Expressivity Landscape for Logics with Probabilistic Interventionist Counterfactuals

Fausto Barbero
University of Helsinki, Finland

Jonni Virtema
University of Sheffield, UK

Abstract

Causal multiteam semantics is a framework where probabilistic dependencies arising from data and causation between variables can be together formalized and studied logically. We discover complete characterizations of expressivity for several logics that can express probabilistic statements, conditioning and interventionist counterfactuals. The results characterize the languages in terms of families of linear equations and closure conditions that define the corresponding classes of causal multiteams. The characterizations yield a strict hierarchy of expressive power. Finally, we present some undefinability results based on the characterizations.

2012 ACM Subject Classification Computing methodologies → Probabilistic reasoning; Mathematics of computing → Causal networks

Keywords and phrases Interventionist counterfactuals, Multiteam semantics, Causation, Probability logic, Linear inequalities, Expressive power

Digital Object Identifier 10.4230/LIPIcs.CSL.2024.15

Related Version *Full Version*: `https://doi.org/10.48550/arXiv.2303.11993`

Funding *Fausto Barbero*: DFG grant VI 1045-1/1; Academy of Finland grants 316460 and 349803
Jonni Virtema: DFG grant VI 1045-1/1; Academy of Finland grant 345634

1 Introduction

The main approach to the study of empirical data in the 20th century has been that of statistics, which makes use of probabilistic notions such as *correlation* and *conditional (in)dependence* between variables. We follow here another line of study – going back at least to Sewall Wright [40] – insisting that the analysis should not stop at correlations, but instead should yield information about causation among variables (conditional on appropriate scientific assumptions). The methods involved in the analysis of causes and effects have gained in popularity in the last decades, and their mathematics has been vastly developed under the label of *causal inference* (see, e.g., [30, 35]). Today the methods of causal inference are heavily utilized, e.g., in epidemiology [23], econometrics [22], social sciences [28] and machine learning [32, 33].

One of the next crucial steps in the development of artificial intelligence will be the capability of AI systems to represent and reason about causal knowledge (see, e.g., [31]). For the development of AI applications of causal inference, the clarification of the related formal logical theory is vital. It turns out that many concepts involved in the analysis of causes can be reduced to the study of *interventionist counterfactuals* in *causal models*. Causal models represent causation between variables using so-called *structural equations*, which describe deterministic causal laws that relate the variables to each other. In their simplest form, interventionist counterfactuals are expressions of the form

“if variables $X_1, \dots, X_n$ were set to values $x_1, \dots, x_n$, then Y would take value y”.

Such conditionals are counterfactual (contrary to fact) in that their evaluation forces us to consider an alternative scenario in which the variables $X_1, \ldots, X_n$ are subtracted to the laws that currently determine their behaviour, and the (possibly new) values taken by such variables are fixed by some external intervention. The causal laws encoded in the model then allow us to find out, computationally, how all the variables in the system are affected in this alternative scenario. Research on logics encompassing interventionist counterfactuals has been active in the past two decades. For example, [13, 17, 7, 9] provided complete axiomatizations for languages of increasing generality. The papers [18, 41] drew precise connections with the earlier Stalnaker-Lewis theory of counterfactuals [36, 27]. In [1] logics for causal reasoning were studied via translations to first-order logic, and the articles [17, 16, 29] discuss the complexity of causal and probabilistic languages.

The classical literature on causal inference does not neatly separate the methods of probability and of causal analysis; many standard concepts in causal inference are expressed by mixing probabilistic and causal statements. In other words, causal inference uses an array of new notational devices that are not entirely reducible to classical probabilistic reasoning; two significant examples (from [30]) of these new notations are the conditional *do* expressions ($Pr(y \mid do(x), z)$) and Pearl's "counterfactuals" ($Pr(Y_{X=x} \mid Z = z)$). We refer the reader to [8] for a detailed discussion of the meaning and use of these expressions. Roughly speaking, they both describe the probability that the variable Y takes value y after intervening to set X to x, conditional upon the observation that Z takes value z; but the two expressions differ subtly in that in the former the conditioning is performed in the system modified by the intervention that sets X to x, while in the latter expression conditioning is relative to the pre-intervention system. To this regard, we follow the proposal of Barbero and Sandu [2, 4] to decompose these complex causal-probabilistic expressions in terms of a minimal set of logical primitives. In particular, probabilistic conditioning and causal interventions will correspond to two distinct logical conditionals, $\supset$ and $\Box\!\!\rightarrow$.

In order to make this decomposition possible, one needs to move from causal models to the more general *causal multiteam semantics*, where all the needed logical operators are available. Team semantics is the semantical framework of modern logics of dependence and independence. Introduced by Hodges [24] and adapted to dependence logic by Väänänen [38], team semantics defines truth in reference to collections of assignments, called *teams*. Team semantics is particularly suitable for the formal analysis of dependencies and independencies in data. Recent developments in the area have broadened the scope of team semantics to cover probabilistic and quantitative notions of dependence and independence. Durand et al. [11, 10] introduced multiset and probabilistic variants of team semantics as frameworks for studying probabilistic dependency notions such as conditional independence logically. Further analysis has revealed that definability and complexity of logics in these frameworks are intimately connected to definability and complexity of Presburger ([14, 39]) and real arithmetic ([21, 20]).

Causal teams, proposed by Barbero and Sandu [3], fuse together teams and causal models, and model inferences encompassing both functional dependencies arising from data and causal dependencies arising from structural equations. The logics considered by Barbero and Sandu use atomic expressions of the form $X = x$ and $=\!(X; Y)$ to state that the variable X takes the value x and that (in the data) the value of the variable Y is functionally determined by the values of the variable X, respectively. *Interventionist counterfactuals* ($X = x \Box\!\!\rightarrow \psi$) and *selective implications* ($\alpha \supset \psi$) then describe consequences of actions and consequences of learning from observations. For example, the intended reading of the formula "Pressure = 300 $\Box\!\!\rightarrow$ Volume = 4" is: *If we raise the pressure to* 300 *kPa, the*

$$\begin{tikzcd}
 & & \mathcal{P}(\supset) \arrow[dr, "\text{Corollary 10}"] & & \\
\mathcal{P}^- \arrow[r, "\text{Lemma 12}"] & \mathcal{P} \arrow[ur, "\text{Prop. 18}"] \arrow[dr, "\text{Corollary 10}"'] & & \mathcal{PCO} \arrow[r, "\text{Cor. 20}"] & \mathcal{PCO}^\omega \\
 & & \mathcal{P}(\Box\!\!\rightarrow) \arrow[ur, "\text{Prop. 18}"'] & &
\end{tikzcd}$$

Figure 1 Arrows denote strict inclusion of expressivity; $\mathcal{P}(\Box\!\!\rightarrow)$ and $\mathcal{P}(\supset)$ are incomparable.

volume of the gas will be $4\ m^3$*.* On the other hand, the intended reading of the formula "Pressure = 20 $\supset$ 10 < Altitude < 30" is: *If we read* 20 *kPa from the barometer, the current altitude is between 10 and 30 km.*

Finally, the *causal multiteam semantics* coined by Barbero and Sandu [4] fuses together multiteams and causal models. The shift from teams to multiteams makes it now possible to study probabilistic conditioning and causal interventions in a unified framework. Barbero and Sandu study a language called $\mathcal{PCO}$ (for Probabilities, Causes and Observations) which they claim to capture a fair portion of the probabilistic causal reasoning that appears in the field of causal inference. It does indeed suffice to capture many forms of probabilistic conditioning, and it suffices to express conditional *do* expressions, the "Pearl counterfactuals" mentioned above and more general kinds of statements. For example, the statement "the probability that a sick untreated patient would be healed when treated is at least $\frac{2}{3}$" can be formalised as (Sick = 1 $\land$ Treated = 0) $\supset$ (Treated = 1 $\Box\!\!\rightarrow$ Pr(Sick = 0) $\geq \frac{2}{3}$). The paper [4] raises however the doubt whether $\mathcal{PCO}$ can express, in general, the comparison of conditional probabilities (e.g., statements of the form $\Pr(\alpha \mid \beta) \geq \Pr(\gamma \mid \delta)$). We show here that it fails to do so; thus, $\mathcal{PCO}$ cannot be used, for instance, to compare the expected efficacy of two distinct (non-enforced) medical treatments.

The cornerstone of this inexpressibility result is an abstract characterization of the expressive power of $\mathcal{PCO}$, which in particular shows that the classes of probability distributions that are consistent with a given $\mathcal{PCO}$ formula can be described in terms of a certain class of linear inequalities. On the other hand, by a geometrical argument we see that there are statements of comparison of conditional probabilities which unavoidably involve inequalities of second degree. The quest for an understanding of language $\mathcal{PCO}$ naturally proceeds via an understanding of the expressivity of its key resources: evaluation atoms ($\Pr(\alpha) \geq \epsilon$), comparison atoms ($\Pr(\alpha) \geq \Pr(\beta)$), observations ($\supset$) and interventions ($\Box\!\!\rightarrow$). This leads us to the study of four fragments $\mathcal{P}^-$, $\mathcal{P}$, $\mathcal{P}(\supset)$, and $\mathcal{P}(\Box\!\!\rightarrow)$. We characterize the expressive power of each of these sublogics, as well as the expressivity of $\mathcal{PCO}$, in terms of closure properties and of an appropriate class of linear inequalities; these results are schematized in Table 1. Together with geometrical reasoning, these characterizations yield a strict hierarchy of expressive power, as summarized in Figure 1. The table and the figure also include a language $\mathcal{PCO}^\omega$ that extends $\mathcal{PCO}$ with (countably) infinite disjunctions. The manuscript [4] already shows that this language is more expressive than $\mathcal{PCO}$; our results yield an alternative proof.

The characterization and hierarchy results can be found in Section 3, after a presentation of the semantics and syntax of the languages (Section 2). Section 4 presents the inexpressibility result for conditional comparison atoms, and briefly discusses the related issue of definability of dependencies and independencies.

Table 1 Characterizations of expressivity of logics. E.g., a class $\mathcal{K}$ of causal multiteams is definable by a $\mathcal{P}(\supset)$-formula iff $\mathcal{K}$ is signed binary, closed under change of laws and rescaling, and has the empty multiteam property. $\mathcal{K}$ is a union of signed binary, when $\mathcal{K} = \bigcup_{\mathcal{F} \in \mathbb{F}_\sigma} \mathcal{K}^{\mathcal{F}}$, for signed binary sets of causal multiteams $\mathcal{K}^{\mathcal{F}}$ of function component $\mathcal{F}$.

Logic	Type of inequalities	Closure properties: change of laws	Closure properties: rescaling & empty multiteam	References
$\mathcal{P}^-$	monic	X	X	Thm. 11
$\mathcal{P}$	signed monic	X	X	Thm. 11
$\mathcal{P}(\supset)$	signed binary	X	X	Thm. 16
$\mathcal{P}(\boxright)$	union of signed monics		X	Thm. 14
$\mathcal{PCO}$	union of signed binary		X	Thm. 19
$\mathcal{PCO}^\omega$	(unrestricted)		X	[4]

2 Logics with causal multiteam semantics

Capital letters such as $X, Y, \ldots$ denote **variables** (standing for specific magnitudes such as "temperature" and "volume") which take **values** denoted by small letters. The values of the variable X will be often denoted by $x, x', \ldots$. Sets (and tuples, depending on the context) of variables and values are denoted by boldface letters such as $\mathbf{X}$ and $\mathbf{x}$. We consider probabilities that arise from the counting measures of finite (multi)sets. For finite sets $S \subseteq T$, we define $P_T(S) := \frac{|S|}{|T|}$.

A **signature** is a pair $(\mathrm{Dom}, \mathrm{Ran})$, where Dom is a finite set of variables and Ran a function mapping each $X \in \mathrm{Dom}$ to a finite set $\mathrm{Ran}(X)$ of values (the **range** of X). We stipulate a fixed ordering on Dom, and write $\mathbf{W}$ for the tuple of all the variables of Dom listed in that order. We write $\mathbf{W}_X$ for the variables of $\mathrm{Dom} \setminus \{X\}$ listed according to the fixed order. For a tuple $\mathbf{X} = (X_1, \ldots, X_n)$ of variables, $\mathrm{Ran}(\mathbf{X})$ denotes the Cartesian product $\mathrm{Ran}(X_1) \times \cdots \times \mathrm{Ran}(X_n)$. An **assignment** of signature σ is a mapping $s : \mathrm{Dom} \to \bigcup_{X \in \mathrm{Dom}} \mathrm{Ran}(X)$ such that $s(X) \in \mathrm{Ran}(X)$ for each $X \in \mathrm{Dom}$. The set of all assignments of signature σ is denoted by $\mathbb{B}_\sigma$. For an assignment s having the variables of $\mathbf{X}$ in its domain, $s(\mathbf{X})$ denotes the tuple $(s(X_1), \ldots, s(X_n))$. For $\mathbf{X} \subseteq \mathrm{Dom}$, $s_{\upharpoonright \mathbf{X}}$ is the restriction of s to the variables in $\mathbf{X}$.

A **team** T of signature σ is a subset of $\mathbb{B}_\sigma$. Intuitively, a multiteam is just a multiset analogue of a team. We represent **multiteams** as (finite) teams with an extra variable Key (not belonging to the signature) ranging over $\mathbb{N}$, which takes different values over different assignments of the team, and which is never mentioned in the formal languages. A multiteam can be then presented as a table; e.g., the following

T :

Key	X	Y
0	0	0
1	0	0
2	0	1

describes a multiteam containing two "copies" of the assignment $s(X, Y) = (0, 0)$ (first two rows) plus another assignment $t(X, Y) = (0, 1)$. We will say that the variable domain of this multiteam T is $\mathrm{Dom} = \{X, Y\}$, and omit mentioning the Key variable. Multiteams will be used to encode probability distributions over the underlying team (in this case, the distribution that assigns probability $\frac{2}{3}$ to assignment s, and probability $\frac{1}{3}$ to t). The "underlying team" (i.e., support of a multiteam) is characterized formally later in Definition 6.

Multiteams by themselves do not encode any solid notion of causation; they do not tell us how a system would be affected by an intervention. We therefore need to enrich multiteams with additional structure. In particular, we will associate to some of the variables a deterministic causal law. The law for variable V takes the form of a function, which describes the way the value of V is generated from the values of other variables in the system. These laws will be used crucially in order to compute how the model is affected by an intervention. Furthermore, we will require that each assignment in the multiteam agrees with these laws.

▶ **Definition 1.** *A **causal multiteam** of signature* (Dom, Ran) *with **endogenous variables*** $\mathrm{End}(T) \subseteq \mathrm{Dom}$ *is a pair* $T = (T^-, \mathcal{F})$ *such that:*

1. T^- *is a multiteam of domain* Dom*,*
2. $\mathcal{F}$ *is a function* $\{(V, \mathcal{F}_V) \mid V \in \mathrm{End}(T)\}$ *that assigns to each endogenous variable* V *a non-constant* $|\mathbf{W}_V|$*-ary function* $\mathcal{F}_V : \mathrm{Ran}(\mathbf{W}_V) \to \mathrm{Ran}(V)$*,*
3. $(T^-, \mathcal{F})$ *satisfies the **compatibility constraint:*** $\mathcal{F}_V(s(\mathbf{W}_V)) = s(V)$*, for all* $s \in T^-$ *and* $V \in \mathrm{End}(T)$*.*

T^- *and* $\mathcal{F}$ *will be called, respectively, the **multiteam component** and the **function component** of* T*. We write* $(\mathrm{Dom}(T), \mathrm{Ran}(T))$ *to denote the signature of the causal multiteam* T*.*

Notice that, due to the compatibility constraint, not all instances for $End(T)$ and T^- give rise to causal multiteams. The function component $\mathcal{F}$ induces a system of structural equations; an equation $V := \mathcal{F}_V(\mathbf{W}_V)$ for each variable $V \in \mathrm{End}(T)$. Note that some of the variables in $\mathbf{W}_V$ may not be necessary for evaluating V. For example, if V is given by the structural equation $V := X + 1$, all the variables in $\mathbf{W}_V \setminus \{X\}$ are irrelevant (we call them **dummy arguments** of $\mathcal{F}_V$). The set of non-dummy arguments of $\mathcal{F}_V$ is denoted as PA_V (the set of **parents** of V).

We associate to each causal multiteam T a **causal graph** G_T, whose vertices are the variables in Dom and where an arrow is drawn from each variable in PA_V to V, whenever $V \in \mathrm{End}(T)$ (see Example 3 and picture 2 for a depiction). The variables in $\mathrm{Dom}(T) \setminus \mathrm{End}(T)$ are called **exogenous** (written $\mathrm{Exo}(T)$).

In the present paper we restrict attention to systems of variables that are connected by causal laws that do not form cycles (e.g., we exclude the possibility that X causally affects Y, Y causally affects Z, and in turn Z affects X); such systems are usually called *recursive*. Concretely, we enforce the following convention:

Throughout the paper we will implicitly assume that causal multiteams have an acyclic causal graph.

While the study of cyclic systems is far from absent from the literature (e.g. [37, 34, 17, 1]), in a probabilistic context it introduces a number of complications that go well beyond the scope of the framework considered in this paper.

▶ **Definition 2.** *A causal multiteam* $S = (S^-, \mathcal{F}_S)$ *is a **causal sub-multiteam** of* $T = (T^-, \mathcal{F}_T)$*, if they have the same signature,* $S^- \subseteq T^-$*, and* $\mathcal{F}_S = \mathcal{F}_T$*. We then write* $S \leq T$*.*

We consider causal multiteams as dynamic models, that can be affected by observations and interventions. Given a causal multiteam $T = (T^-, \mathcal{F})$ and a formula α of some formal language (evaluated over causal multiteams according to some semantic relation $\models$), "observing α" produces the causal sub-multiteam $T^\alpha = ((T^\alpha)^-, \mathcal{F})$ of T, where $(T^\alpha)^- := \{s \in T^- \mid (\{s\}, \mathcal{F}) \models \alpha\}$.[1] An intervention on T will *not*, in general, produce a sub-multiteam of T. It will instead

[1] Throughout the paper, the semantic relation in terms of which T^α is defined will be the semantic relation for language CO, which shall be defined below.

T^-:

Key	X	Y	Z
0	0	1	1
1	1	2	3
2	1	2	3
3	2	3	5
4	2	3	5
5	2	3	5

$\rightsquigarrow$

Key	X	Y	Z
0	0	**1**	...
1	1	**1**	...
2	1	**1**	...
3	2	**1**	...
4	2	**1**	...
5	2	**1**	...

$\rightsquigarrow$ $T^-_{Y=1}$:

Key	X	Y	Z
0	0	1	**1**
1	1	1	**2**
2	1	1	**2**
3	2	1	**3**
4	2	1	**3**
5	2	1	**3**

Figure 2 Causal multiteams for Example 3, showing how the multiteam component $T^-_{Y=1}$ of a causal multiteam is computed from T^- given an intervention $do(Y = 1)$. The figure also describes the associated causal graphs.

modify the values that appear in some of the columns of T. We consider interventions that are described by conjunctions of the form $X_1 = x_1 \wedge \dots \wedge X_n = x_n$ (or, shortly, $\mathbf{X} = \mathbf{x}$). Such a formula is **inconsistent** if there are two indexes i, j such that X_i and X_j denote the same variable, while x_i and x_j denote distinct values; it is **consistent** otherwise. Applying an intervention $do(\mathbf{X} = \mathbf{x})$, where $\mathbf{X} = \mathbf{x}$ is consistent, to a causal multiteam $T = (T^-, \mathcal{F})$ of endogenous variables $\mathbf{V}$ will produce a causal multiteam $T_{\mathbf{X}=\mathbf{x}} = (T^-_{\mathbf{X}=\mathbf{x}}, \mathcal{F}_{\mathbf{X}=\mathbf{x}})$, where the function component is $\mathcal{F}_{\mathbf{X}=\mathbf{x}} := \mathcal{F}_{\upharpoonright(\mathbf{V}\setminus\mathbf{X})}$ (the restriction of $\mathcal{F}$ to the set of variables $\mathbf{V} \setminus \mathbf{X}$) and the multiteam component is $T^-_{\mathbf{X}=\mathbf{x}} := \{s^{\mathcal{F}}_{\mathbf{X}=\mathbf{x}} \mid s \in T^-\}$, where each $s^{\mathcal{F}}_{\mathbf{X}=\mathbf{x}}$ is the unique assignment compatible with $\mathcal{F}_{\mathbf{X}=\mathbf{x}}$ defined (recursively) as

$$s^{\mathcal{F}}_{\mathbf{X}=\mathbf{x}}(V) = \begin{cases} x_i & \text{if } V = X_i \in \mathbf{X} \\ s(V) & \text{if } V \in \mathrm{Exo}(T) \setminus \mathbf{X} \\ \mathcal{F}_V(s^{\mathcal{F}}_{\mathbf{X}=\mathbf{x}}(\mathbf{W}_V)) & \text{if } V \in \mathrm{End}(T) \setminus \mathbf{X}. \end{cases}$$

We emphasize that the uniqueness of $s^{\mathcal{F}}_{\mathbf{X}=\mathbf{x}}$, and thus the correctness of this definition, hinges on our assumption that the causal graphs are acyclic. For an explanation of how interventions may be defined in the cyclic (non-probabilistic) case, see [1].

▶ **Example 3.** Consider the causal multiteam $T = (T^-, \mathcal{F})$ depicted in Figure 2, where each row of the leftmost table depicts an assignment of T^- (e.g., the third row represents an assignment s with $s(Key) = 2$, $s(X) = 1$, $s(Y) = 2$, $s(Z) = 3$). The rows of the table are compatible with the laws $\mathcal{F}_Z(X, Y) = X + Y$ and $\mathcal{F}_Y(X) = X + 1$, while X is exogenous. T encodes probabilities for formulas that discuss variables X, Y, Z and their possible values; for example, $P_T(Z = 3) = \frac{1}{3}$.

Suppose we can enforce the variable Y to take the value 1. The effect of such an intervention, depicted in the right-hand side of Figure 2, is to first set the value of Y to 1 (in all rows) and then to recompute the values of Z using the function $\mathcal{F}_Z$. The probability distribution has changed: now $P_{T_{Y=1}}(Z = 3) = \frac{1}{2}$. Furthermore, the function $\mathcal{F}_Y$ is omitted from $T_{Y=1}$, and thus the arrow from X to Y has been omitted from the causal graph.

Given two languages $\mathcal{L}, \mathcal{L}'$ of signature σ, whose semantics is defined over causal multiteams, and formulae $\varphi \in \mathcal{L}$ and $\varphi' \in \mathcal{L}'$, we write $\varphi \equiv_\sigma \varphi'$ if $T \models \varphi \Leftrightarrow T \models \varphi'$ holds for all causal multiteams T of signature σ. We omit the index σ if it is clear from the context. Similarly, we may write $\mathcal{L}_\sigma$ to emphasise that the signature of $\mathcal{L}$ is σ.

We write $\mathcal{L} \leq \mathcal{L}'$ if for every $\varphi \in \mathcal{L}$ there is $\varphi' \in \mathcal{L}'$ with $\varphi \equiv \varphi'$. We write $\mathcal{L} < \mathcal{L}'$ if $\mathcal{L} \leq \mathcal{L}'$ but $\mathcal{L}' \nleq \mathcal{L}$. Finally, we write $\mathcal{L} \equiv \mathcal{L}'$ if $\mathcal{L} \leq \mathcal{L}'$ and $\mathcal{L}' \leq \mathcal{L}$. $\mathcal{K}^\sigma_\varphi$ is the set of all causal multiteams of signature σ that satisfy φ. $\mathcal{K}^\sigma_\varphi$ will be (with the exception of contradictory formulae) a countably infinite set.

A class $\mathcal{K}$ of causal multiteams is **definable** in $\mathcal{L}_\sigma$ if $\mathcal{K} = \mathcal{K}^\sigma_\varphi$ for some $\varphi \in \mathcal{L}_\sigma$.

A class $\mathcal{K}$ is **flat** if $(T^-, \mathcal{F}) \in \mathcal{K}$ iff $(\{s\}, \mathcal{F}) \in \mathcal{K}$ for every $s \in T^-$. A class $\mathcal{K}$ of causal multiteams of signature σ has the **empty multiteam property**, if $\mathcal{K}$ includes all empty causal multiteams of signature σ (we say that a causal multiteam $(T^-, \mathcal{F})$ is **empty** if the multiteam T^- is). A σ-formula φ has one of the above (or to be defined) properties, if $\mathcal{K}^\sigma_\varphi$ has it. A language $\mathcal{L}$ is flat (resp. has the empty team property), if every $\varphi \in \mathcal{L}$ is flat (resp. has the empty team property). In general, we say that $\mathcal{L}$ has a certain property if and only if each $\varphi \in \mathcal{L}$ has it.

The language $\mathcal{CO}$, introduced in [3], is defined by the following BNF grammar:

$$\alpha ::= Y = y \mid Y \neq y \mid \alpha \wedge \alpha \mid \alpha \vee \alpha \mid \alpha \supset \alpha \mid \mathbf{X} = \mathbf{x} \mathrel{\Box\!\!\rightarrow} \alpha,$$

where $\mathbf{X} \cup \{Y\} \subseteq \mathrm{Dom}$, $y \in \mathrm{Ran}(Y)$, and $\mathbf{x} \in \mathrm{Ran}(\mathbf{X})$. It is a language for the description of facts. We will later introduce extensions that allow us to talk about the probabilities of the facts that are expressible in $\mathcal{CO}$. Formulae of the forms $Y = y$ and $Y \neq y$ are **literals**. Semantics for $\mathcal{CO}$ is given by the following clauses:

$T \models Y = y$	iff	$s(Y) = y$ for all $s \in T^-$.
$T \models Y \neq y$	iff	$s(Y) \neq y$ for all $s \in T^-$.
$T \models \alpha \wedge \beta$	iff	$T \models \alpha$ and $T \models \beta$.
$T \models \alpha \vee \beta$	iff	there are $T_1, T_2 \leq T$ s.t. $T_1^- \cup T_2^- = T^-$, $T_1^- \cap T_2^- = \emptyset, T_1 \models \alpha$ and $T_2 \models \beta$.
$T \models \alpha \supset \beta$	iff	$T^\alpha \models \beta$.
$T \models \mathbf{X} = \mathbf{x} \mathrel{\Box\!\!\rightarrow} \beta$	iff	$T_{\mathbf{X}=\mathbf{x}} \models \beta$ or $\mathbf{X} = \mathbf{x}$ is inconsistent.

where T^α is defined simultaneously with the clauses, as previously explained.

The intuitive readings of the conditional formulas $\alpha \supset \beta$ and $\mathbf{X} = \mathbf{x} \mathrel{\Box\!\!\rightarrow} \beta$ are, respectively, "After observing (or learning) α, we know that β holds" and "After setting $\mathbf{X}$ to $\mathbf{x}$, we know that β holds". Some of the semantic clauses for the other connectives may look unusual to a reader unaccustomed to team semantics, but they are natural lifts of the usual Tarskian clauses from a setting in which formulas are evaluated on single assignments to a setting where they are evaluated on a multiplicity of assignments (for an overview of team semantics, the reader may consult e.g. [12]). As an example, the clause for a disjunction $\alpha \vee \beta$ is just stating that each assignment in T satisfies either α or β. It says so by saying that T can be split into two parts, one containing assignments that satisfy α and one containing assignments that satisfy β. This reading of the clauses is made possible by the fact that language $\mathcal{CO}$ is flat. The proof of the following result is similar to that of the analogous result for causal teams [3, Thm. 2.10].

▶ **Theorem 4.** *$\mathcal{CO}_\sigma$ is flat and therefore has the empty multiteam property.*

In a sense, flatness tells us that $\mathcal{CO}$ behaves as a classical language. The probabilistic languages that we shall consider later will not be flat; probabilistic statements are meaningful at the level of teams but not at the level of the single assignments.

We also remark that in [3] the operator $\vee$ was defined without insisting that $T_1^- \cap T_2^- = \emptyset$. This was done since the paper considered set-based semantics. As our semantics is based on multisets, the appropriate definition of $\vee$ uses a union that is sensitive to multiplicities (i.e. disjoint union). Theorem 4 entails that this distinction is irrelevant for $\mathcal{CO}$, but it will have an impact in forthcoming works that apply $\vee$ to formulae φ that do not have the following property called *downward closure*: if $T \models \varphi$ and $S \leq T$, then $S \models \varphi$.

If we pick a variable X in the signature and a value $x \in Ran(X)$, we can abbreviate the formulae $X = x \vee X \neq x$ and $X = x \wedge X \neq x$ as $\top$, resp. $\bot$ (the former is a valid formula because it just says that the multiteam can be split in two parts, the assignments where X takes value x and those where it does not). The so-called **dual negation** of a formula α, $T \models \alpha^d$ iff $(\{s\}, \mathcal{F}) \not\models \alpha$ for all $s \in T^-$, is then definable in $\mathcal{CO}$ as $\alpha \supset \bot$.

Next, we introduce a language with **probabilistic atoms** $\Pr(\alpha) \geq \epsilon$, $\Pr(\alpha) > \epsilon$, $\Pr(\alpha) \geq \Pr(\beta)$, $\Pr(\alpha) > \Pr(\beta)$, where $\alpha, \beta \in \mathcal{CO}$ and $\epsilon \in [0,1] \cap \mathbb{Q}$. The first two are called **evaluation atoms**, and the latter two **comparison atoms**. Probabilistic atoms together with literals of $\mathcal{CO}$ are called **atomic formulae**. The probabilistic language $\mathcal{PCO}$ is then given by the following grammar:

$$\varphi ::= \eta \mid \varphi \wedge \varphi \mid \varphi \sqcup \varphi \mid \alpha \supset \varphi \mid \mathbf{X} = \mathbf{x} \mathrel{\Box\!\!\rightarrow} \varphi,$$

where $\mathbf{X} \subseteq \mathrm{Dom}$, $\mathbf{x} \in \mathrm{Ran}(\mathbf{X})$, η is an atomic formula, and α is a $\mathcal{CO}$ formula. Note that the antecedents of $\supset$ and the arguments of probability operators are $\mathcal{CO}$ formulae. The semantic clauses for the additional operators are given below:

$T \models \psi \sqcup \chi$	iff	$T \models \psi$ or $T \models \chi$
$T \models \Pr(\alpha) \geq \epsilon$	iff	$T^- = \emptyset$ or $P_T(\alpha) \geq \epsilon$
$T \models \Pr(\alpha) > \epsilon$	iff	$T^- = \emptyset$ or $P_T(\alpha) > \epsilon$
$T \models \Pr(\alpha) \geq \Pr(\beta)$	iff	$T^- = \emptyset$ or $P_T(\alpha) \geq P_T(\beta)$
$T \models \Pr(\alpha) > \Pr(\beta)$	iff	$T^- = \emptyset$ or $P_T(\alpha) > P_T(\beta)$,

where $P_T(\alpha)$ is a shorthand for $P_{T^-}((T^\alpha)^-)$.[2] The language $\mathcal{PCO}$ still has the empty team property but it is not flat. The definability of the dual negation in $\mathcal{CO}$ allows us to introduce many useful abbreviations:

$$\Pr(\alpha) \leq \epsilon := \Pr(\alpha^d) \geq 1 - \epsilon \qquad \Pr(\alpha) = \epsilon := \Pr(\alpha) \geq \epsilon \wedge \Pr(\alpha) \leq \epsilon$$
$$\Pr(\alpha) < \epsilon := \Pr(\alpha^d) > 1 - \epsilon \qquad \Pr(\alpha) \neq \epsilon := \Pr(\alpha) > \epsilon \sqcup \Pr(\alpha) < \epsilon$$

We will see in Section 4 that the $\supset$ operator enables us to express some statements involving conditional probabilities.

We consider the following syntactic fragments of $\mathcal{PCO}$, which preserve the syntactic restrictions yielded by its two level syntax – that the antecedents of $\supset$ and the arguments of Pr are always $\mathcal{CO}$ formulae. $\mathcal{P}$ is the fragment without $\supset$ and $\Box\!\!\rightarrow$. $\mathcal{P}^-$ is the fragment of $\mathcal{P}$ without comparison atoms. $\mathcal{P}(\Box\!\!\rightarrow)$ and $\mathcal{P}(\supset)$ are fragments of $\mathcal{PCO}$ without $\supset$ and $\Box\!\!\rightarrow$, respectively. Finally, $\mathcal{PCO}^\omega$ is the extension of $\mathcal{PCO}$ with countable disjunctions of the form $\bigsqcup_{i \in I} \psi_i$, where the ψ_i are $\mathcal{PCO}$ formulae.

▶ **Example 5.** Let $T = (T^-, \mathcal{F})$ be a causal multiteam over variables GroundSpeed, DescentAngle, StructuralIntegrity, SafeLanding depicting data related to landing an Airbus A350-900 aircraft. The first three variables are numerical, while the last is Boolean. The structural equation $F_{\mathrm{SL}}(\mathrm{GS}, \mathrm{DA}, \mathrm{SI})$ outputs a Boolean value "true" when a plane of given structural integrity is expected to make a safe landing at a given speed and angle. The formula "$\mathrm{SI} \neq 0 \supset [(\mathrm{GS} = 300 \wedge \mathrm{DA} = 4) \mathrel{\Box\!\!\rightarrow} \Pr(\mathrm{SL} = \mathrm{false}) < 0.01]$" expresses that the

[2] We remark that in $\mathcal{PCO}$ (but not in $\mathcal{CO}$!) is it also possible to define, inductively, an operator that behaves as classical negation on nonempty causal multiteams (*weak contradictory negation*). Details can be found in [6]; we will not use it here.

probability of landing failure is less than 1% when setting a landing speed of $300km/h$ and descent angle of 4 degrees, conditional on the plane not being grounded due to structural condition ($\mathrm{SI} = 0$).

Since we can assume that SI is exogenous (the assessment of structural integrity is not affected by the speed and angle set during the flight), this statement can be equivalently written as "($\mathrm{GS} = 300 \land \mathrm{DA} = 4) \;\Box\!\!\rightarrow (\mathrm{SI} \neq 0 \supset \Pr(\mathrm{SL} = \mathrm{false}) < 0.01)$". This would not be legitimate if SI was causally affected by GS or DA; the operators $\Box\!\!\rightarrow$ and $\supset$ do not in general commute with each other.

3 Expressive power of fragments of $\mathcal{PCO}$

We start by rephrasing the known characterizations from the literature. A number of results appear in the literature (e.g. in [7]) that characterize causal languages in the context of causal *team* semantics. A causal team (of signature σ) is, essentially, a pair $(T^-, \mathcal{F})$, where T^- is a team instead of a multiteam (i.e., a set of assignments on Dom instead of $Dom \cup \{Key\}$), satisfying the conditions given in Definition 1. Each causal multiteam can be seen as a causal team enriched with a probability distribution. This correspondence is expressed precisely as follows:

▶ **Definition 6.** *The **support** of a causal multiteam $T = (T^-, \mathcal{F})$ is the causal team* $\mathrm{Team}(T) = (\mathrm{Team}(T^-), \mathcal{F})$*, where* $\mathrm{Team}(T^-) := \{s_{\upharpoonright Dom} \mid s \in T^-\}$.

It is immediate to see that a language without probabilistic features (such as $\mathcal{CO}$) cannot tell apart two causal multiteams that have the same support. From this, it is straightforward but tedious (see the extended version of the paper, [5]) to show that the characterization of $\mathcal{CO}$ given in [7, Theorem 4.4] in terms of causal teams holds unchanged over causal multiteams:

▶ **Theorem 7** (Characterization of $\mathcal{CO}$)**.** *Let σ be a finite signature, and $\mathcal{K}$ a class of causal multiteams of signature σ. Then $\mathcal{K}$ is definable by a $\mathcal{CO}_\sigma$ formula (resp. a set of $\mathcal{CO}_\sigma$ formulae) if and only if $\mathcal{K}$ is flat.*

$\mathcal{PCO}$ is a purely probabilistic language; it cannot tell apart multiteams representing the same distribution. Given an assignment t and a causal team $T = (T^-, \mathcal{F})$, we write $\#(t, T)$ for the number of copies of t in T^- and (provided T is nonempty) $\epsilon_t^T := \frac{\#(t,T)}{|T^-|}$ for the probability of t in T. Two causal teams $S = (S^-, \mathcal{F})$ and $T = (T^-, \mathcal{G})$ are **rescalings** of each other ($S \sim T$) if $\mathcal{F} = \mathcal{G}$ and either $S^- = T^- = \emptyset$ or $\epsilon_t^T = \epsilon_t^S$ for each assignment t. A class $\mathcal{K}$ of causal multiteams of signature σ is **closed under rescaling** if, whenever $S \in \mathcal{K}$ and $S \sim T$, also $T \in \mathcal{K}$. An ideal language for purely probabilistic reasoning should be characterized just by this condition. It turns out that $\mathcal{PCO}$ is not expressive enough for the task, however its extension with countable global disjunctions $\mathcal{PCO}^\omega$ is.

▶ **Theorem 8** ([4])**.** *A* nonempty *class $\mathcal{K}$ of multiteams of signature σ is definable in $\mathcal{PCO}^\omega_\sigma$ iff $\mathcal{K}$ has the empty multiteam property and is closed under rescaling.*

The key to the proof is the fact that for any causal multiteam $(T^-, \mathcal{F})$ one can write $\mathcal{PCO}$-formulae Θ_{T^-} and $\Phi^{\mathcal{F}}$ that characterize the properties of having team component T^- (up to rescaling) and function component $\mathcal{F}$, respectively. A set $\mathcal{K}$ of causal multiteams is then defined by the formula $\bigsqcup_{(T^-,\mathcal{F})\in\mathcal{K}}(\Theta_{T^-} \land \Phi^{\mathcal{F}})$. Since $\mathcal{K}$ can be countably infinite, the proof crucially depends on the use of infinitary disjunctions and gives us no hints on how to obtain a finitary logic with such expressivity. Actually, a counting argument given in [4] shows that such a language must be uncountable, and thus that $\mathcal{PCO} < \mathcal{PCO}^\omega$. Our characterization of the expressivity of $\mathcal{PCO}$ will provide an alternative proof for the strict inclusion.

In order to characterize the expressivity of $\mathcal{PCO}$ and its fragments, we need to introduce some classes of linear equations and closure properties of classes of causal multiteams. For the latter, we have already seen closure under rescaling and the empty multiteam property. A class $\mathcal{K}$ of causal multiteams of signature σ is **closed under change of laws** if, whenever $(T^-, \mathcal{F}) \in \mathcal{K}$ and $\mathcal{G}$ is a system of functions of signature σ such that $(T^-, \mathcal{G})$ satisfies the compatibility constraint (point 3. of definition 1), then $(T^-, \mathcal{G}) \in \mathcal{K}$.

It is self-evident that the logics without $\boxright$ are closed under change of laws, while the logics with $\boxright$ are not. Thus, the following hold.

▶ **Lemma 9.** *$\mathcal{P}^-$, $\mathcal{P}$, and $\mathcal{P}(\supset)$ are closed under change of laws. $\mathcal{PCO}$, $\mathcal{P}(\boxright)$, and $\mathcal{CO}$ are not closed under change of laws.*

▶ **Corollary 10.** *$\mathcal{P} < \mathcal{P}(\boxright)$, $\mathcal{P}(\supset) < \mathcal{PCO}$, and $\mathcal{P}(\boxright) \not\leq \mathcal{P}(\supset)$.*

3.1 Monic and signed monic probability sets: $\mathcal{P}^-$, $\mathcal{P}$, and $\mathcal{P}(\boxright)$

We characterize the expressivity of fragments of $\mathcal{PCO}$ by investigating the families of subsets of $\mathbb{Q}^n$ that are definable in the logics. For a given signature σ, we fix an enumeration $s_1, \ldots, s_n$ of the assignments of $\mathbb{B}_\sigma$; every *nonempty* causal multiteam T can then be associated with a **probability vector** $\overline{\mathrm{p}}_T = (\epsilon^T_{s_1}, \ldots, \epsilon^T_{s_n}) \in \mathbb{Q}^n$. Similarly, a class $\mathcal{K}$ of causal multiteams of signature σ has an associated **probability set** $\overline{\mathrm{P}}_\mathcal{K} = \{\overline{\mathrm{p}}_T \mid T \in \mathcal{K}, T \text{ nonempty }\}$. Note that $\overline{\mathrm{p}}_T$ and $\overline{\mathrm{P}}_\mathcal{K}$ are, respectively, a point and a subset of the **standard $n-1$-simplex** Δ^{n-1} (i.e. the set of points of $[0,1]^n \cap \mathbb{Q}^n$ that satisfy the equation $\epsilon_{s_1} + \cdots + \epsilon_{s_n} = 1$), respectively. To each formula φ, we can associate a probability set $\overline{\mathrm{P}}_\varphi := \overline{\mathrm{P}}_{\mathcal{K}_\varphi}$. Note that if S, T are causal multiteams of the same signature and same function component, such that $\overline{\mathrm{p}}_S = \overline{\mathrm{p}}_T$, then S is a rescaling of T. Similarly, a class $\mathcal{K}$ of causal multiteams of signature σ that is closed under change of laws and rescaling is the largest class of causal multiteams of signature σ having probability set $\overline{\mathrm{P}}_\mathcal{K}$.

A **linear inequality** is an expression of the form $a_1\epsilon_1 + \cdots + a_n\epsilon_n \rhd b$, where $\rhd \in \{\geq, \leq, >, <\}$, $a_1, \ldots, a_n, b \in \mathbb{Q}$, and $\epsilon_1, \ldots \epsilon_n$ are variables (in the usual algebraic sense). A linear inequality is **signed monic** if each of the a_i is in $\{0, 1, -1\}$. It is **monic** if each of the a_i is in $\{0, 1\}$. A probability set $\overline{\mathrm{P}}$ is **(signed) monic** if it is a finite union of subsets of Δ^{n-1} defined by finite systems of (signed) monic inequalities. A class $\mathcal{K}$ of causal multiteams of a fixed signature is **(signed) monic** if $\overline{\mathrm{P}}_\mathcal{K}$ is a (signed) monic probability set.

We will show that being monic and closed under change of laws and rescaling characterizes expressibility in $\mathcal{P}^-$, whereas being signed monic and closed under change of laws and rescaling characterizes expressibility in $\mathcal{P}$. The full proofs of the following theorem and the subsequent lemma can be found in the extended version of the paper ([5]). A crucial role in the proofs is played by the fact that there are only finitely many assignments of signature σ (say $s_1, \ldots, s_n$) and that we can describe each such assignment s_i with a formula $\hat{\alpha}_i := \mathbf{W} = s_i(\mathbf{W})$, where $\mathbf{W}$ lists all the variables in *Dom*.

▶ **Theorem 11.** *A class $\mathcal{K}$ of multiteams of signature σ is definable in $\mathcal{P}^-$ if and only if $\mathcal{K}$ is monic, has the empty multiteam property, and is closed under change of laws and rescaling. $\mathcal{K}$ is definable in $\mathcal{P}$ if and only if $\mathcal{K}$ is signed monic, has the empty multiteam property, and closed under change of laws and rescaling.*

Proof (sketch). The fact that $\mathcal{P}^-$ and $\mathcal{P}$ have the empty multiteam property and are closed under rescaling follows from Theorem 8. Since $T \models \Pr(\alpha) \rhd \epsilon$ (resp. $T \models \Pr(\alpha) \rhd \Pr(\beta)$) iff the monic inequality $\sum_{s \in \mathrm{Team}((T^\alpha)^-)} \epsilon^T_s \rhd \epsilon$ (resp. the signed monic inequality $\sum_{s \in \mathrm{Team}((T^\alpha)^-)} \epsilon^T_s + \sum_{s \in \mathrm{Team}((T^\beta)^-)} (-1) \cdot \epsilon^T_s \rhd 0$) holds, we obtain that $\mathcal{P}^-$ (resp. $\mathcal{P}$) is monic (resp. signed monic) by induction on the syntax of formulae.

For $\mathcal{P}^-$, the right-to-left entailment is proved via a direct translation from finite unions of finite systems of signed monic inequalities into $\mathcal{P}^-$ formulae. The union of systems, which defines the probability set of $\mathcal{K}$, is expressed via a formula of the form $\varphi := \bigsqcup_{1\leq j\leq m} \bigwedge_{i\in I_j} \psi_i$, where each $\psi_i := \Pr(\bigvee_{s_k\in\mathbb{B}_\sigma|a^i_k=1} \hat{\alpha}_k) \triangleleft b^i$ expresses an inequality of the form $a^i_1\epsilon_1 + \cdots + a^i_n\epsilon_n \triangleleft b^i$ (it is easy to see that b^i can always be assumed to be in $[0,1]\cap\mathbb{Q}$). The fact that $\mathcal{K}$ has the empty team property and closure under rescaling guarantees that $\mathcal{K}$ is "maximal", i.e. it contains *all* the causal multiteams whose probability set is defined by this system; thus $\mathcal{K} = \mathcal{K}^\sigma_\varphi$.

In contrast, for $\mathcal{P}$, we do not construct any general direct translations of signed monic inequalities into $\mathcal{P}$ formulas. However, the signed monic inequalities with constant coefficient 0, say $\sum_{i\in I}\epsilon_i - \sum_{j\in J}\epsilon_j \triangleleft 0$ with $I\cap J=\emptyset$, are easily translated as $\Pr(\bigwedge_{i\in I}\hat{\alpha}_i) \triangleleft \Pr(\bigwedge_{j\in J}\hat{\alpha}_j)$. In order to extend the argument to inequalities with nonzero constant coefficient, we first use the simplex equality $\epsilon_1+\cdots+\epsilon_n = 1$ in order to show that we can assume that such inequality e has at least one null variable coefficient – say, it is of the form $a_1\epsilon_1 + \cdots + a_{n-1}\epsilon_{n-1} \triangleleft b$ (one must be careful to ensure that in this simplified inequality we still have $a_i \in \{0,1,-1\}$ and $b\in[0,1]\cap\mathbb{Q}$). But now e is equivalent to a system of three inequalities:

$$\begin{cases} a_1\epsilon_1 + \cdots + a_{n-1}\epsilon_{n-1} - \epsilon_n \triangleleft 0 \\ \epsilon_n \leq b \\ \epsilon_n \geq b \end{cases}$$

the first of which is expressible in $\mathcal{P}$ (since its constant coefficient is zero), while the second and third are even expressible in $\mathcal{P}^-$. ◂

It is not immediate to see whether $\mathcal{P}^- \leq \mathcal{P}$ is strict. However, by analyzing the geometry of Δ^{n-1} we are able show that there are signed monic classes of causal multiteams that are not monic. The following lemma establishes that not all signed monic probability sets can be captured by monic inequalities (more specifically, that this happens for sets defined by a single signed monic inequality). Together with the previous theorem this implies that $\mathcal{P}^- < \mathcal{P}$.

▸ **Lemma 12.** *Consider a nonempty probability set $\overline{\mathrm{P}} \subset \Delta^{n-1}$ which is defined by an inequality $a_1\epsilon_1 + \cdots + a_n\epsilon_n \leq b$, where there are indexes i, j such that a_i is 1 and a_j is -1, and b is a rational number in $[0,1]$. Then $\overline{\mathrm{P}}$ is not a monic probability set.*

Proof (sketch). The projection of the set described in the statement on the (i,j)-plane has as its frontier a line that is perpendicular to the segment of extremes $(0,1),(1,0)$. On the other hand, monic equalities describe, in this projection, only lines that are either parallel to this segment or parallel to one of the axis. ◂

Next we turn to characterize the expressivity of $\mathcal{P}(\boxright)$. First note that while $\mathcal{P}(\boxright)$ is in general more expressive than $\mathcal{P}$ (Corollary 10), if we restrict attention to causal multiteams with a fixed function component, all occurrences of $\boxright$ can be eliminated from $\mathcal{P}(\boxright)$ formulae (or even $\mathcal{PCO}$ formulae). The following result is proven in the extended version of the paper ([5]).

▸ **Proposition 13.** *Let $\varphi \in \mathcal{P}(\boxright)_\sigma$ (resp. $\mathcal{PCO}_\sigma$), and $\mathcal{F}$ a function component of signature σ. Then there is a formula $\varphi^\mathcal{F} \in \mathcal{P}_\sigma$ (resp. $\mathcal{P}(\supset)_\sigma$) such that, for every causal multiteam T of signature σ and function component $\mathcal{F}$, $T \models \varphi \Leftrightarrow T \models \varphi^\mathcal{F}$.*

Proof (sketch). Write α_s for the formula $\mathbf{W} = s(\mathbf{W})$. First, for every subformulae of φ of the form $\beta \supset \psi$, replace β with $\bigvee_{(\{s\},\mathcal{F})\models\beta} \alpha_s$ (this removes occurrences of $\boxright$ from antecedents of $\supset$). Next, we use the fact that $\boxright$ distributes over $\wedge, \sqcup, \supset$ to guarantee that the consequents of $\boxright$ are atoms. The atoms can be assumed to be probabilistic (since $X = x \equiv \Pr(X = x) \geq 1$, and similarly for $X \neq x$). Then, we use the equivalences

$$\mathbf{X} = \mathbf{x} \boxright \Pr(\alpha) \lhd \epsilon \equiv \Pr(\mathbf{X} = \mathbf{x} \boxright \alpha) \lhd \epsilon$$

$$\mathbf{X} = \mathbf{x} \boxright \Pr(\alpha) \lhd \Pr(\beta) \equiv \Pr(\mathbf{X} = \mathbf{x} \boxright \alpha) \lhd \Pr(\mathbf{X} = \mathbf{x} \boxright \beta)$$

to ensure that all the occurences of $\boxright$ are inside arguments of Pr. Finally, we replace each subformula of the form $\Pr(\alpha) \lhd \epsilon$ with $\Pr(\bigvee_{(\{s\},\mathcal{F})\models\alpha} \alpha_s) \lhd \epsilon$, and similarly for comparison atoms. ◀

Notice that, for any fixed finite signature σ, there is only a finite number of distinct function components. We denote the set they form as $\mathbb{F}_\sigma$.

▶ **Theorem 14.** *Let $\mathcal{K}$ be a class of causal multiteams of signature σ. $\mathcal{K}$ is definable by a $\mathcal{P}(\boxright)_\sigma$ formula if and only if 1) $\mathcal{K}$ has the empty multiteam property, 2) $\mathcal{K}$ is closed under rescaling, and 3) $\mathcal{K} = \bigcup_{\mathcal{F}\in\mathbb{F}_\sigma} \mathcal{K}^{\mathcal{F}}$, where each $\mathcal{K}^{\mathcal{F}}$ is a signed monic set of causal multiteams of function component $\mathcal{F}$.*

Proof. We have already mentioned that there is a $\mathcal{PCO}$ formula $\Phi^{\mathcal{F}}$ characterizing the property of having function component $\mathcal{F}$. We can obtain an equivalent formula (call it $\Psi^{\mathcal{F}}$) in $\mathcal{P}(\boxright)$ by replacing each subformula of $\Phi^{\mathcal{F}}$ of the form $\alpha \supset \beta$ with $\Pr(\alpha^d \vee \beta) = 1$ (the trick works because no consequent of $\supset$ in $\Phi^{\mathcal{F}}$ contains probabilistic atoms).

$\Rightarrow$) Suppose $\mathcal{K} = \mathcal{K}_\varphi$, where $\varphi \in \mathcal{P}(\boxright)_\sigma$. Now define, for each $\mathcal{F} \in \mathbb{F}_\sigma$, $\mathcal{K}^{\mathcal{F}} := \mathcal{K}_{\varphi\wedge\Psi^{\mathcal{F}}}$, where $\Psi^{\mathcal{F}}$ is as described above. Clearly $\varphi \equiv \bigsqcup_{\mathcal{F}\in\mathbb{F}_\sigma}(\varphi \wedge \Psi^{\mathcal{F}})$, so $\mathcal{K}_\varphi = \bigcup_{\mathcal{F}\in\mathbb{F}_\sigma} \mathcal{K}^{\mathcal{F}}$.

Now, by Theorem 8, $\mathcal{K}_\varphi$ is closed under rescaling and has the empty multiteam property. Next, observe that, by Proposition 13, for every $\mathcal{F} \in \mathbb{F}_\sigma$ there is a formula of $\mathcal{P}_\sigma$, call it $\varphi^{\mathcal{F}}$, which is satisfied by the same causal multiteams of function component $\mathcal{F}$ as $\varphi \wedge \Psi^{\mathcal{F}}$ is. In other words, $\mathcal{K}^{\mathcal{F}}$ is the restriction of $\mathcal{K}_{\varphi^{\mathcal{F}}}$ to causal multiteams of function component $\mathcal{F}$. Thus, since $\mathcal{K}_{\varphi^{\mathcal{F}}}$ is closed under change of laws (Lemma 9), we have $\overline{\mathrm{P}}_{\mathcal{K}^{\mathcal{F}}} = \overline{\mathrm{P}}_{\mathcal{K}_{\varphi^{\mathcal{F}}}}$. Now $\mathcal{K}_{\varphi^{\mathcal{F}}}$ is signed monic (Theorem 11), and thus by $\overline{\mathrm{P}}_{\mathcal{K}^{\mathcal{F}}} = \overline{\mathrm{P}}_{\mathcal{K}_{\varphi^{\mathcal{F}}}}$ we conclude that also $\mathcal{K}^{\mathcal{F}}$ is signed monic.

$\Leftarrow$) Suppose $\mathcal{K}$ is closed under rescaling, has the empty multiteam property and $\mathcal{K} = \bigcup_{\mathcal{F}\in\mathbb{F}_\sigma} \mathcal{K}^{\mathcal{F}}$ for some sets $\mathcal{K}^{\mathcal{F}}$ as in the statement. Write $\hat{\mathcal{K}}^{\mathcal{F}}$ for the set of all causal multiteams of signature σ whose team component appears in $\mathcal{K}^{\mathcal{F}}$. It is straightforward then that also $\hat{\mathcal{K}}^{\mathcal{F}}$ is closed under rescaling, has the empty multiteam property and is signed monic; however, $\hat{\mathcal{K}}^{\mathcal{F}}$ is also, by definition, closed under change of laws. Thus, by Theorem 11, there is a $\mathcal{P}$ formula $\varphi^{\mathcal{F}}$ such that $\hat{\mathcal{K}}^{\mathcal{F}} = \mathcal{K}_{\varphi^{\mathcal{F}}}$. Note that, $\mathcal{K}^{\mathcal{F}}$ is the set of all causal multiteams of $\mathcal{K}_{\varphi^{\mathcal{F}}}$ that have function component $\mathcal{F}$. Thus $\mathcal{K}^{\mathcal{F}} = \mathcal{K}_{\varphi^{\mathcal{F}}\wedge\Psi^{\mathcal{F}}}$. Thus $\mathcal{K}$ is defined by the $\mathcal{P}(\boxright)_\sigma$ formula $\bigsqcup_{\mathcal{F}\in\mathbb{F}_\sigma}(\varphi^{\mathcal{F}} \wedge \Psi^{\mathcal{F}})$. ◀

Note that the sets $\mathcal{K}^{\mathcal{F}}$ in the statement of the theorem are themselves closed under rescaling if $\mathcal{K}$ is. This immediately follows from the fact that any two causal multiteams $(T, \mathcal{F}), (S, \mathcal{G})$ with $\mathcal{F} \neq \mathcal{G}$ are *not* rescalings of each other.

3.2 Signed binary probability sets: $\mathcal{P}(\supset)$ and $\mathcal{PCO}$

A subset $\overline{\mathrm{P}}$ of Δ^{n-1} is **signed binary** if it is a finite union of sets defined by finite systems of inequalities of the form

$$c^- \sum_{i\in I} \epsilon_i + c^+ \sum_{j\in J} \epsilon_j \lhd b$$

where $I \cap J = \emptyset$, $c^-, c^+ \in \mathbb{Z}$, $c^- \leq 0$, $c^+ \geq 0$, $b \in \mathbb{Q}$. Likewise, a class $\mathcal{K}$ of causal multiteams of signature σ is signed binary if $\overline{\mathrm{P}}_{\mathcal{K}}$ is.

▶ **Lemma 15.** *Every formula $\varphi \in \mathcal{P}(\supset)$ is signed binary.*

Proof. The proof proceeds by induction on φ. We only discuss the most difficult case, when φ is of the form $\alpha \supset \psi$. Write $\lhd$ for any symbol in $\{\leq, \geq, <, >\}$. Using the distributivity of $\supset$ over $\wedge$ and $\vee$, and the equivalences $X = x \equiv \Pr(X = x) = 1, X \neq x \equiv \Pr(X \neq x) = 1, \mathbf{X} = \mathbf{x} \boxright \Pr(\alpha) \lhd \epsilon \equiv \Pr(\mathbf{X} = \mathbf{x} \boxright \alpha) \lhd \epsilon$ and $\mathbf{X} = \mathbf{x} \boxright \Pr(\alpha) \lhd \Pr(\beta) \equiv \Pr(\mathbf{X} = \mathbf{x} \boxright \alpha) \lhd \Pr(\mathbf{X} = \mathbf{x} \boxright \beta)$, we can assume ψ to be a probabilistic atom. Hence we have two cases.

1) Assume ψ is $\Pr(\beta) \lhd b$. Now $T = (T^-, \mathcal{F}) \in \mathcal{K}_\varphi$ iff either $P_T(\alpha) \leq 0$ or $P_T(\beta \mid \alpha) \lhd b$. The latter is equivalent to $P_T(\beta \wedge \alpha) \lhd b \cdot P_T(\alpha)$, which can be rewritten as

$$\sum_{\substack{s \in \mathbb{B}_\sigma \\ \{s\} \models \beta \wedge \alpha}} \epsilon_s^T \lhd b \cdot \sum_{\substack{s \in \mathbb{B}_\sigma \\ \{s\} \models \alpha}} \epsilon_s^T$$

where we write e.g. $\{s\} \models \alpha$ as a shorthand for $(\{s\}, \mathcal{F}) \models \alpha$.

The above can be rewritten as

$$\sum_{\substack{s \in \mathbb{B}_\sigma \\ \{s\} \models \beta \wedge \alpha}} \epsilon_s^T \lhd b \cdot \Big(\sum_{\substack{s \in \mathbb{B}_\sigma \\ \{s\} \models \beta \wedge \alpha}} \epsilon_s^T + \sum_{\substack{s \in \mathbb{B}_\sigma \\ \{s\} \models \neg\beta \wedge \alpha}} \epsilon_s^T \Big)$$

which again is equivalent to

$$(1 - b) \cdot \sum_{\substack{s \in \mathbb{B}_\sigma \\ \{s\} \models \beta \wedge \alpha}} \epsilon_s^T + (-b) \cdot \sum_{\substack{s \in \mathbb{B}_\sigma \\ \{s\} \models \neg\beta \wedge \alpha}} \epsilon_s^T \lhd 0. \tag{1}$$

Now, since $b \in [0, 1]$, we have $1 - b \geq 0$ and $-b \leq 0$. Then, by multiplying both sides of (1) by a common denominator of $1 - b$ and $-b$, we obtain a signed binary inequality.

On the other hand, the inequality $P_T(\alpha) \leq 0$ can be rewritten as $\sum_{\{s\} \models \alpha} \epsilon_s \leq 0$. Thus $\overline{\mathrm{P}}_\varphi$ is the union of two sets defined by signed binary inequalities.

2) Assume ψ is $\Pr(\beta) \lhd \Pr(\gamma)$. Now $T \in \mathcal{K}_\varphi$ iff either $P_T(\alpha) \leq 0$ or $P_T(\beta \mid \alpha) \lhd P_T(\gamma \mid \alpha)$. The proof then proceeds as in the previous case. ◀

▶ **Theorem 16.** *A class $\mathcal{K}$ of multiteams of signature σ is definable by a formula of $\mathcal{P}(\supset)$ if and only if $\mathcal{K}$ is signed binary, has the empty multiteam property and is closed under change of laws and rescaling.*

Proof (sketch). $\Rightarrow$) By Theorem 8, $\mathcal{K}$ is closed under rescaling. Closure under change of laws follows from Lemma 9. Lemma 15 shows that $\overline{\mathrm{P}}_\mathcal{K}$ is signed binary. The empty multiteam property is given by Theorem 4.

$\Leftarrow$) The proof strategy is analogous to that used for the characterization of $\mathcal{P}$ (in Theorem 11), although it involves more difficult calculations. We need to show that every constraint of the form

$$c^- \sum_{i \in I} \epsilon_i + c^+ \sum_{j \in J} \epsilon_j \lhd b$$

where $I \cap J = \emptyset$, $c^-, c^+ \in \mathbb{Z}$, $c^- \leq 0$, $c^+ \geq 0$, $b \in \mathbb{Q}$, can be expressed in $\mathcal{P}(\supset)$.

First of all, let us prove it in the special case when b is 0. Write d for $c^+ - c^-$. Notice that $-d \leq c^- \leq 0 \leq c^+ \leq d$. We can also assume that $d > 0$ (the case when $d = 0$ is covered by Theorem 11). Then $-\frac{c^-}{d}$ is a rational number in $[0, 1]$, and thus the following is a $\mathcal{P}(\supset)$ formula (where, as before, $\hat{\alpha}_j$ stands for $\mathbf{W} = s_j(\mathbf{W})$):

$$\Big(\bigvee_{k \in I \cup J} \hat{\alpha}_k \Big) \supset \Pr\Big(\bigvee_{j \in J} \hat{\alpha}_j\Big) \lhd -\frac{c^-}{d}.$$

Now we have

$$\begin{aligned}
& T \models \Big(\bigvee_{k \in I \cup J} \hat{\alpha}_k \Big) \supset \Pr(\bigvee_{j \in J} \hat{\alpha}_j) \lhd -\frac{c^-}{d} \\
\iff & P_T(\bigvee_{j \in J} \hat{\alpha}_j \mid \bigvee_{k \in I \cup J} \hat{\alpha}_k) \lhd -\frac{c^-}{d} \\
\iff & d \cdot P_T(\bigvee_{j \in J} \hat{\alpha}_j \wedge \bigvee_{k \in I \cup J} \hat{\alpha}_k) \lhd -c^- \cdot P_T(\bigvee_{k \in I \cup J} \hat{\alpha}_k) \\
\iff & d \cdot P_T(\bigvee_{j \in J} \hat{\alpha}_j) \lhd -c^- \cdot P_T(\bigvee_{k \in I \cup J} \hat{\alpha}_k) \\
\iff & d \sum_{j \in J} \epsilon_j^T \lhd -c^- \sum_{k \in I \cup J} \epsilon_k^T \\
\iff & c^- \sum_{i \in I} \epsilon_i^T + (d + c^-) \sum_{j \in J} \epsilon_j^T \lhd 0 \\
\iff & c^- \sum_{i \in I} \epsilon_i^T + c^+ \sum_{j \in J} \epsilon_j^T \lhd 0,
\end{aligned}$$

as required.

Now let us consider the case when $b \neq 0$. Suppose, first, that we have an inequality of the form $c^- \sum_{i \in I} \epsilon_i + c^+ \sum_{j \in J} \epsilon_j \lhd b$ that satisfies the additional constraint that $I \cup J = \{1, \ldots, n\}$, i.e. it contains all variables. We show that then it is equivalent to an inequality of the same form, but with coefficient 0 for at least one variable. Assuming that I is nonempty, let us pick a variable in I (that we may assume wlog to be ϵ_n). Thus the inequality can be rewritten as:

$$c^- \sum_{i \in I \setminus \{n\}} \epsilon_i + c^+ \sum_{j \in J} \epsilon_j + c^- \epsilon_n \lhd b.$$

Using the fact that, in Δ^{n-1}, $\epsilon_1 + \cdots + \epsilon_n = 1$, we can rewrite the inequality as

$$c^- \sum_{i \in I \setminus \{n\}} \epsilon_i + c^+ \sum_{j \in J} \epsilon_j + c^- - c^- \epsilon_1 - \cdots - c^- \epsilon_{n-1} \lhd b$$

i.e.,

$$\sum_{j \in J} (c^+ - c^-) \epsilon_j \lhd b - c^-,$$

which is of the correct form. In case I is empty, we can perform analogous transformations to eliminate a variable indexed in J.

Thus we can always assume that an inequality $c^- \sum_{i \in I} \epsilon_i + c^+ \sum_{j \in J} \epsilon_j \lhd b$ (as above) has coefficient 0 for ϵ_n. Let k be a positive integer such that $kb \in \mathbb{Z}$. Then, it is easy to see that our inequality is equivalent to the following system:

$$\begin{cases}
(kbc^-) \sum_{i \in I} \epsilon_i + (kbc^+) \sum_{j \in J} \epsilon_j + (kbc^-) \epsilon_n \lhd 0 \\
\epsilon_n \leq -\frac{b}{c^-} \\
\epsilon_n \geq -\frac{b}{c^-}
\end{cases}$$

The first of these inequalities is expressible with a $\mathcal{P}(\supset)$ formula by the discussion above. By theorem 11 the other two inequalities are expressed by $\mathcal{P}^-$ formulae. ◀

In order to prove that $\mathcal{P}(\supset)$ is strictly more expressive than $\mathcal{P}$, we can follow a similar strategy as for separating $\mathcal{P}$ and $\mathcal{P}^-$. In other words, we use Theorem 16 together with the fact that there are signed binary probability sets that are not signed monic, as established by the following lemma.

▶ **Lemma 17.** *Let $\overline{\mathrm{P}} \subset \Delta^{n-1}$ be a probability set defined by a single inequality $a_1\epsilon_1 + \cdots + a_n\epsilon_n \leq b$, where $0 \neq a_i, a_j \in \mathbb{Z}$ and $|a_i| \neq |a_j|$, for some indices i, j. Then $\overline{\mathrm{P}}$ is not signed monic.*

Proof (sketch). The proof is analogous to that of Lemma 12, using the fact that the projection on the (i, j)-plane of the figure described in the statement is neither parallel to any axis, nor parallel or orthogonal to the segment of extremes $(0, 1), (1, 0)$. ◀

Actually, the lemma immediately yields multiple separation results.

▶ **Proposition 18.** *1) $\mathcal{P} < \mathcal{P}(\supset)$, 2) $\mathcal{P}(\supset) \not\leq \mathcal{P}(\Box\!\!\rightarrow)$, 3) $\mathcal{P}(\Box\!\!\rightarrow) < \mathcal{PCO}$.*

We are finally ready to characterize the expressive power of $\mathcal{PCO}$; the proof is analogous to that of Theorem 14.

▶ **Theorem 19.** *Let $\mathcal{K}$ be a class of causal multiteams of signature σ. $\mathcal{K}$ is definable by a $\mathcal{PCO}_\sigma$ formula if and only if 1) it has the empty multiteam property, 2) it is closed under rescaling, and 3) $\mathcal{K} = \bigcup_{\mathcal{F}\in\mathbb{F}_\sigma} \mathcal{K}^{\mathcal{F}}$, where each $\mathcal{K}^{\mathcal{F}}$ is a signed binary set of causal multiteams of function component $\mathcal{F}$.*

By Theorem 8, $\mathcal{PCO}^\omega$ formulae may characterize arbitrary probability sets. By Theorem 19, instead, we know that the probability sets of $\mathcal{PCO}$ formulae are all definable in terms of linear inequalities. A strict inclusion of languages immediately follows. An alternative proof for this using a counting argument was given in [4].

▶ **Corollary 20.** *$\mathcal{PCO} < \mathcal{PCO}^\omega$.*

4 Definability of probabilistic and dependence atoms

Next we briefly explore the relationships of our logics and the probabilistic atoms studied in probabilistic and multiteam semantics. We consider the dependence atom by Väänänen [38], and marginal distribution identity and probabilistic independence atoms by Durand et al. [10].

The dependence atom $=(\mathbf{X};\mathbf{Y})$ expresses that the values of $\mathbf{X}$ functionally determine the values of $\mathbf{Y}$. Dependence atoms can be expressed already in $\mathcal{P}(\supset)$:

$$=(\mathbf{X};\mathbf{Y}) := \bigwedge_{\mathbf{x}\in\mathrm{Ran}(\mathbf{X})} \bigsqcup_{\mathbf{y}\in\mathrm{Ran}(\mathbf{Y})} \mathbf{X} = \mathbf{x} \supset \mathbf{Y} = \mathbf{y}$$

The marginal distribution identity atom $\mathbf{X} \approx \mathbf{Y}$ states that the marginal distributions induced by $\mathbf{X}$ and $\mathbf{Y}$ are identical. This can be defined in $\mathcal{P}$ by

$$\begin{aligned}\mathbf{X} \approx \mathbf{Y} := &\bigwedge_{\mathbf{x}\in\mathrm{Ran}(\mathbf{X})\cap\mathrm{Ran}(\mathbf{Y})} \Pr(\mathbf{X}=\mathbf{x}) = \Pr(\mathbf{Y}=\mathbf{x}) \land \\ &\bigwedge_{\mathbf{x}\in\mathrm{Ran}(\mathbf{X})\setminus\mathrm{Ran}(\mathbf{Y})} \Pr(\mathbf{X}=\mathbf{x}) = 0 \land \bigwedge_{\mathbf{y}\in\mathrm{Ran}(\mathbf{Y})\setminus\mathrm{Ran}(\mathbf{X})} \Pr(\mathbf{Y}=\mathbf{y}) = 0.\end{aligned}$$

The conditional probabilistic atoms inherit their semantics from probability theory:

$$\begin{aligned}T &\models \Pr(\alpha \mid \beta) \rhd \epsilon && \text{iff} && (T^\beta)^- = \emptyset \text{ or } P_{T^\beta}(\alpha) \rhd \epsilon.\\ T &\models \Pr(\alpha \mid \beta) \rhd \Pr(\gamma \mid \delta) && \text{iff} && (T^\beta)^- = \emptyset \text{ or } (T^\delta)^- = \emptyset \text{ or } P_{T^\beta}(\alpha) \rhd P_{T^\delta}(\beta),\end{aligned}$$

and we may also write e.g. $\Pr(\alpha \mid \beta) \rhd \Pr(\gamma)$ as an abbreviation for $\Pr(\alpha \mid \beta) \rhd \Pr(\gamma \mid \top)$. Related to these, the atom $\mathbf{X} \perp\!\!\!\perp_{\mathbf{Z}} \mathbf{Y}$ (*conditional independence atom*) states that for any given value for the variables in $\mathbf{Z}$ the variable sets $\mathbf{X}$ and $\mathbf{Y}$ are probabilistically independent. Its special case with $\mathbf{Z} = \emptyset$ is called *marginal independence atom.* We can define these atoms in terms of conditional comparison atoms:

$$\mathbf{X} \perp\!\!\!\perp \mathbf{Y} := \bigwedge_{\substack{\mathbf{x} \in \mathrm{Ran}(\mathbf{X}) \\ \mathbf{y} \in \mathrm{Ran}(\mathbf{Y})}} \Pr(\mathbf{X} = \mathbf{x}) = \Pr(\mathbf{X} = \mathbf{x} \mid \mathbf{Y} = \mathbf{y})$$

$$\mathbf{X} \perp\!\!\!\perp_{\mathbf{Z}} \mathbf{Y} := \bigwedge_{\substack{\mathbf{x} \in \mathrm{Ran}(\mathbf{X}) \\ \mathbf{y} \in \mathrm{Ran}(\mathbf{Y}) \\ \mathbf{z} \in \mathrm{Ran}(\mathbf{Z})}} \Pr(\mathbf{X} = \mathbf{x} \mid \mathbf{Z} = \mathbf{z}) = \Pr(\mathbf{X} = \mathbf{x} \mid \mathbf{YZ} = \mathbf{yz})$$

Hence the atoms (and the dependence atom expressed as $\mathbf{Y} \perp\!\!\!\perp_{\mathbf{X}} \mathbf{Y}$) are expressible in $\mathcal{P}$ extended with the conditional probability comparison atoms. It is an open question whether the probabilistic independence atoms are already expressible in $\mathcal{PCO}$.

The above definitions of atoms imply that our languages, if enriched with conditional probability atoms and arbitrary applications of the disjunction $\vee$, are strong enough to the express properties of multiteams that are expressible in the quantifier free fragments of the logics $\mathrm{FO}(\perp\!\!\!\perp)$ (*probabilistic independence logic*) and $\mathrm{FO}(\approx)$ (*probabilistic inclusion logic*), over any fixed finite structure. The expressivity and complexity of these logics have been thoroughly studied in the probabilistic and multiteam semantics literature (see [10, 11, 14, 19, 20, 21, 39]).

It was observed in [4] that $\Pr(\alpha \mid \gamma) \rhd \epsilon$ and $\Pr(\alpha \mid \gamma) \rhd \Pr(\beta \mid \gamma)$ can be defined by $\gamma \supset \Pr(\alpha) \rhd \epsilon$ and $\gamma \supset \Pr(\alpha) \rhd \Pr(\beta)$, respectively. The latter result concerns comparison atoms in which both probabilities are conditioned over the same formula, γ. We establish that this restriction is necessary, and that $\Pr(\alpha \mid \gamma) \geq \Pr(\beta \mid \delta)$ is not, in general, expressible in $\mathcal{PCO}$. The full proof of the theorem is given in the extended version of the paper ([5]).

▶ **Theorem 21.** *The comparison atoms* $\Pr(\alpha \mid \beta) \lhd \Pr(\gamma \mid \delta)$ *and* $\Pr(\alpha \mid \beta) \lhd \Pr(\gamma)$*, (where* $\lhd \in \{\leq, \geq, <, >, =\}$*) are not, in general, expressible in* $\mathcal{PCO}$*.*

Proof (sketch). Due to the equivalence $\Pr(\alpha \mid \beta) \lhd \Pr(\gamma \mid \top) \equiv \Pr(\alpha \mid \beta) \lhd \Pr(\gamma)$, it suffices to prove the theorem for $\Pr(\alpha \mid \beta) \lhd \Pr(\gamma)$.

Fix a signature σ, $\delta \in [0, 1] \cap \mathbb{Q}$ and take four distinct assignments $s_i, s_j, s_k, s_l \in \mathbb{B}_\sigma$. The proof proceeds by showing that the conjunction

$$\Xi := \Pr(\hat{\alpha}_k \vee \hat{\alpha}_i \mid \hat{\alpha}_l \vee \hat{\alpha}_i) \lhd \Pr(\hat{\alpha}_l \vee \hat{\alpha}_j) \ \wedge \ \Pr(\hat{\alpha}_i) = \delta \ \wedge \ \Pr(\hat{\alpha}_j \vee \hat{\alpha}_k \vee \hat{\alpha}_l) = 1 - \delta$$

has a probability set that cannot be characterized in terms of systems of linear inequalities, and thus is not expressible in $\mathcal{P}(\supset)$; extending the result to the whole $\mathcal{PCO}$ is then straightforward. Calculation shows that any T satisfies Ξ if and only if the two-variable inequality

$$2\epsilon_k \epsilon_l + 2\delta\epsilon_k + (2\delta - 2)\epsilon_l + 2\delta^2 \rhd 0$$

holds. Standard geometric techniques (analysis of the homogeneous discriminant) tell us that, in the intersection of the (k, l)-plane with Δ^3, the frontier of the set defined by this inequality is a segment of a nondegenerate conic (a hyperbola). But, clearly, no linear set can have a segment of hyperbola as a subset of its frontier. ◀

5 Conclusion

We embarked for a comprehensive study of the expressive power of logics of probabilistic reasoning and causal inference in the unified setting of causal multiteam semantics. We focused on the logic $\mathcal{PCO}$ that can express probability comparisons in a dataset, and encompasses interventionist counterfactuals and selective implications for describing consequences of actions and consequences of learning from observations, respectively. In addition, we considered the syntactic fragments $\mathcal{P}^{-}$, $\mathcal{P}$, $\mathcal{P}(\supset)$, and $\mathcal{P}(\Box\!\!\rightarrow)$ of $\mathcal{PCO}$ and proved that they form a strict expressivity hierarchy (see Figure 1 on page 3). Moreover, we discovered natural complete characterizations, for each of the aforementioned logics, based on the families of linear equations needed to define the corresponding classes of causal multiteams (satisfying some invariances); these results are summarized in Table 1 (on page 4). Finally, we established that conditional probability statements of the forms $\Pr(\alpha \mid \beta) \leq \Pr(\gamma \mid \delta)$ and $\Pr(\alpha \mid \beta) \leq \Pr(\gamma)$ are not in general expressible in $\mathcal{PCO}$, and separated $\mathcal{PCO}$ from its extension $\mathcal{PCO}^{\omega}$ with infinitary disjunctions.

Analogous to the folklore result that the logic $L_{\infty\omega}$ can define all classes of finite structures, it was shown in [4] that the same holds for $\mathcal{PCO}^{\omega}$ with respect to all classes of causal multiteams that are closed under *rescaling* and have the *empty multiteam property*. While any logic that is expressively complete in this sense is uncountable, it is an interesting task to identify more expressive finitary languages. We describe some future directions of research:

- In the languages we considered, the usage of the *strict tensor* $\vee$ was restricted to $\mathcal{CO}$ formulae. What impact would removing this restriction have on the expressivity of the languages? We conjecture that liberalizing this operator would allow to capture probability sets described by any linear inequality.
- Can (conditional) probabilistic independence atoms be expressed in $\mathcal{PCO}$? We conjecture the negative in line with [21, Proposition 26], which establishes that it is not expressible in $\mathrm{FO}(\approx)$, the *probabilistic inclusion logic* of [19] (although the proof in [21] relies on the use of quantifiers).
- How can our results be extended to cover infinite signatures? Here one might need to extend the languages with quantifiers ranging over data values.
- Our characterizations cover only logics that express linear properties of data. Can we generalize our results if some natural source of multiplication, such as conditional probabilistic independence or the conditional comparison atoms, are added to the logics? It was shown by Hannula et al. [20] that the so-called *probabilistic independence logic* is equiexpressive with a variant of existential second-order logic that has access to addition and multiplication of reals.
- Finally, a promising direction for future work would be to study temporal aspects of causal inference (see e.g., [25]) via (probabilistic) temporal logics by generalising *temporal team semantics* introduced by Krebs et al. [26] and further developed by Gutsfeld et al. [15].

We conclude by pointing out the formal similarity of our work with some results obtained for first-order logics with probabilistic dependencies, such as the aforementioned language $\mathrm{FO}(\approx)$. Such languages do not formalize causation, and yet we can conjecture that $\mathcal{PCO}$ might be embeddable in $\mathrm{FO}(\approx)$ (similarly as the language $\mathcal{CO}$ is embedded into first-order logic in [1]). This idea is supported by a result of Hannula and Virtema ([21]) that establishes that definability in $\mathrm{FO}(\approx)$ can be reformulated in linear programming. It is however unknown which exact fragment of linear programming corresponds (in the sense of our Table 1) to the language $\mathrm{FO}(\approx)$; such a characterization would give precise limits to the possibility of embedding results.

References

1 Fausto Barbero and Pietro Galliani. Embedding causal team languages into predicate logic. *Annals of Pure and Applied Logic*, pages 103–159, 2022. doi:10.1016/j.apal.2022.103159.

2 Fausto Barbero and Gabriel Sandu. Interventionist counterfactuals on causal teams. In *CREST 2018 Proceedings – Electronic Proceedings in Theoretical Computer Science*, volume 286, pages 16–30. Open Publishing Association, January 2019. doi:10.4204/eptcs.286.2.

3 Fausto Barbero and Gabriel Sandu. Team semantics for interventionist counterfactuals: observations vs. interventions. *Journal of Philosophical Logic*, 50:471–521, 2021.

4 Fausto Barbero and Gabriel Sandu. Multiteam semantics for interventionist counterfactuals: probabilities and causation. pre-print, 2023. arXiv:2305.02613.

5 Fausto Barbero and Jonni Virtema. Expressivity landscape for logics with probabilistic interventionist counterfactuals. *CoRR*, abs/2303.11993, 2023. doi:10.48550/arXiv.2303.11993.

6 Fausto Barbero and Jonni Virtema. Strongly complete axiomatization for a logic with probabilistic interventionist counterfactuals. In Sarah Gaggl, Maria Vanina Martinez, and Magdalena Ortiz, editors, *Logics in Artificial Intelligence*, pages 649–664, Cham, 2023. Springer Nature Switzerland.

7 Fausto Barbero and Fan Yang. Characterizing counterfactuals and dependencies over (generalized) causal teams. *Notre Dame Journal of Formal Logic*, 63(3), 2022. doi:10.1215/00294527-2022-0017.

8 Elias Bareinboim, Juan Correa, Duligur Ibeling, and Thomas Icard. On Pearl's hierarchy and the foundations of causal inference (1st edition). In Hector Geffner, Rina Dechter, and Joseph Y. Halpern, editors, *Probabilistic and Causal Inference: the Works of Judea Pearl*, pages 507–556. ACM Books, 2022.

9 Rachael Briggs. Interventionist counterfactuals. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 160(1):139–166, 2012.

10 Arnaud Durand, Miika Hannula, Juha Kontinen, Arne Meier, and Jonni Virtema. Approximation and dependence via multiteam semantics. *Ann. Math. Artif. Intell.*, 83(3-4):297–320, 2018. doi:10.1007/s10472-017-9568-4.

11 Arnaud Durand, Miika Hannula, Juha Kontinen, Arne Meier, and Jonni Virtema. Probabilistic team semantics. In Flavio Ferrarotti and Stefan Woltran, editors, *Foundations of Information and Knowledge Systems*, pages 186–206, Cham, 2018. Springer International Publishing.

12 Arnaud Durand, Juha Kontinen, and Heribert Vollmer. Expressivity and complexity of dependence logic. *Dependence Logic: Theory and Applications*, pages 5–32, 2016.

13 David Galles and Judea Pearl. An axiomatic characterization of causal counterfactuals. *Foundations of Science*, 3(1):151–182, January 1998.

14 Erich Grädel and Richard Wilke. Logics with multiteam semantics. *ACM Trans. Comput. Log.*, 23(2):13:1–13:30, 2022. doi:10.1145/3487579.

15 Jens Oliver Gutsfeld, Arne Meier, Christoph Ohrem, and Jonni Virtema. Temporal team semantics revisited. In Christel Baier and Dana Fisman, editors, *LICS '22: 37th Annual ACM/IEEE Symposium on Logic in Computer Science, Haifa, Israel, August 2 - 5, 2022*, pages 44:1–44:13. ACM, 2022. doi:10.1145/3531130.3533360.

16 Joseph Halpern. *Actual causality*. MIT Press, 2016.

17 Joseph Y. Halpern. Axiomatizing causal reasoning. *J. Artif. Int. Res.*, 12(1):317–337, May 2000.

18 Joseph Y. Halpern. From causal models to counterfactual structures. *Review of Symbolic Logic*, 6(2):305–322, 2013. doi:10.1017/s1755020312000305.

19 Miika Hannula, Åsa Hirvonen, Juha Kontinen, Vadim Kulikov, and Jonni Virtema. Facets of distribution identities in probabilistic team semantics. In *European Conference on Logics in Artificial Intelligence*, pages 304–320. Springer, 2019.

20 Miika Hannula, Juha Kontinen, Jan Van den Bussche, and Jonni Virtema. Descriptive complexity of real computation and probabilistic independence logic. In Holger Hermanns, Lijun Zhang, Naoki Kobayashi, and Dale Miller, editors, *LICS '20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science, Saarbrücken, Germany, July 8-11, 2020*, pages 550–563. ACM, 2020. doi:10.1145/3373718.3394773.

21 Miika Hannula and Jonni Virtema. Tractability frontiers in probabilistic team semantics and existential second-order logic over the reals. *Ann. Pure Appl. Log.*, 173(10):103108, 2022. doi:10.1016/j.apal.2022.103108.

22 James J Heckman and Edward J Vytlacil. Econometric evaluation of social programs, part i: Causal models, structural models and econometric policy evaluation. *Handbook of econometrics*, 6:4779–4874, 2007.

23 MA Hernan and J Robins. *Causal Inference: What if.* Boca Raton: Chapman & Hill/CRC, forthcoming.

24 Wilfrid Hodges. Compositional semantics for a language of imperfect information. *Logic Journal of the IGPL*, 5:539–563, 1997.

25 Samantha Kleinberg. A logic for causal inference in time series with discrete and continuous variables. In Toby Walsh, editor, *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, pages 943–950. IJCAI/AAAI, 2011. `doi:10.5591/978-1-57735-516-8/IJCAI11-163`.

26 Andreas Krebs, Arne Meier, Jonni Virtema, and Martin Zimmermann. Team semantics for the specification and verification of hyperproperties. In Igor Potapov, Paul G. Spirakis, and James Worrell, editors, *43rd International Symposium on Mathematical Foundations of Computer Science, MFCS 2018, August 27-31, 2018, Liverpool, UK*, volume 117 of *LIPIcs*, pages 10:1–10:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. `doi:10.4230/LIPIcs.MFCS.2018.10`.

27 David Lewis. *Counterfactuals.* Oxford: Blackwell Publishers, 1973.

28 Stephen L Morgan and Christopher Winship. *Counterfactuals and causal inference.* Cambridge University Press, 2015.

29 Milan Mossé, Duligur Ibeling, and Thomas Icard. Is causal reasoning harder than probabilistic reasoning? *The Review of Symbolic Logic*, pages 1–26, 2022.

30 Judea Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, New York, NY, USA, 2000.

31 Judea Pearl and Dana Mackenzie. *The book of why: the new science of cause and effect.* Basic books, 2018.

32 Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference: foundations and learning algorithms.* MIT Press, 2017.

33 Bernhard Schölkopf. Causality for machine learning. In *Probabilistic and Causal Inference: The Works of Judea Pearl*, pages 765–804. Association for Computing Machinery, 2022.

34 Peter Spirtes. Directed cyclic graphical representations of feedback models. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pages 491–498, 1995.

35 Peter Spirtes, Clark Glymour, and Richard N. Scheines. *Causation, Prediction, and Search*, volume 81 of *Lecture Notes in Statistics.* Springer New York, 1993.

36 Robert C. Stalnaker. A theory of conditionals. *Americal Philosophical Quarterly*, pages 98–112, 1968.

37 Robert H Strotz and Herman OA Wold. Recursive vs. nonrecursive systems: An attempt at synthesis (part i of a triptych on causal chain systems). *Econometrica: Journal of the Econometric Society*, pages 417–427, 1960.

38 Jouko Väänänen. *Dependence Logic: A New Approach to Independence Friendly Logic*, volume 70 of *London Mathematical Society Student Texts.* Cambridge University Press, 2007.

39 Richard Wilke. On the Presburger fragment of logics with multiteam semantics. *Ann. Pure Appl. Log.*, 173(10):103120, 2022. `doi:10.1016/j.apal.2022.103120`.

40 Sewall Wright. Correlation and causation. *Journal of agricultural research*, 20:557–585, 1921.

41 Jiji Zhang. A Lewisian logic of causal counterfactuals. *Minds and Machines*, 23(1):77–93, 2013.