causal model on the original DAG does generally not inflate to a "quantum inflation model" or "general probabilistic inflation model" on the inflation DAG.

Some inflations, such as the one of Fig. 6, do not require such broadcasting.

Definition 4. *$G' \in \mathsf{Inflations}[G]$ is* ***non-broadcasting*** *if every latent node in G' has at most one copy of each $A \in \mathsf{Nodes}[G]$ among its children.*

It follows that every quantum causal model can be inflated to a non-broadcasting DAG, so that one obtains a quantum and general probabilistic analogue of Lemma 3 in the non-broadcasting case. Constraints derived from non-broadcasting inflations are therefore valid also for quantum and even general probabilistic distributions. In the specific case of the entropic monogamy inequality for the Triangle scenario, i.e. Eq. (26) here, this was originally noticed in Ref. [?]. Another example is Eq. (24), which was derived from the non-broadcasting inflation of Fig. 6. Eq. (24) too, therefore, is also a causal infeasibility criterion which holds for the Triangle scenario even when the latent nodes are allowed to carry quantum or general probabilistic systems. Since the perfect-correlation distribution considered in Eq. (13) violates both of these inequalities, it evidently cannot be generated within the Triangle scenario, even with quantum or general probabilistic states on the hidden nodes. This was also pointed out in Ref. [?].

On the other hand, by intentionally using broadcasting in an inflation DAG, we can specifically try to witness certain quantum or general probabilistic distributions as non-classical. This is exactly what happens in Bell's theorem.

Even when using broadcasting inflation DAGs, it may still be possible to derive inequalities valid for quantum distributions if one appropriately the nonnegativity inequalities in the marginal problem, e.g. such as Eq. (32). The modification would replace demanding nonnegativity of the full joint distribution with instead demanding the nonnegativity of only quantum-physically-meaningful marginal probability distributions.

Even when using broadcasting inflation DAGs, it may still be possible to derive inequalities valid for quantum distributions if one appropriately modifies Sec. V to generate a different initial set of nonnegativity inequalities. This new set should capture the nonnegativity of only quantum-physically-meaningful marginal probability distributions. Indeed, a quantum causal model on the original DAG can potentially be inflated to a quantum inflation model on the inflation DAG in terms of the logical broadcasting maps of [?]. From this perspective, a broadcasting inflation DAG is an abstract logical concept, as opposed to a feasible physical construct. However, this would result in a joint distribution over all observable variables that may have some negative probabilities, and one cannot expect Eq. (32) to hold in general. But one can still try to reformulate the marginal problem so as to refer only to the existence of joint distributions on non-broadcastings sets rather than the existence of a full joint distribution from which the marginal distributions might be recovered. Here, a set $\boldsymbol{U}$ of observable nodes is non-broadcasting if $\mathsf{An}(\boldsymbol{U})$ does not two distinct copies of a node which have a parent in common.

An analysis along these lines has already been carried out successfully by [?] in the derivation of entropic inequalities that are valid for all quantum distributions. Although [?] do not invoke inflation DAGs, they do seem to employ a similar type of structure to model the conditioning of a variable on a "setting" variable, and this also gives rise to non-broadcasting sets. [?] take pains to avoid including full joint probability distributions in any of their initial entropic inequalities, precisely as we would want to do in constructing our initial probability inequalities, and they successfully derive quantumly valid entropic inequalities. But so far, no inequalities polynomial in the probabilities have been derived using this method.

A tight set of inequalities characterizing quantum distributions would provide the ultimate constraints on what quantum theory allows. Deriving additional inequalities that hold for quantum distributions is therefore a priority for future research.

VIII. DERIVING HARDY-TYPE CONSTRAINTS FOR THE MARGINAL PROBLEM

In the literature on Bell inequalities, it has been noticed that the inconsistency with the Bell scenario DAG can sometimes be witnessed by only looking at which joint probabilities are zero and which ones are nonzero. In other words, instead of considering the probability of a joint outcome, some distributions can be witnesses as infeasible by only considering the *possibility* or *impossibility* of each joint outcome. This was originally noticed by [?], and hence such **possibilistic constraints** are also known as **Hardy-type paradoxes**. For a systematic account of Hardy-type paradoxes in Bell scenarios, see Ref. [?], although see also Refs. [? ? ? ? ?].

The usual presentation of a Hardy-type paradox takes the following form: Even though events $E_1..E_N$ never occur (are not possible), nevertheless some other event E_0 *does* occur sometimes (is possible). The impossibility of events $E_1..E_N$ *should* prohibit the possibility of E_0, under the marginal distributions of all the variables in all the events are compatible.

The following observational data illustrates a Hardy-type paradox. Suppose a plague has suddenly wiped out a population of rats. Three kinds of autopsies are performed on different samples of the dead rats. Autopsies checking for heart and brain disease find that rats always presented with one condition but never with *both* conditions, i.e. the events $[Brain_{-}, Heart_{-}]$ and $[Brain_{+}, Heart_{+}]$ are taken to be *not possible*. Suppose the autopsies checking for heart and lung diseases similarly find that all dead rats are afflicted by some disease but with no possibility of dual diagnoses, i.e. the events $[Heart-, Lungs_{-}]$ and $[Heart_{+}, Lungs_{+}]$ are not possible. Logic should then dictate lung disease ought to be perfectly correlated with brain disease, because both those conditions are perfectly anticorrelated with heart disease. If some medical examiner checking rats for brain and lung disease then finds that *it is possible* for a rat to have brain disease without heart disease, this would be considered a violation of a Hardy-type paradox.

Hardy-type paradoxes are predicated on assuming the existence of a solution to the marginal problem. Thus, ensuring the observational data avoids any Hardy-type paradoxes is a necessary condition for solving the marginal problem. [? LSW, Section III.C] discuss the use of possibilistic constraints to certify the incompatibility of marginal distributions in contexts of than Bell scenarios.

The possibilistic constraint which is equivalent to a Hardy-type paradox is as follows,

$$\mathsf{Never}[E_1] \bigwedge \mathsf{Never}[E_2] \bigwedge ...\mathsf{Never}[E_N] \implies \mathsf{Never}[E_0] \tag{40}$$

although it can be expressed more fundamentally in conjuctive form, i.e.

$$[E_0] \implies [E_1] \bigvee [E_2] \bigvee ...[E_N]. \tag{41}$$

Any possibilistic constraint can be immediately translated into a stronger probabilistic one, as noted by [? Mansfield2012]. The probabilistic variant states than *whenever* the event E_0 occurs than at least one of the events $E_1..E_N$ should also occur. Applying the union bound to the probability of the right-hand side we obtain

$$p(E_0) \leq \sum_{n=1}^{N} p(E_n). \tag{42}$$

In order to derive causal infeasibility criteria via marginal compatibility, therefore, we may consider the task of **enumerating** Hardy-type paradoxes, each of which we can then express as an inequalities in terms of probabilities. In the following, we explain how to determine *all* such constraints for *any* marginal problem.

To start with a simple example, suppose that we are in a marginal scenario where the pairwise joint distributions of three variables A, B and C are given.One Hardy-type possibilistic constraint which we would want to enumerate is

$$[\mathbf{A{=}1}, \mathbf{C{=}1}] \implies [\mathbf{A{=}1}, B{=}1] \bigvee [B{=}0, \mathbf{C{=}1}] \tag{43}$$

corresponding to the probabilistic constraint

$$P_{AC}(\mathbf{11}) \leq P_{AB}(\mathbf{1}1) + P_{BC}(0\mathbf{1}). \tag{44}$$ {eq:trivm

Eq. (44) is a necessary condition for the marginal compatibility of $P(AB)$, $P(AC)$, and $P(BC)$; it is equivalent to Eq. (23).

We outline the general procedure using a slightly more sophisticated example. Consider the marginal scenario of Fig. 10, where the contexts are $\{A_1B_1C_1\}$, $\{A_1B_2C_2\}$, $\{A_2B_1C_2\}$, $\{A_2B_2C_1\}$ and $\{A_2B_2C_2\}$, pursuant to Eq. (30). Now a possibilistic constraint on this marginal problem consists of a logical implication with one joint outcome as the **antecedent** and a disjunction of joint outcomes as the **consequent**. In the following, we explain how to generate *all* such implications which are tight in the sense that their right-hand sides are minimal.

First we fix the antecedant by choosing some context and a composite outcome for it. In order to generate all possibilistic constraints, one will have to perform this procedure for *every* context as the antecedent, and every choice of joint outcome thereupon. For the sake of concreteness we take $[\mathbf{A_2{=}1}, \mathbf{B_2{=}1}, \mathbf{C_2{=}1}]$ to be the fixed antecedant.

The consequent will be a conjunction of composite outcomes restricted to marginal contexts, and further restricted so that all marginal composite outcomes in the consequent are compatible with that specified by the consequent. For the implication to be valid, however, the consequent must further be such that **for any *joint* composite outcome which extends the antecentent's marginal composite outcome, also at least one of the marginal composite**

outcomes in the consequent must occur.

To formally determine all valid consequents, we first consider two hypergraphs: The nodes in the first hypegraph correspond to every possible joint outcome for every possible marginal context. The hyperedges in the first hypergraph correspond to every possible joint outcome for the joint context over all variables. A hyperedge (joint compositive outcome) contains a node (marginal composite outcome) iff the hyperedge is an extension of the node, for example the hyperedge $[A_1{=}0, \mathbf{A_2{=}1}, B_1{=}0, \mathbf{B_2{=}1}, C_1{=}1, \mathbf{C_2{=}1}]$ is an extension of the node $[A_1{=}0, \mathbf{B_2{=}1}, \mathbf{C_2{=}1}]$. In our example following Fig. 10, this initial hypergraph has $5 \cdot 2^3 = 40$ nodes and $2^6 = 64$ hyperedges.

The second hypergraph is a sub-hypergraph of the first one. We delete from the first graph all nodes and hyperedges which contradict the supposition of the composite event per our fixed antecedent. For example, the node $[\mathbf{A_2{=}1}, \mathbf{B_2{=}0}, C_1{=}1]$ contradicts the antecedent $[\mathbf{A_2{=}1}, \mathbf{B_2{=}1}, \mathbf{C_2{=}1}]$. We also delete the node corresponding to the antecedent itself. In our example, this final resulting hypergraph has $2^3 + 3 \cdot 2^1 = 14$ nodes and $2^3 = 8$ hyperedges.

All valid (minimal) consequents are (minimal) **transversals** of this latter hypergraph. A tranversal is the sets of nodes which have the property that they intersect every hyperedge in at least one node. In order to get implications which are as tight as possible, it is sufficient to enumerate only the minimal transversals. Doing so is a well-studied problem in computer science with various natural reformulations and for which manifold algorithms have been developed [?]eiter_dualization_2008.

In our example, it is not hard to check that the right-hand side of

$$\begin{aligned}[\mathbf{A_2{=}1}, \mathbf{B_2{=}1}, \mathbf{C_2{=}1}] \Longrightarrow &[A_1{=}0, B_1{=}0, C_1{=}0] \bigvee \mathsf{And}[A_1{=}1, \mathbf{B_2{=}1}, \mathbf{C_2{=}1}] \\ &\bigvee [\mathbf{A_2{=}1}, B_1{=}1, \mathbf{C_2{=}1}] \bigvee [\mathbf{A_2{=}1}, \mathbf{B_2{=}1}, C_1{=}1]\end{aligned} \tag{45}$$ {eq:F3imp

is such a minimal transversal: every assignment of values to all variables which extends the assignment on the left-hand side satisfies at least one of the terms on the right, but this ceases to hold as soon as one removes any one term on the right.

We convert the implications into inequalities in the usual way, by replacing "$\Rightarrow$" by "$\leq$" at the level of probabilities and the disjunctions by sums, so the possibilistic constraint Eq. (45) translates to the probabilistic constraint

$$P_{A_2B_2C_2}(\mathbf{111}) \leq P_{A_1B_1C_1}(111) + P_{A_1B_2C_2}(a_1\mathbf{11}) + P_{A_2B_1C_2}(\mathbf{1}1\mathbf{1}) + P_{A_2B_2C_1}(\mathbf{11}1) \tag{46}$$ {eq:F3raw

Eq. (46) is equivalent to Eq. (28); therefore applying it the inflation depicted in Fig. 3 recovers Eq. (29).

Since also many Bell inequalities are of this form—such as the CHSH inequality which follows in this way from Hardy's original implication—we conclude that this method is still sufficiently powerful to generate plenty of interesting inequalities, and at the same time significantly easier to perform in practice than the full-fledged linear (let alone nonlinear) quantifier elimination.

We note that the connection between classical propositional logic and linear inequalities has been used previously in the task of causal inference. Noteworthy examples of works deriving causal infeasibility criteria via classical logic are [?]Pitowsky1989 and [?]Ghirardi08, see also Refs. [? ? ? ?]pitowsky_boole_1994,kellerer_marginal_1964,leggett_garg_1985,araujo_cycle_2013. Novel this work, however, is the use of the hypergraph transversals problem to formally enumerate relevant logical implications.

We expect that this enumeration of minimal transversals will be computationally much more tractable than the linear quantifier elimination, even if one does it for every possible left-hand side of the implication. We reiterate that inequalities resulting from hypergraph transversals, however, are a subset of the inequalities that result from completely solving the marginal problem. Consequently, linear quantifier elimination is the preferable tool for deriving inequalities whenever the more general linear elimination strategy is computationally tractable.

IX. CONCLUSIONS

Our main contribution is a new way of deriving causal infeasibility criteria, namely the inflation DAG approach. An inflation DAG naturally carries inflation models, and the existence of an inflation model implies inequalities, 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, or through further relaxation to purely *possibilistic* constraints.

These inequalities are necessary conditions on a joint distribution to be explained by the causal structure. We currently do not know to what extent they can also be considered sufficient, and there is somewhat conflicting evidence: as we have seen, the inflation DAG approach reproduces all Bell inequalities; but on the other hand, we have not been able to use it to rederive Pearl's instrumental inequality, although the instrumental scenario also contains only one latent node. By excluding the W-type distribution on the Triangle scenario, we have seen that our polynomial inequalities are stronger than entropic inequalities in at least some cases.

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 some 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 a good inflation from which strong polynomial inequalities can be derived 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 may be 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 crucially relying on the fact that classical hidden variables *can* be cloned. The possibility of classical cloning 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,BeyondBellII [? ? ? ? ?]. 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 Bell inequality NPA2008Long,I3322NPA1 [? ?] is violated in quantum theory, but even finding a quantum violation of one of our *polynomial* inequalities presents a new task for which we currently lack a systematic approach.

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 an alternative strategy that holds some promise.

ACKNOWLEDGMENTS

TF would like to thank Guido Montúfar for discussion and references. 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: Algorithms for Solving the Marginal Problem

lgorithms

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 jones2004equality,JonesThesis2005,Jones2008 [? ? ?]. The oldest-known method for polytope projection, i.e. linear quantifier elimination, is an algorithm known as Fourier-Motzkin (FM) elimination fordan1999projection,DantzigEaves [? ?], although Fourier-Černikov elimination variant Shapot2012,Bastrakov2015 [? ?], as well as Block Elimination and Vertex Enumeration Avis2000lrs [?], are also fairly popular. More advanced polytope projection algorithms, such as Equality Set Projection (ESP) and Parametric Linear Programming, have also recently become available jones2004equality,JonesThesis2005,Jones2008 [? ? ?]. ESP could be an interesting algorithm to use in practice, because each internal iteration of ESP churns out a new facet; by contrast, FM algorithms only generate the entire list of facets after their final internal iteration, after all the quantifiers have been eliminated one by one.

Linear quantifier elimination routines are available in many software tools[10]. We have found custom-coding an linear elimination routine in *Mathematica*^{TM} to be most efficient, see Appendix A for further detail.

The generic task of polytope projection assumes that the initial polytope is given *only* in halfspace representation. If, however, a **dual description** of the initial polytope is available, i.e. we are also given its extremal vertices, then the projection problem can be significantly optimized projectiondual,Avis2000lrs [? ?]. Such dual-description algorithms are used, for example, by modern convex hull solvers. The marginal problem can be explicitly recast as a a special convex hull problem, which can be seen as follows.

A generic convex hull problem takes a matrix of vertices $\hat{V}$, where each row is a vertex, and asks what is the set of coordinates $\vec{x}$ such that $\vec{x}$ can be achieved as a convex combination of the vertices. Formally, it amounts to resolving a partial-existential-closure problem pertaining the existing of nonnegative weights, i.e.

$$\exists_{\vec{w}} : \ \mathsf{And}[\vec{w}.\hat{V} = \vec{x}, \ \sum_i w_i = 1, \ \forall_i w_i \geq 0]. \tag{A-1}$$

It should be evident from Eqs. (32-33) eq:nonnegativity eq:marginalinequalities222 that the marginal problem is precisely of this form. The individual probabilities of the joint outcome correspond to the weights in the convex hull problem. Recasting the marginal problem as a convex hull problem means that optimized convex hull algorithms can be used to directly solve the marginal problem. Fine's Theorem FineTheorem [?] also follows along these lines. Fine's theorem states that the existence of a joint distribution is equivalent to having the observables marginal distribution lie inside the convex hull of all deterministic joint distributions. Tobias, can you say the previous sentence better perhaps?

Indeed, the authors found that the software *lrs* [lrs] was capable of solving the marginal problem rather efficiently.

TABLE I. A comparison of different approaches for constraining the distributions on the pre-injectable sets. The primary divide is quantifier elimination, which is more difficult but produces inequalities, versus satisfiability which can witness the infeasibility of a specific distribution. The approaches subdivide further subdivided into nonlinear, linear, and possibilistic variants.

Approach	General problem	Standard algorithm(s)	Difficulty
Nonlinear quantifier elimination	Real quantifier elimination	Cylindrical algebraic decomposition, see ChavesPolynomial [?]	Very hard
Nonlinear satisfiability	Nonlinear optimization	See BarFT-SMTLIB [?], and semidefinite relaxations laurent_polynomial_2012 [?]	Easy
Linear quantifier elimination	Polytope projection	Fourier-Motzkin elimination, see fordan1999projection,DantzigEaves,Bastrakov20 [? ? ? ? ?]	Hard
Solve marginal problem	Convex Hull	Dual-description linear elimination, see projectiondual,Avis2000lrs [? ?]	Moderate
Linear satisfiability	Linear programming	Simplex method, see Korovin2012ImplementingCRA,Bobot2012SimplexSAT [? ?]	Very easy
Enumerate Hardy paradoxes	Hypergraph transversals	See eiter_dualization_2008 [?]	Very easy

[10] 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 Shapot2012,Bastrakov2015 [? ?] is implemented in *qskeleton* qskeleton [?]. ESP jones2004equality,Jone [? ? ?] 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.

Appendix B: On Identifying All Coinciding Marginal Distributions

ngdetails

We call two sets of observable nodes $\boldsymbol{X}$ and $\boldsymbol{Y}$ **inflationarily isomorphic** (relative to some inflation DAG G') if there exists an isomorphism $\mathsf{AnSubDAG}_{G'}[\boldsymbol{X}] \sim \mathsf{AnSubDAG}_{G'}[\boldsymbol{Y}]$ which has the additional property that, when restricting to $\boldsymbol{X}$, the morphism maps $\boldsymbol{X}$ bijectively to $\boldsymbol{Y}$. For example in Fig. 3, the sets $\boldsymbol{X} = \{A_1A_2B_1\}$ and $\boldsymbol{Y} = \{A_1A_2B_2\}$ are inflationarily isomorphic.

If two sets $\boldsymbol{X}$ and $\boldsymbol{Y}$ are inflationarily isomorphic then we can conclude that $P(\boldsymbol{X}) = P(\boldsymbol{Y})$ in any inflation model.

Coinciding ancestral subgraphs is not, by itself, a strong enough condition to justify coinciding distributions; rather the variables in question must play similar *roles* in their coinciding ancestral subgraphs, hence the definition of inflationary isomorphism. To illustrate this, consider the inflation DAG in Fig. 11. We have $\mathsf{AnSubDAG}_{(\text{Fig. 11})}[X_1Y_2Z_1] \sim \mathsf{AnSubDAG}_{(\text{Fig. 11})}[X_1Y_2Z_2]$, since both of these ancestral subgraphs are the entire DAG. But there is no isomorphism which in addition takes $\{X_1Y_1Z_1\}$ to $\{X_1Y_2Z_2\}$, since the induced subgraphs on these two nodes are different, and therefore $\{X_1Y_1Z_1\} \not\sim \{X_1Y_2Z_2\}$.

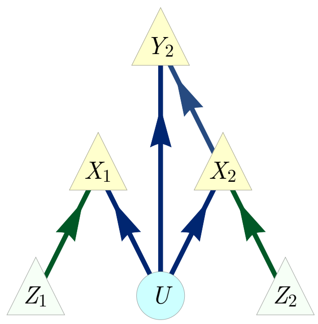

FIG. 11. An inflation DAG which illustrates why coinciding ancestral subgraphs don't necessarily imply coinciding distributions. fig:ancestralsubgraphnotenough

Equality constraint expressing coinciding distributions can be used to supplement the marginal problem, and may be of practical use for reducing the dimensionality of the problem (number of quantifiers which need to be eliminated).

However, we are not aware of any case in which this would actually result in tighter exclusions for the distributions on the original DAG. In many cases, this can be explained by the following argument. Suppose that there is an automorphism of the inflation DAG which takes copies of nodes to copies[11] and restricts to an inflationary isomorphism $\mathsf{AnSubDAG}_{G'}[\boldsymbol{X}] \sim \mathsf{AnSubDAG}_{G'}[\boldsymbol{Y}]$ as above. If $\hat{P}$ solves the unsupplemented marginal problem, then switching the variables in $\hat{P}$ according to the automorphism still solves the unsupplemented marginal problem, since the given marginal distributions are preserved by the automorphism. Now taking the uniform mixture of this new distribution with $\hat{P}$ results in a distribution that still solves the marginal problem, and in addition satisfies the supplemented constraint $P_{\boldsymbol{X}} = P_{\boldsymbol{Y}}$.

Note that the argument does not apply if there is no automorphism which restricts to $\mathsf{AnSubDAG}_{G'}[\boldsymbol{X}] \sim \mathsf{AnSubDAG}_{G'}[\boldsymbol{Y}]$, and it also does not apply if one uses the conditional independence relations on the inflation DAG as well, since this destroys linearity. We do not know what happens in either of these cases.

[11] Note the similarity to *deck transformations* in covering space theory.

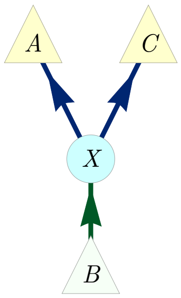

FIG. 12. A causal structure that is compatible with any distribution P_{ABC}. fig:beforecopy

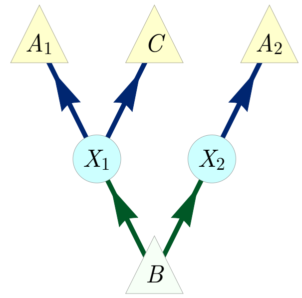

FIG. 13. An inflation. fig:aftercopy

Appendix C: The Copy Lemma and Non-Shannon type Entropic Inequalities

onShannon

As it turns out, the inflation DAG technique is also useful outside of the problem of causal inference. As we argue in the following, inflation is secretly what underlies the **copy lemma** in the derivation of non-Shannon type entropic inequalities yeung_network_2008 [?, Chapter 15]. The following formulation of the copy lemma is the one of Kaced kaced_equivalence_2013 [?].

Lemma 5. *Let A, B and C be random variables with distribution P_{ABC}. Then there exists a fourth random variable A' and joint distribution $P_{AA'BC}$ such that:*

1. $P_{AB} = P_{AB'}$,
2. $A' \perp AC \mid B$. copylemma

Proof. Consider the original DAG of Fig. 12 and the associated inflation DAG of Fig. 13. If the original distribution P_{ABC} is compatible with Fig. 12, then the associated inflation model marginalizes to a distribution $P_{AA'BC}$ which has the required properties. Hence it remains to be shown that every P_{ABC} is compatible with Fig. 12. But this is easy: take λ to be any sufficient statistic for the joint variable (A, B) given C, such as $\lambda := (A, B, C)$. $\square$

While it is also not hard to write down a distribution with the desired properties explicitly yeung_network_2008 [?, Lemma 15.8], our purpose of rederiving the lemma via inflation is our hope that more sophisticated applications of the inflation technique will result in *new* non-Shannon type entropic inequalities.

Appendix D: Classifying polynomial inequalities for the Triangle scenario

c:38ineqs

The following polynomial inequalities for the Triangle scenario have been derived via the linear quantifier elimination method of Sec. V using the inflation DAG of Fig. 3. Initially this has resulted in 64 symmetry classes of inequalities, where the symmetries are given by permuting the variables and inverting the outcomes. For the resulting 64 inequalities, numerical checks have found violations of only 38 of them: although they are all facets of the marginal polytope over the distributions on pre-injectable sets, there is no guarantee that they are also nontrivial inequalities at the level of the original DAG, and this has indeed turned out not to be the case for 26 of these symmetry classes of inequalities. Moreover, it is still likely to be the case that some of these inequalities are redundant; we have not yet checked whether for every inequality there is a distribution which violates the inequality but satisfies all others.

In the following table, the inequalities are listed in expectation-value form, where we assume the two possible outcomes of each variables to be $\{-1, +1\}$.

T: it may be better to list this as a table of coefficients, as e.g. in arXiv:1101.2477, p.14/15?

TABLE II. List of coefficients

constant	$\langle A\rangle$	$\langle B\rangle$	$\langle C\rangle$	$\langle AB\rangle$	$\langle AC\rangle$	$\langle BC\rangle$	$\langle ABC\rangle$	$\langle A\rangle\langle B\rangle$	$\langle A\rangle\langle C\rangle$	$\langle B\rangle\langle C\rangle$	$\langle C\rangle\langle AB\rangle$	$\langle B\rangle\langle AC\rangle$	$\langle A\rangle\langle BC\rangle$	$\langle A\rangle\langle B\rangle\langle C\rangle$
1	0	0	0	1	1	0	0	0	0	1	0	0	0	0
2	0	0	0	0	-2	0	0	0	0	0	-1	0	0	1
3	1	1	1	3	-1	0	0	0	0	-1	1	-1	0	1
3	1	1	-1	3	1	0	0	0	0	1	-1	-1	0	1
3	0	0	1	-2	0	-2	0	1	0	0	-1	-1	0	1
3	0	1	0	1	0	-2	0	-1	1	0	1	-1	0	1
3	0	1	0	1	0	-2	0	1	-1	0	1	1	0	-1
3	1	1	1	2	2	2	-1	1	1	1	1	1	1	-1
3	1	1	1	2	0	-2	1	1	-1	1	1	1	-1	-1
4	0	0	2	-2	-2	0	-1	2	0	2	1	1	1	0
4	0	-2	0	-2	0	-3	1	0	0	1	1	-1	0	1
4	0	0	-2	-2	-2	-3	1	2	0	1	1	1	0	-1
4	0	0	0	2	-2	1	1	2	2	-1	1	-1	0	-1
4	0	0	0	2	-2	1	1	-2	2	-1	1	1	0	1
4	0	0	0	-2	0	3	1	2	0	1	-1	-1	0	1
4	0	0	-2	-2	-2	-2	1	2	0	0	1	1	-1	0
4	0	0	0	-2	-2	-2	1	2	2	2	1	1	1	0
5	1	1	1	3	1	-4	0	-2	0	1	1	-1	0	1
5	1	1	1	3	-1	-4	0	2	-2	1	1	1	0	-1
5	1	-1	1	1	2	-2	-2	-2	-1	1	1	-2	-2	0
5	3	1	1	1	3	1	-1	2	0	0	-2	0	0	2
5	1	1	1	1	2	-2	-1	0	-1	-1	2	1	1	-2
5	-1	1	1	1	1	-1	1	-2	-2	2	-2	-2	-2	0
5	1	1	1	2	1	-1	1	-1	0	2	-1	-2	-2	1
5	1	1	1	-1	2	2	1	-2	-1	-1	2	1	-1	-2
6	0	0	0	-3	-4	0	0	1	2	2	-1	-2	-2	1
6	0	2	0	3	-4	0	0	1	2	0	1	-2	-2	1
6	-2	2	0	-3	-5	0	0	1	1	0	1	-1	-2	2
6	0	0	0	1	-3	2	0	1	1	-4	1	-1	-2	-2
6	0	0	2	0	3	-5	0	-2	1	1	2	1	1	-2
6	0	0	-2	2	-2	1	0	-4	2	-1	2	2	1	1
6	0	0	0	-3	-2	-2	-2	1	0	4	-1	-2	0	1
7	1	1	1	2	1	-3	3	1	-2	2	3	2	-2	-1
8	0	0	0	-4	-2	-2	-3	4	2	-2	1	-1	-3	2
8	2	-2	0	1	-6	0	1	-1	0	2	2	1	-3	3
8	2	0	0	6	1	-2	1	0	1	2	-1	-2	-3	3
8	0	-2	-2	0	-6	1	1	2	0	-1	3	1	-2	-3
8	0	0	2	2	1	-6	1	-2	1	0	3	2	-1	-3

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

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

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

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

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

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

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

(#10): $0 \le 4 + 2 \langle A\rangle \langle B\rangle + 2 \langle C\rangle + 2 \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle + \langle A\rangle \langle BC\rangle - \langle ABC\rangle$

(#11): $0 \le 4 - 2 \langle B\rangle + \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - \langle B\rangle \langle AC\rangle - 3 \langle BC\rangle + \langle ABC\rangle$

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

(#13): $0 \le 4 + 2 \langle A\rangle \langle B\rangle + 2 \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle - \langle B\rangle \langle AC\rangle + \langle BC\rangle + \langle ABC\rangle$

(#14): $0 \le 4 - 2 \langle A\rangle \langle B\rangle + 2 \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle + \langle BC\rangle + \langle ABC\rangle$

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

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

(#20): $0 \le 5 + \langle A\rangle - \langle B\rangle - 2 \langle A\rangle \langle B\rangle + \langle C\rangle - \langle A\rangle \langle C\rangle + \langle B\rangle \langle C\rangle + \langle AB\rangle + \langle C\rangle \langle AB\rangle + 2 \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle - 2 \langle ABC\rangle$

(#21): $0 \le 5 + 3 \langle A\rangle + \langle B\rangle + 2 \langle A\rangle \langle B\rangle + \langle C\rangle + 2 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle - 2 \langle C\rangle \langle AB\rangle + 3 \langle AC\rangle + \langle BC\rangle - \langle ABC\rangle$

(#22): $0 \le 5 + \langle A\rangle + \langle B\rangle + \langle C\rangle - \langle A\rangle \langle C\rangle - \langle B\rangle \langle C\rangle - 2 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle + 2 \langle AC\rangle + \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle + \langle A\rangle \langle BC\rangle - \langle ABC\rangle$

(#23): $0 \le 5 - \langle A\rangle + \langle B\rangle - 2 \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle A\rangle \langle C\rangle + 2 \langle B\rangle \langle C\rangle + \langle AB\rangle - 2 \langle C\rangle \langle AB\rangle + \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

(#24): $0 \le 5 + \langle A\rangle + \langle B\rangle - \langle A\rangle \langle B\rangle + \langle C\rangle + 2 \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle - \langle C\rangle \langle AB\rangle + \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

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

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

(#28): $0 \le 6 - 2 \langle A\rangle + 2 \langle B\rangle + \langle A\rangle \langle B\rangle + \langle A\rangle \langle C\rangle + 2 \langle A\rangle \langle B\rangle \langle C\rangle - 3 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 5 \langle AC\rangle - \langle B\rangle \langle AC\rangle - 2 \langle A\rangle \langle BC\rangle$

(#29): $0 \le 6 + \langle A\rangle \langle B\rangle + \langle A\rangle \langle C\rangle - 4 \langle B\rangle \langle C\rangle - 2 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + \langle C\rangle \langle AB\rangle - 3 \langle AC\rangle - \langle B\rangle \langle AC\rangle + 2 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle$

(#30): $0 \le 6 - 2 \langle A\rangle \langle B\rangle + 2 \langle C\rangle + \langle A\rangle \langle C\rangle + \langle B\rangle \langle C\rangle - 2 \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle C\rangle \langle AB\rangle + 3 \langle AC\rangle + \langle B\rangle \langle AC\rangle - 5 \langle BC\rangle + \langle A\rangle \langle BC\rangle$

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

(#32): $0 \le 6 + \langle A\rangle \langle B\rangle + 4 \langle B\rangle \langle C\rangle + \langle A\rangle \langle B\rangle \langle C\rangle - 3 \langle AB\rangle - \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle - 2 \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 2 \langle ABC\rangle$

(#33): $0 \le 7 + \langle A\rangle + \langle B\rangle + \langle A\rangle \langle B\rangle + \langle C\rangle - 2 \langle A\rangle \langle C\rangle + 2 \langle B\rangle \langle C\rangle - \langle A\rangle \langle B\rangle \langle C\rangle + 2 \langle AB\rangle + 3 \langle C\rangle \langle AB\rangle + \langle AC\rangle + 2 \langle B\rangle \langle AC\rangle - 3 \langle BC\rangle - 2 \langle A\rangle \langle BC\rangle + 3 \langle ABC\rangle$

(#34): $0 \le 8 + 4 \langle A\rangle \langle B\rangle + 2 \langle A\rangle \langle C\rangle - 2 \langle B\rangle \langle C\rangle + 2 \langle A\rangle \langle B\rangle \langle C\rangle - 4 \langle AB\rangle + \langle C\rangle \langle AB\rangle - 2 \langle AC\rangle - \langle B\rangle \langle AC\rangle - 2 \langle BC\rangle - 3 \langle A\rangle \langle BC\rangle - 3 \langle ABC\rangle$

(#35): $0 \le 8 + 2 \langle A\rangle - 2 \langle B\rangle - \langle A\rangle \langle B\rangle + 2 \langle B\rangle \langle C\rangle + 3 \langle A\rangle \langle B\rangle \langle C\rangle + \langle AB\rangle + 2 \langle C\rangle \langle AB\rangle - 6 \langle AC\rangle + \langle B\rangle \langle AC\rangle - 3 \langle A\rangle \langle BC\rangle + \langle ABC\rangle$

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

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