

Axiomatizations of causal reasoning under indeterministic causal laws

Fausto Barbero*

Department of Philosophy, History and Art Studies, University of Helsinki, Unioninkatu 40, 00170 Helsinki, Finland

* Correspondence: fausto.barbero@helsinki.fi (Fausto B)

Abstract

We investigate the generalization of causal models to the case of indeterministic causal laws that was suggested in Halpern (2000). We give an overview of what differences in modeling are enforced by this more general perspective, and propose an implementation of generalized models in the style of the causal team semantics of Barbero & Sandu (2021). In these models, the laws are not represented by functions (as in the deterministic case), but more generally by relations. We point out significant differences in the indeterministic vs. the deterministic case, both for what concerns the axiomatization of counterfactuals and the definability of causal notions in terms of them. We notice, for instance, that the notions of parenthood and direct cause, differently from the deterministic case, come apart. We provide two main strongly complete axiomatizations, one for the class of indeterministic causal models and one for their team generalizations, which can also represent uncertainty about the state of the system. Then, we also identify axioms that allow specializing the axiomatizations to a number of significant subclasses, characterized by properties such as acyclicity, totality and determinism of the causal laws. We also see that a minor change in the notion of signature of the models has a dramatic impact; in particular, it makes so that the deterministic subclass of models become undefinable in the usual Halpern-style counterfactual language. We see that the definability of determinism can be recovered either by a restriction of the class of models, by allowing right-nested counterfactuals, or by introducing an operator for information update.

Keywords: Indeterministic causal models, Interventionist counterfactuals, Causal teams, Axiomatizations, Nested counterfactuals

Citation: Barbero F. 2026. Axiomatizations of causal reasoning under indeterministic causal laws. *The Knowledge Engineering Review* 41: e008
<https://doi.org/10.48130/ker-0026-0005>

1 Introduction

When the existence of causal laws is considered at all, the discussion of causality in the philosophy of science tends to focus on *deterministic* causal mechanisms, which uniquely determine an effect given full knowledge of the causes. The restriction to deterministic laws is, in some cases, a deliberate choice to avoid technical complications (see, e.g., Paul et al.^[1], section 4.3), as many important and sometimes puzzling aspects of causation already emerge at this level. However, often, the laws considered in applied sciences are not of this kind. A physical law may allow us to predict, for example, that

A cannonball shot at a certain angle will fall *within a certain range*.

In contrast to more idealized laws of mechanics, this kind of law tells us that, *even without taking into account some factor that is not mentioned by the law* (such as the direction or intensity of the wind), we can predict that the effect will (quite literally) fall within a certain range of possible values. These kinds of causal dependencies are discussed at length, e.g., in Bohm's book *Causality and chance in modern physics*^[2], chapter I.7, under the name of *one-to-many* causal relationships.

Note that in the cannonball example, the law is indeterministic due to the impossibility of taking into account all factors, i.e., it accounts for *epistemic* limitations; the indeterministic law might, e.g., just be a rough approximation of an underlying deterministic law. On the other hand, many physicists would think that the following statement expresses an irreducibly indeterministic law about the behaviour of a particle:

If an atom of silver passes through a Stern-Gerlach apparatus, it will be detected in either of two specific areas of the screen.

In other words, a physical law may tell us that the specifics of the outcome of a certain experiment are unpredictable in principle, even if we had a complete knowledge of the initial conditions of a physical

system. In this case, an indeterministic law tells us that the particle will be detected somewhere within two areas, but nothing more. Indeterministic laws do not just arise within scientific theories, but also justify everyday statements such as

If I toss the coin, it will land on heads or tails.

Far from being a platitude, such a statement implicitly assumes an indeterministic causal law, which allows two values for the future state of the coin, while excluding other alternatives (such as "the coin staying in my pocket" or "the coin breaking into pieces"). Notice also that all these examples do not *require* probabilities for their formulation and justification. We might not know what probabilistic distributions are involved, or, in some context, it might even be mathematically impossible to associate a probability distribution to a given phenomenon (see e.g., Wysocki's work^[3] for some examples of this kind). Thus, probabilistic laws are only a special case of indeterministic laws.

In the last decades, causal reasoning has been mathematically formalized in the fields of statistics and computer science known as *causal discovery*^[4] and *causal inference*^[5], which have found applications, e.g., in machine learning^[6,7], epidemiology^[8], econometrics^[9], and social sciences^[10]. Do indeterministic laws feature in these approaches? There are two main types of models considered in this context. The first are the *Bayesian networks*; these consist of probabilistic distributions paired with graphs, and are intrinsically indeterministic models; but laws or mechanisms are not specified in any way by such models. Instead, *structural equation models* (or *causal models*), use systems of equations to encode the causal laws that link together the relevant variables of the scenario under examination. The equations take the form:

$$Y := f(X_1, \dots, X_n)$$

where the term $f(X_1, \dots, X_n)$ stands for a *function* of the variables $X_1, \dots, X_n$. The equations, together with some information about the state of the variables, allow one to formulate and often answer queries about

the deterministic and probabilistic behaviour of the variables. A significant example of such a query is establishing the truth or falsity of an *interventionist counterfactual*^[11]:

If variables $Z_1, \dots, Z_m$ were set to values $z_1, \dots, z_m$, then condition ψ would hold.

These kinds of expressions can be studied with the methods of logic. The classic work of Halpern^[12,13] provided a wealth of complete deduction systems for classes of causal models, and the idea has since been extended in various directions^[14–24]. Now, the very fact that causal laws are represented by functions amounts to a restriction to deterministic laws. The natural way to extend these kinds of models to the indeterministic case is to replace these functional constraints with *relational* constraints. If we write R_Y for the causal law describing the behaviour of variable Y , we shall interpret the statement that $(x_1, \dots, x_n, y) \in R_Y$ as expressing the fact that, if variables $X_1, \dots, X_n$ are set to values $x_1, \dots, x_n$, then Y *might* take the value y (but it might also take any other value y' such that $(x_1, \dots, x_n, y') \in R_Y$). Equivalently, we can recover the equational form by encoding the causal law as a *multivalued function* $\mathcal{F}_Y$ that associates to each tuple of values for $X_1, \dots, X_n$ a set of possible values for Y .

The possibility of such an extension is hinted at by Halpern in the conclusions of his paper on the axiomatization of interventionist counterfactuals:

(...) *a more general approach to modeling causality would allow there to be more than one value of X once we have set all other variables. This would be appropriate if we model things at a somewhat coarser level of granularity, where the values of all the variables other than X do not suffice to completely determine the value of X . I believe the results of this paper can be extended in a straightforward way to deal with this generalization, although I have not checked the details*^[12].

Perhaps stating that this generalization is "straightforward" discouraged researchers from pursuing this direction: or was it the discovery that, after all, the details are not so straightforward? To the best of our knowledge, the idea has been taken up again only in 2021, when Peters and Halpern^[17,25] introduced the more complex framework of *generalized structural equation models* (GSEMs). In that context, causal laws are not mentioned at all, but it is evident that indeterministic causal laws can be modeled in a GSEM as a special case. More recently, a simpler framework, closer to Halpern's original idea for extending causal models, was considered by Wysocki^[3], who provided a cursory examination of the differences between deterministic and indeterministic counterfactuals^①. The purpose of the present paper is to sharpen the understanding of the logic of interventionist counterfactuals in the indeterministic (but not probabilistic) context and ultimately provide complete axiomatizations for significant classes of indeterministic causal models. The models we use (*relational causal teams*) are a generalization of the *causal teams* proposed in Barbero & Sandu^[28] and are somewhat more general than those of Wysocki (one key difference being that our models—like structural equation models—do not necessarily obey the Markov condition^②). In section 2, we will describe, in general terms, some key differences between deterministic and indeterministic frameworks, and how they will affect our modeling choices. We then use the insights just gained to provide (section 3.1) a precise definition of indeterministic models. We then introduce (section 3.2) a language $\mathcal{H}$ similar to those considered in Halpern's work^[12], and its extension $\mathcal{H}^+$ that also allows right-nesting of counterfactuals. In section 4, we provide a strongly complete

axiomatization of language $\mathcal{H}$ over the class of all *singleton* models (of a given signature); this allows us to make a direct comparison with the usual case of deterministic causal models. Later, section 6 provides an axiomatization of $\mathcal{H}^+$ over the class of *all* models; section 5 lays down some necessary technical tools for the proof, by clarifying which models are indistinguishable using $\mathcal{H}$ or $\mathcal{H}^+$, and explaining to what extent the causal laws can be described in these languages. In section 7, we analyze notions such as parenthood, direct cause, and (non-) dummy arguments, we see that (differently from the deterministic case) they come apart, and investigate their mutual relationships. Some of these insights help us produce complete axiomatizations for models where cyclic causal relationships are forbidden (section 8); more precisely, we axiomatize two classes of models that we call, respectively, *strictly recursive* and *visibly recursive*. In section 9, we axiomatize two other interesting subclasses, that of *deterministic* models (those in which the causal laws associate at most one output to each input value for the parents) and that of *total* models (where the causal laws associate at least one possible effect to each input value for the parents). We also see that the *recursive* class of models, which forbids cycles of direct causation, can be axiomatized relatively to the class of total models. Combining all these results, we obtain an alternative axiomatization for the (usual, deterministic) recursive causal models. In section 10, we consider a slight variation of the framework, which adopts a notion of signature closer to that which is often used in the philosophical literature. Such simplified signatures only fix a variable domain and the allowed value ranges for the variables; they do not decide which variables are going to be associated with a causal law. While the use of such signatures is tempting, as it can potentially lead to more general completeness results, we see (section 10.1) that, in the indeterministic context, it leads to a barrage of undefinability results (which, among other things, prevent us from extending the axiomatization strategies used in the rest of the paper). Some of these results—e.g., the undefinability of direct cause—extend to full signatures. Others, like the undefinability of the deterministic class of models, affect $\mathcal{H}$ only if simplified signatures are adopted. We show three different strategies for recovering the definability of determinism while keeping the simplified signatures: First, we see it is again definable if we restrict attention to singleton models (section 10.2). Second, we see the definability is regained if one allows for right-nested counterfactuals (section 10.3). Third, the same is achieved by extending $\mathcal{H}$ with a second conditional operator which describes learning or information updates (section 10.4); we analyze the potential axioms and rules for this language extension. Lastly, in section 11, we compare our approach and results to the recent literature on indeterministic counterfactuals.

The reader who is interested in the conceptual aspects of indeterministic causation and counterfactuals but not in the technicalities of proving the completeness results can pick the following shorter route: reading sections 2, 3, the first two pages of 4, sections 7, 8, 10, and 11.

2 Differences with the deterministic case

We describe here a few differences and challenges that arise when trying to describe models that are not limited to deterministic causal laws.

We need to introduce some notational conventions, for which we follow usage from the field of causal inference. We adopt capital

① There is also recent work by Beckers^[26,27] inspired by the conference version of this paper. We will discuss it in section 11.

② Which, in the context of non-probabilistic models, amounts to the fact that there may be data dependencies among exogenous variables.

letters $X, Y, \dots$ to denote **variables**, which are used e.g. to represent magnitudes such as pressure, temperature, etc.; binary variables may also be used to represent whether an event occurs or not. The **values** that a variable X may take will be denoted by small letters such as x, x', x'' . We use atomic formulas of the form $X = x$ to assert that (in a given context) X takes the value x ; the set of atomic formulas of signature σ will be denoted by Atom_σ . We assume that the variables come from two disjoint finite sets $\mathcal{E}$ (**external** variables) and $\mathcal{I}$ (**internal** variables); we write Dom for $\mathcal{E} \cup \mathcal{I}$. Intuitively, the internal variables are those whose causal laws will be described by the models, while the external variables are those whose causes are not further analyzed.^③ Furthermore, we assume the existence of a function Ran that associates to each variable X a finite set $\text{Ran}(X)$ of possible values. The triple $(\mathcal{E}, \mathcal{I}, \text{Ran})$ is called a **signature**. An **assignment** of signature $\sigma = (\mathcal{E}, \mathcal{I}, \text{Ran})$ is a function that assigns to each variable an allowed value (i.e., a function $f : \text{Dom} \rightarrow \bigcup_{V \in \text{Dom}} \text{Ran}(V)$ such that, for all $V \in \text{Dom}$, $s(V) \in \text{Ran}(V)$). We call $\mathcal{A}_\sigma$ the set of all such assignments.

Boldface letters such as $\mathbf{X}, \mathbf{x}$ denote (depending on the context) either finite sets or finite tuples of variables, resp. of values. If $\mathbf{X} = (X_1, \dots, X_n)$, then $\text{Ran}(\mathbf{X})$ abbreviates $\text{Ran}(X_1) \times \dots \times \text{Ran}(X_n)$. If furthermore $\mathbf{x} = (x_1, \dots, x_n)$, we abbreviate as $\mathbf{X} = \mathbf{x}$ either a multiset of atomic formulas $X_1 = x_1, \dots, X_n = x_n$ or their conjunction. When $\mathbf{X}$ is a nonempty tuple of variables and s is an assignment, $s(\mathbf{X})$ denotes the tuple of values that s associates to the variables in $\mathbf{X}$.

We use $\mathbf{W}$, resp. $\mathbf{W}_X, \mathbf{W}_{XY}$ to denote tuples listing without repetitions the variables in Dom , resp. $\text{Dom} \setminus \{X\}, \text{Dom} \setminus \{X, Y\}$.

2.1 Uncertainty

While working with deterministic causal laws, it has been customary to describe a scenario by means of an assignment of values to the variables of the system.^④ This may be appropriate, in some cases, also in the presence of indeterministic laws. Consider the scenario "Alice tosses a coin, and it comes out heads"; we would model it by one causal law (saying that the coin tossing will lead either to a "heads" or to a "tails" outcome) together with the following assignment:

$A \rightarrow C$
1 heads

where the Boolean variable A tells us whether Alice tossed the coin ($A = 1$) or not ($A = 0$), and C records the outcome of the toss; we drew the symbol $\rightarrow$ to emphasize that A is an indeterministic cause of C . However, one might also want to model the scenario "Alice tossed a coin". The indeterministic law does not allow us to infer the outcome of the toss; thus, to represent this kind of scenario, we need *two* assignments, describing the two alternative situations that are not excluded by the description of the scenario:

$A \rightarrow C$
1 heads
1 tails

These kinds of considerations lead us to shift attention away from causal models to the more general *causal teams*^[28], models which allow *sets* ("teams") of variable assignments compatible with the

causal laws. This perspective is not alien to the previous literature on causal inference: it is implicit, e.g., in the treatment of interventions in the presence of cyclic causal laws. Even when considering a full description of a scenario (i.e., a single assignment), such an intervention may produce a multiplicity of possible new scenarios^[11,12].

In the presence of indeterministic causal laws, the need for teams also emerges when asking counterfactual queries about a fully described scenario, e.g., "Alice did not toss and the coin is still in her pocket". If we want to know what would have happened if Alice had tossed the coin, we will need to consider an intervention that sets A to 1 ($do(A = 1)$), which produces the same 2-assignment model that we considered in the second scenario.^⑤ We thus see that the class of causal models is not closed under interventions: intervening on an (indeterministic) causal *model* produces an (indeterministic) causal *team*. Our definitions will ensure that the class of indeterministic causal teams is closed under interventions.

2.2 Specifying the causal laws

In the deterministic case, the causal laws can be specified in at least two different ways, which are, for most purposes, equivalent:

1. First, pick out some variables, which will be considered *endogenous*. For each endogenous variable V , we specify which other variables are *direct causes* or *parents* of V ; call this set PA_V . We then, specify the law for V as a function $\mathcal{F}_V : \text{Ran}(PA_V) \rightarrow \text{Ran}(V)$.

2. Assign to each variable V a function $\mathcal{F}_V : \text{Ran}(\mathbf{W}_V) \rightarrow \text{Ran}(V)$. Observe that some of the variables of $\mathbf{W}_V$ are dummy arguments of $\mathcal{F}_V$; define PA_V as the set of variables of $\mathbf{W}_V$ that are *not* dummy for $\mathcal{F}_V$. Define the set of endogenous variables as those whose parent set is nonempty.

The first approach is more natural and direct, but it has technical disadvantages: prominently, the fact that it allows a proliferation of essentially equivalent^⑥ models (for example, we might have two models that differ only in that, in the former, variable Z is generated by the law $\mathcal{F}_Z(X, Y) = X + Y$, while in the latter, the law is $\mathcal{F}_Z(X, Y, U, V, W) = X + Y$, i.e., the same function with three dummy arguments). For this reason, in technical papers the second approach is usually preferred.

Analogously, in the indeterministic case, we might want to encode the laws as relations in $\text{Ran}(PA_V) \times \text{Ran}(V)$ or, instead, in $\text{Ran}(\mathbf{W}_V) \times \text{Ran}(V)$. Unfortunately, in the indeterministic case, Approach 2 seems not to be viable. We consider two examples that raise problems for this approach.

Example 2.1 (Jumpin' Alice). *In this scenario, we have Boolean variables A (whether Alice jumps), B (whether Bob tosses a coin), and C representing three possible states of the coin (whether it is on heads, tails, or stays in Bob's pocket). Suppose we represent the causal law determining the state of the coin by the relation $\mathcal{F}_C \subseteq \text{Ran}(A) \times \text{Ran}(B) \times \text{Ran}(C)$, $\mathcal{F}_C = \{(0, 0, \text{in-pocket}), (1, 0, \text{in-pocket}), (0, 1, \text{heads}), (0, 1, \text{tails}), (1, 1, \text{heads}), (1, 1, \text{tails})\}$. Suppose also that we know that Alice did not jump, B tossed the coin and it came heads:*

A	B	C
0	1	heads

^③ The more common terminology, distinguishing between exogenous and endogenous variables, will be (re)defined later.

^④ At least, this has been most common in the philosophical literature; but also e.g., Halpern's work^[13], section 2.7 follows this convention.

^⑤ We will assume that the same team is produced by the intervention $do(A = 1)$ when applied to our initial scenario "Alice tossed the coin and it came heads." i.e., when considering a subjunctive conditional with a *true* antecedent, we will also use non-actual worlds in the evaluation of the statement. This move seems justified because, in any case, the "actual world" cannot really be obtained via the intervention $do(A = 1)$; rather, we obtain (one of the) worlds in which Alice's toss is dictated by an external intervention and not by its natural causes. As observed in Briggs's work^[14], causal models do not even satisfy *weak centering*.

^⑥ In the sense that two such models are not distinguishable by any reasonable formal language of interventionist counterfactuals.

Suppose we intervene on the system by forcing Alice to jump. According to Halpern's definition of intervention^[13], the possible scenarios after such an intervention are those that agree 1) with the law, 2) with the condition $A = 1$, or 3) with the current state of the exogenous variables^{⑦⑧} (different from A), i.e., $B = 1$. There are two assignments consistent with these conditions, namely:

A	B	C
1	1	heads
1	1	tails

In other words, forcing Alice to jump—or imagining her to counterfactually jump—makes us to lose information about the outcome of a coin toss. This may be unwanted, for example, if Alice's and Bob's feats take place, simultaneously, on distant planets.

There seems to be a straightforward way to address this problem. If Alice's and Bob's deeds are obviously unrelated in the scenario we wish to model, then A should not be treated as a direct cause of B (nor, more generally, there should there be a path of direct causes from A to B). This can be guaranteed using the Approach 1 outlined at the beginning of the section. Now, it seems reasonable that an intervention on a variable A should only affect variables that are (directly or indirectly) causally dependent on A . Then, we should reject the second assignment in the table, which modifies the value of C , a non-descendant of A . To this end, we will adopt a definition of intervention in the style of Barbero & Galliani^[30], which, differently from Halpern's, does not violate this constraint on cyclic models.

There is more. As the following example shows, and contrarily to what happens in the deterministic case, identifying the dummy arguments of the relational laws is not sufficient for identifying the direct causes and parents of a variable.

Example 2.2 (Two-coin Bob). In this scenario, Bob has two coins, say coin 1 and coin 2, and may toss one of the two ($B = 1$ or $B = 2$). Variable O represents the outcome of the toss (heads or tails), and its behaviour is described by the relational law $\mathcal{F}_O = \{(1, \text{heads}), (1, \text{tails}), (2, \text{heads}), (2, \text{tails})\}$. Furthermore, we know that Bob has tossed coin 1 and got heads:

B	O
1	heads

Now, B is a dummy argument of the law $\mathcal{F}_O$: changing the value of B does not change the range of values that O may attain ($\mathcal{F}_O(1) = \mathcal{F}_O(2) = \{\text{heads}, \text{tails}\}$). Seen as a multivalued function, $\mathcal{F}_O$ is a constant-valued law. But our intuition about the real world seems to disagree with the idea that B is not a cause of O , and that O should be considered external (uncaused) in this context. Indeed, we presume that, if we force Bob to make another toss, the outcome may change. And we usually take the future behaviour of a mechanism to be a reliable indicator of its counterfactual behaviour. Thus, B should be considered a direct cause of O , even if it is a dummy argument of $\mathcal{F}_O$. Our definition of intervention will ensure that the intervention $\text{do}(B = 1)$, forcing Bob to repeat the toss, will produce two alternative scenarios:

B	O
1	heads
1	tails

where, the range of possible values for O has changed from $\{\text{heads}\}$ to $\{\text{heads}, \text{tails}\}$.

This second example should clarify that a relational causal law $\mathcal{F}_V$, by itself, determines the set of its own non-dummy arguments, but

does not determine the (possibly larger) set of direct causes of V . Learning the set of direct causes is possible once we have a notion of intervention—we may then check whether intervening on a given argument Z may change the range of allowed values for V in some model (when everything else is held fixed). But on the other hand, to know whether intervening on Z may affect V , we need to know whether Z is a causal descendant of V . We can escape this vicious circle by declaring explicitly the set of parents of a variable as part of a specification of the model, i.e., by following the Approach 1 delineated at the beginning of this section (the set of direct causes of V will then be a subset of PA_V).

We remark that the problems raised by the two examples in this section are not artificially induced by switching from causal models to causal teams. The issue from Example 2.1 only concerns how an indeterministic causal law is modeled—this only concerns the function component of a model. Example 2.2 concerns a single-assignment model, which is the straightforward generalization to the indeterministic case of the causal models from the literature, for example in Halpern's book^[13] and in Briggs^[14]. It will not arise by following the perspective of Halpern's original paper^[12], according to which all statements essentially concern the special case of teams that feature a single tuple of values for the exogenous variables and *all* compatible tuples of values for the endogenous ones. We just take this to signal a limit to the expressiveness of the said frameworks.^⑨

3 Formal framework

3.1 Models: relational causal teams

The considerations from the previous section lead us to the following definition of a model. A **team** of signature σ is a set of assignments of signature σ .

Definition 3.1. A **relational causal team** (of signature $\sigma = (\mathcal{E}, \mathcal{I}, \text{Ran})$) is a pair $T = (T^-, \mathcal{F})$, where:

- T^- is a team of signature σ (team component)
- $\mathcal{F}$ (law component) is a function that associates to each $V \in \mathcal{I}$:
 - a set of variables $PA_V \subseteq \text{Dom} \setminus \{V\}$ (parents of V)
 - a relation $\mathcal{F}_V \subseteq \text{Ran}(PA_V) \times \text{Ran}(V)$ (V -generating law)
- For all $V \in \mathcal{I}$, and for all $s \in T^-$, $(s(PA_V), s(V)) \in \mathcal{F}_V$ (compatibility constraint).

For brevity, we will often simply call it a **model**. The last condition (the **compatibility constraint**) can be thought of as admitting in the team only "solutions" of the system of causal laws. When dealing with a model T , we will often bypass mention of the signature by denoting its set of internal variables as $\text{Int}(T)$.

We remark that the causal law for an internal variable V can also be, equivalently, represented by showing what set of values is associated by $\mathcal{F}_V$ to each tuple pa in PA_V ; we may denote such a set as $\mathcal{F}_V(pa)$, thus treating the relation $\mathcal{F}_V$ as a multivalued function. Note that we make an exception to our earlier conventions and denote a tuple of values for the parents of a given variable by pa, pa' , and so on.

We will say that a relational causal team is **total** if all causal laws are total multivalued functions, i.e., for each $V \in \text{Int}(T)$ and for each $pa \in \text{Ran}(PA_V)$, $\mathcal{F}_V(pa) \neq \emptyset$. We will say it is **deterministic** if, for each $V \in \text{Int}(T)$ and for each $pa \in \text{Ran}(PA_V)$, $|\mathcal{F}_V(pa)| \leq 1$. The total

⑦ Roughly speaking, a variable is exogenous if causally unaffected by any other variable in the model. Later we will have a formal definition.

⑧ Applying Halpern's rule here is an extrapolation, since Halpern usually only considers interventions on endogenous variables. The reader who is troubled by this may extend the example with an additional variable U representing the factors that determine Alice's decision, so that A becomes an internal variable.

⑨ In any case, Halpern asserts (personal communication) that his intended semantics is that given in Halpern^[13] and not the earlier one from Halpern^[12].

deterministic models essentially coincide with the causal teams introduced in Barbero & Sandu^[28]. If we add the constraint that the team component must contain exactly one assignment, then what we obtain are essentially the usual causal models.¹⁰ In general, a model with exactly one assignment will be called a **singleton** model; singleton models are the most straightforward generalization of causal models to the indeterministic case (we might also call them **indeterministic causal models**), and for some time they will be our main concern, in section 4. We will also say that a model is **empty** if such is its team component. We remark that empty models are not uninteresting: as we shall see, interventions on empty models may produce nonempty models.

We define the **parenthood graph** of T to be the graph whose vertices are the variables of the system, and where an arrow (directed edge) connects X to Y if and only if X is a parent of Y . The parenthood graph allows us to apply some graph theoretic terminology to the variables; for example, we will say that Y is a *descendant* of X if either Y and X are the same variable, or there is a path $X \rightarrow Z \rightarrow \dots \rightarrow Z_n \rightarrow Y$ in the parenthood graph; otherwise, Y is a *nondescendant* of X . If the parenthood graph is acyclic, we will say that the model is **strictly recursive**. We will see later (section 7) how to define a notion of *causal graph* closer to that used for the deterministic case (and related notions of endogeneity, exogeneity, and recursivity).

Interventions can be defined similarly as in the deterministic case, using the parenthood graph as a guide. It was, however, argued in Barbero & Galliani^[30] that Halpern's general defining clause from Halpern^[12] leads to paradoxical outcomes: if the causal graph has cycles, intervening on a variable X may affect a variable Y that is not its descendant. We then adopt the modifications suggested in Barbero & Galliani^[30].¹¹ We will need some additional notational conventions. If T^- is a team, and $\mathbf{Y}$ a nonempty set of variables in its domain, we $T^-(\mathbf{Y}) := \{\mathbf{y} \in \text{Ran}(\mathbf{Y}) \mid \text{for some } s \in T^-, s(\mathbf{Y}) = \mathbf{y}\}$ be the set of (tuples of) values for $\mathbf{Y}$ that occur in T^- . Given a finite multiset of atomic formulas, say $X_1 = x_1, \dots, X_n = x_n$, we say it is **consistent** if, whenever X_i and X_j are the same variable, then x_i and x_j also coincide. Given a consistent $\mathbf{X} = \mathbf{x}$, and a relational causal team $T = (T^-, \mathcal{F})$ of signature $(\mathcal{E}, \mathcal{I}, \text{Ran})$, the effect of an intervention $do(\mathbf{X} = \mathbf{x})$ on T is to produce a new model $T_{\mathbf{X}=\mathbf{x}} = ((T_{\mathbf{X}=\mathbf{x}})^-, \mathcal{F}_{\mathbf{X}=\mathbf{x}})$ of signature $(\mathcal{E} \cup \mathbf{X}, \mathcal{I} \setminus \mathbf{X}, \text{Ran})$, with components:

- $\mathcal{F}_{\mathbf{X}=\mathbf{x}} := \mathcal{F}|_{\text{Int}(T) \setminus \mathbf{X}}$ (the restriction of $\mathcal{F}$ to $\text{Int}(T) \setminus \mathbf{X}$)
- $(T_{\mathbf{X}=\mathbf{x}})^- := \{s \in \mathbb{A}_{\sigma} \text{ compatible with } \mathcal{F}_{\mathbf{X}=\mathbf{x}} \mid s(\mathbf{X}) = \mathbf{x} \text{ and } s(\mathbf{N}_{\mathbf{X}}) \in T^-(\mathbf{N}_{\mathbf{X}})\}$

where $\mathbf{N}_{\mathbf{X}}$ is the set of nondescendants of $\mathbf{X}$ in the parenthood graph (in the special case where $\mathbf{N}_{\mathbf{X}} = \emptyset$, we conventionally take the condition $s(\emptyset) \in T^-(\emptyset)$ to be "trivially true"). Note also that the parenthood graph G of T is replaced with its subgraph $G_{\mathbf{X}}$ that omits all arrows going into $\mathbf{X}$ (i.e., (V, Z) is an arrow of $G_{\mathbf{X}}$ iff it is an arrow of G and $Z \notin \mathbf{X}$). The definition above includes as a special case the empty intervention $do()$ corresponding to the empty multiset of atomic formulas. Such an intervention leaves both components of the causal team unchanged.¹²

It is easy to prove from the definitions that the team component of intervened *nonempty* teams can be presented more concretely as the

union $(T_{\mathbf{X}=\mathbf{x}})^- = \bigcup_{s \in T^-} s_{\mathbf{X}=\mathbf{x}}^{\mathcal{F}}$, where each $s_{\mathbf{X}=\mathbf{x}}^{\mathcal{F}}$ is the outcome of the intervention applied to the single assignment s ,¹³ which can be described as follows:

$$\begin{aligned} s_{\mathbf{X}=\mathbf{x}}^{\mathcal{F}} &:= \{t \in \mathbb{A}_{\sigma} \mid \forall X \in \mathbf{X} : t(X) = x \\ &\quad \forall V \in \mathbf{N}_{\mathbf{X}} : t(V) = s(V) \\ &\quad \forall V \in \text{Int}(T) \setminus (\mathbf{X} \cup \mathbf{N}_{\mathbf{X}}) : (t(PA_V), t(V)) \in \mathcal{F}_V\}. \end{aligned}$$

In the case of strictly recursive models, this presentation can be used to generate intervened models via a recursive procedure (analogously to the case of the usual recursive causal models). Indeed, one can define the "distance" of a variable Y from a tuple $\mathbf{X}$ as follows:

$$\begin{cases} d(\mathbf{X}, Y) = 0 & \text{if } Y \in \mathbf{X} \cup \mathbf{N}_{\mathbf{X}} \\ d(\mathbf{X}, Y) = \max\{d(\mathbf{X}, Z) \mid Z \in PA_Y\} + 1 & \text{otherwise.} \end{cases}$$

If $T = (T^-, \mathcal{F})$ is strictly recursive (and the set of variables is finite), then clearly each variable has a finite distance from $\mathbf{X}$, and there is a maximum distance attained; furthermore, if $Z \in PA_Y$, then $d(\mathbf{X}, Z) < d(\mathbf{X}, Y)$. The team $T_{\mathbf{X}=\mathbf{x}}^-$ can then be produced by starting (step 0) from the set T_0 of all assignments t (over $\mathbf{X} \cup \mathbf{N}_{\mathbf{X}}$) such that $t(\mathbf{X}) = \mathbf{x}$ and $t(\mathbf{N}_{\mathbf{X}}) \in T^-(\mathbf{N}_{\mathbf{X}})$. Supposing a team T_n has been defined (over all variables Y with $d(\mathbf{X}, Y) \leq n$), one then produces T_{n+1} by adding to each assignment of T_n all values of the $n+1$ -distance variables that are compatible with $\mathcal{F}$. That is, if we write $\mathbf{Z} = (Z_1, \dots, Z_k)$ for a list of the variables at distance $n+1$ from $\mathbf{X}$, we may let $T_{n+1} := \{t(\mathbf{z}/\mathbf{Z}) \mid t \in T_n, \mathbf{z} \in \mathcal{F}_{Z_1}(t(PA_{Z_1})) \times \dots \times \mathcal{F}_{Z_k}(t(PA_{Z_k}))\}$. Since our assumptions entail the existence of a maximum distance $\hat{n}$, this procedure ends, and we may let $T_{\mathbf{X}=\mathbf{x}}^- := T_{\hat{n}}$.¹⁴ We illustrate this construction in a simple case.

Example 3.1. Consider the following game. I can toss, or not, a coin with my left hand ($L = 0$ or 1); the outcome is recorded by a variable C_L with 3 possible values: heads (h), tails (t) or none (n) in case I do not toss. If I toss and get heads, then I will toss the coin that I hold in my right hand ($R = 1$), whose outcome is recorded analogously in C_R . The parent sets and causal laws are then as follows: $PA_{C_L} = \{L\}$, $PA_R = \{C_L\}$, $PA_{C_R} = \{R\}$; $\mathcal{F}_{C_L} = \mathcal{F}_{C_R} = \{(0, n), (1, h), (1, t)\}$; $\mathcal{F}_R = \{(n, 0), (h, 1), (t, 0)\}$. The current situation is that neither coin has been tossed, which can be represented by a single assignment s :

$L \rightarrow C_L \rightarrow R \rightarrow C_R$
$0 \quad n \quad 0 \quad n$

(the arrow $\rightarrow$ emphasizes that the law for R is deterministic). Let $T = (\{s\}, \mathcal{F})$. To see what would have happened had I decided to toss with my left hand (intervention $do(L = 1)$), we need to update all columns (since all variables are descendants of L):

$L \rightarrow C_L \rightarrow R \rightarrow C_R$	$\sim$	$L \rightarrow C_L \rightarrow R \rightarrow C_R$
$1 \quad \dots \quad \dots \quad \dots$		$1 \quad h \quad \dots \quad \dots$
$1 \quad \dots \quad \dots \quad \dots$		$1 \quad t \quad \dots \quad \dots$

$L \rightarrow C_L \rightarrow R \rightarrow C_R$	$\sim$	$L \rightarrow C_L \rightarrow R \rightarrow C_R$	$\sim$	$T_{L=1}:$
$1 \quad h \quad 1 \quad \dots$		$1 \quad h \quad 1 \quad h$		$1 \quad h \quad 1 \quad h$
$1 \quad t \quad 0 \quad \dots$		$1 \quad t \quad 0 \quad n$		$1 \quad t \quad 0 \quad n$

¹⁰ However, as we shall see, our official definition of intervention will diverge from that used in Barbero & Sandu^[28] and Halpern^[13] when causal cycles are allowed—even in the deterministic case.

¹¹ Judea Pearl seems to have been aware of this problem, as he adopts in Pearl^[31] a definition of intervention close to ours.

¹² This is in contrast with Halpern's semantics, according to which new assignments may appear after an empty intervention (see Barbero & Galliani^[30]). In our case, this does not happen because our interventions leave unchanged the values of all nondescendants of the intervened variables, and in the case of empty interventions, this set of variables is the whole variable domain.

¹³ To see that nonemptiness is a necessary requirement, consider an arbitrary empty model T , say with variables $\mathbf{W}$. Now, $T_{\mathbf{W}=\mathbf{w}}$ is a singleton model (it contains the assignment $s(\mathbf{W}) = \mathbf{w}$), while $\bigcup_{s \in T^-} s_{\mathbf{W}=\mathbf{w}}^{\mathcal{F}}$, being the union of an empty family of sets, is empty.

¹⁴ If some of the causal laws are not total, the procedure might also end before this stage, producing an empty model.

From the modified model, we see, among other things, that if I had tossed the coin in my left hand, then all outcomes for the right-hand coin would have been possible, while outcome n would have been impossible for the left-hand coin.

It must be remarked that, while in Barbero & Galliani^[30] it could be proved that our definition of intervention coincides with that of Halpern in the recursive deterministic case, the same does not hold in the (strictly) recursive indeterministic case.¹⁵

Example 3.2. Consider a model T with four binary variables X, Y, A, B , $PA_Y = X$, $PA_B = A$, laws $\mathcal{F}_Y(x) = \{x\}$ (for all $x \in \text{Ran}(X)$) and $\mathcal{F}_B(a) = \{0, 1\}$ (for all $a \in \text{Ran}(A)$), and the following team component, with a single assignment s :

$X \rightarrow Y$	$A \rightarrow B$
0	0

Note that the parenthood graph, sketched above the table, is acyclic. It is easy to see that, with our definitions, $T_{X=0} = T$. This is due to our requirement that only assignments that agree with s on nondescendants of X be included. If we had just required (in the style of Halpern^[12]) agreement over the non-intervened exogenous variable (A), we would have obtained a larger team:

$X \rightarrow Y$	$A \rightarrow B$
0	0
0	1

3.2 Formal languages

We consider a family of languages that is close to the language in Halpern^[12]. Given a consistent multiset $\mathbf{X} = \mathbf{x}$, we call the expression $[\mathbf{X} = \mathbf{x}]$ a *modal operator*.¹⁶ We also allow the empty multiset, in which case the operator will be written as $\square$. For any given signature $\sigma = (\mathcal{E}, \mathcal{I}, \text{Ran})$, we define a corresponding language $\mathcal{H}_\sigma$. The suffix σ will be omitted when the signature is clear.

Language $\mathcal{H}_\sigma$: $X = x \mid \sim\psi \mid \psi \& \chi \mid [\mathbf{X} = \mathbf{x}]\eta$

where, $\mathbf{X} \cup \{X\} \subseteq \text{Dom}$, $x \in \text{Ran}(X)$, $\mathbf{x} \in \text{Ran}(\mathbf{X})$, and η has no occurrences of modal operators (i.e., it is a Boolean combination of atoms like $X = x$). For simplicity, and without loss of generality, we assume that $\mathbf{X} = \mathbf{x}$ is nonredundant (each variable appears in it at most once); it is then, *a fortiori*, consistent.

Language $\mathcal{H}$ differs from Halpern's in some minor respects: 1) we allow for atomic formulas not prefixed by modal operators,¹⁷ 2) we allow arbitrary variables in atomic formulas and inside the modal operators.¹⁸ This alternative set of conventions is often followed in the philosophical literature, e.g., in Briggs's work^[14].

The restriction on η , which forbids the right-nesting of counterfactuals, comes from the original paper^[11], and has been preserved in large parts of the literature on causal inference. The reasons behind this self-imposed limitation can only be guessed; on one hand, interestingly, in the recursive deterministic case, it is easy to show that any nested counterfactual is equivalent to an unnested one, see Briggs^[14] and Barbero & Sandu^[28]; thus, nested counterfactuals may appear as

superfluous. On the other hand, this reduction is not straightforward if the recursivity constraint is removed, and in that case, the axiomatization of right-nested counterfactuals is a considerably more difficult task than the unnested case.¹⁹ We will tackle this issue and consider, starting from section 5, the language $\mathcal{H}_\sigma^+$ that is obtained by allowing right-nesting of the modal operators. In other words, $\mathcal{H}_\sigma^+$ has the same formal definition as $\mathcal{H}_\sigma$, with the exception that no restriction is put on η . Thus, for example, formulas $[\mathbf{X} = \mathbf{x}][\mathbf{Y} = \mathbf{y}]W = w$ and $[\mathbf{X} = \mathbf{x}](Z = z \& [\mathbf{Y} = \mathbf{y}]W = w)$ are $\mathcal{H}_\sigma^+$ formulas but not $\mathcal{H}_\sigma$ formulas. This type of language is frequently used in the philosophical literature (e.g., Briggs's work^[14]).

The semantics for languages $\mathcal{H}_\sigma$ and $\mathcal{H}_\sigma^+$ is given by the following clauses:

- $T \models X = x$ iff $s(X) = x$ for each $s \in T^-$.
- $T \models \sim\psi$ iff $T \not\models \psi$.
- $T \models \psi \& \chi$ iff $T \models \psi$ and $T \models \chi$.
- $T \models [\mathbf{X} = \mathbf{x}]\psi$ iff for all $s \in T_{\mathbf{X}=\mathbf{x}}^-, (\{s\}, \mathcal{F}) \models \psi$.

Formulas of the form $[\mathbf{X} = \mathbf{x}]\psi$ will be called *counterfactuals*; formulas without occurrences of modal operators are said to be *modal-free*. We can also define the *might-counterfactual* $\langle \mathbf{X} = \mathbf{x} \rangle \psi$ as an abbreviation for $\sim[\mathbf{X} = \mathbf{x}]\sim\psi$; it is easy then to see that the semantics of such formulas is:

- $T \models \langle \mathbf{X} = \mathbf{x} \rangle \psi$ iff there is an $s \in T_{\mathbf{X}=\mathbf{x}}^-$ such that $(\{s\}, \mathcal{F}) \models \psi$.

As a special case, we write $\diamond$ for $\sim\square\sim$. Our definition of intervention yields the following derived semantic clauses:

- $T \models \square\psi$ iff for all $s \in T^-$, $(\{s\}, \mathcal{F}) \models \psi$.
- $T \models \diamond\psi$ iff for some $s \in T^-$, $(\{s\}, \mathcal{F}) \models \psi$.

We can then define a few additional operators:

- $\psi \sqcup \chi$ as $\sim(\sim\psi \& \sim\chi)$
- $\perp$ as $X = x \& \sim X = x$ ²⁰
- $\psi \rightarrow \chi$ as $\sim\psi \sqcup \chi$
- $\top$ as $\sim\perp$
- $\psi \leftrightarrow \chi$ as $(\psi \rightarrow \chi) \& (\chi \rightarrow \psi)$
- $X \neq x$ as $\sim X = x$

Note that $X \neq x$ is just the classical negation of $X = x$, asserting that *at least one assignment* does not assign x to X .²¹

We will use index sets for iterated conjunctions and disjunctions; conventionally, if $I = \emptyset$, $\&_{i \in I} \psi_i$ stands for $\top$ and $\sqcup_{i \in I} \psi_i$ stands for $\perp$.

4 An axiomatization for singleton models

We will see that the general class of singleton models (which also allow for cyclic causation) obeys a set of axioms that is quite similar to that given by Halpern for the general class of *deterministic* causal models^[12]. There is a significant omission: Halpern had an axiom stating that, if you intervene on all (endogenous) variables except one, say Y , then Y will take a single value in the resulting model. This is false in our context, even restricting attention to singleton models, because if Y is internal, an indeterministic law for Y may produce multiple or no values upon intervention.²² On the other hand, we have a principle that describes the fact that the *external* variables are not affected by interventions on *all* other variables. It does so by saying that

¹⁵ We thank Sander Beckers for this insight.

¹⁶ Inside modal operators, we identify strings of symbols that represent the same multiset. This saves us the somewhat trivial issue of axiomatizing this form of equivalence. See Barbero & Sandu^[28] for a complete list of axioms for the antecedents of counterfactuals.

¹⁷ In Halpern^[12], atomic formulas $Y = y$ are treated as abbreviations for $\square Y = y$.

¹⁸ In Halpern^[12], only *endogenous* variables play an active role in the formal languages.

¹⁹ This is illustrated by sections 5 and 6.

²⁰ The reader used to team semantics must be warned that, with the present definition, $\perp$ is not satisfied by empty causal teams. It is satisfied by no model at all.

²¹ This is in contrast with other papers involving causal team semantics.

²² If we look at causal teams in general, external variables also disagree with Halpern's axiom. If Y is external, a causal team may record more than one value for it, in distinct assignments; and many interventions do not make such values disappear.

the available values for Y are the same either before or after such an intervention:

$$\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \leftrightarrow \Diamond Y = y.$$

This axiom scheme needs to be restricted to instances in which Y is external according to the signature. This is the only axiom that requires having the distinction between $\mathcal{E}$ and $\mathcal{I}$ as part of the signature.²³ Halpern's system does not need such an axiom because the external (exogenous) variables do not appear in its syntax.

A second new axiom (*Flatness*) needs to be added to account for the fact that in $\mathcal{H}$ we also have non-modal formulas; it tells us that a formula $\mathbf{Y} = \mathbf{y}$ (which, remember, is a conjunction of atoms) can be converted into modal statements $\Box \mathbf{Y} = \mathbf{y}$; i.e., formula $\mathbf{Y} = \mathbf{y}$ just states that $\mathbf{Y}$ take values $\mathbf{y}$ in all assignments of the model. A third new axiom (*Full intervention*) states that intervening on all variables produces (at least) an assignment; we believe that such an axiom should have already been included in the axiomatizations for general deterministic models in Halpern^[12].²⁴ The fourth new axiom (*Nonemptiness*) is specific for singleton models and states that they contain at least an assignment; it is simply $\Diamond \top$. We remark that Halpern's framework allowed evaluating formulas over assignments that are *not* compatible with the causal laws; $\Diamond \top$ is unsound over such "impossible models", which are excluded here.

For each signature σ we define a Hilbert-style axiom system A_σ consisting of rules and axiom schemes as follows; the symbol A_σ will also often be used just to denote the set of axioms.

$$\text{Rule MP. } \frac{\psi \quad \psi \rightarrow \chi}{\chi} \quad \text{Rule NEC. } \frac{\vdash \psi}{\vdash [\mathbf{X} = \mathbf{x}] \psi} \text{ (if } \psi \text{ modal-free)}$$

- 10. $\mathcal{H}_\sigma$ instances of classical tautologies in $\sim, \&$.
- 11. $[\mathbf{X} = \mathbf{x}] Y = y \rightarrow [\mathbf{X} = \mathbf{x}] Y \neq y'$ (when $y \neq y'$) [Uniqueness]
- 11b. $Y = y \rightarrow Y \neq y'$ [Uniqueness]
- 12. $[\mathbf{X} = \mathbf{x}] \bigcup_{y \in \text{Ran}(Y)} Y = y$ [Definiteness]
- 12b. $\bigcup_{y \in \text{Ran}(Y)} Y = y$ [Definiteness]
- 13. $\langle \mathbf{X} = \mathbf{x} \rangle (Z = z \& \mathbf{Y} = \mathbf{y}) \rightarrow \langle \mathbf{X} = \mathbf{x}, Z = z \rangle \mathbf{Y} = \mathbf{y}$ [Weak composition]
- 14. $[\mathbf{X} = \mathbf{x}, Y = y] Y = y$ [Effectiveness]
- 15. $([\mathbf{X} = \mathbf{x}] \psi \& [\mathbf{X} = \mathbf{x}] (\psi \rightarrow \chi)) \rightarrow [\mathbf{X} = \mathbf{x}] \chi$ [K-axiom]
- 16. $(\langle \mathbf{X} = \mathbf{x}, V = v \rangle (Y = y \& \mathbf{Z} = \mathbf{z}) \& \langle \mathbf{X} = \mathbf{x}, Y = y \rangle (V = v \& \mathbf{Z} = \mathbf{z})) \rightarrow \langle \mathbf{X} = \mathbf{x} \rangle (V = v \& Y = y \& \mathbf{Z} = \mathbf{z})$
(for $V \neq Y$, and $\mathbf{Z} = \text{Dom} \setminus (\mathbf{X} \cup \{V, Y\})$) [Weak reversibility]
- 17. $\mathbf{Y} = \mathbf{y} \leftrightarrow \Box \mathbf{Y} = \mathbf{y}$. [Flatness]
- 18. $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \leftrightarrow \Diamond Y = y$ (if $Y \in \mathcal{E}$) [External variables]
- 19. $\langle \mathbf{W} = \mathbf{w} \rangle \top$ [Full intervention]
- 110. $\Diamond \top$ [Nonemptiness]

In the axioms, $\mathbf{X}$ and $\mathbf{Y}$ are allowed to be empty tuples; in such a case, the modal $[\mathbf{X} = \mathbf{x}]$ reduces to $\Box$, $\langle \mathbf{X} = \mathbf{x} \rangle$ to $\Diamond$, and, e.g., the conjunction $\mathbf{Y} = \mathbf{y}$ is taken to stand for the formula $\top$.

Note that rule NEC is restricted to non-modal ψ ; this is to guarantee that $[\mathbf{X} = \mathbf{x}] \psi$ will be an $\mathcal{H}$ formula. Axioms 10–16 essentially coincide with the part of Halpern's axiomatization (for the general class of causal models) that is meaningful for our language and sound on indeterministic singleton models (we will see that I6 fails in the case of causal teams with multiple assignments).²⁵ We can also see

that schemes I1b and I2b (which can be thought of as no-intervention cases of I1 and I2) characterize the fact that the team contains at least one assignment, resp. at most one assignment.

Lemma 4.1. T is a singleton model iff for all $Y \in \mathcal{E} \cup \mathcal{I}$, 1) $T \models \bigcup_{y \in \text{Ran}(Y)} Y = y$ and 2) for all $y, y' \in \text{Ran}(Y)$ with $y \neq y'$, $T \models Y = y \rightarrow Y \neq y'$.

Furthermore, condition 2) may be replaced by 2'): $T \models \Diamond \top$.

Proof. $\Leftarrow$) Suppose T is empty; then, $T \models Y = y \& Y = y'$, contradicting 2).

Since T satisfies 1), its team component is constant over all variables in the domain; thus, it can contain at most one assignment.

$\Rightarrow$) Straightforward.²⁶

Finally, it is easy to see that $\Diamond \top \equiv Y = y \rightarrow Y \neq y'$. $\square$

It seems to us, however, that these two non-modal formula schemes are insufficient for deriving formally all the modal consequences of the nonemptiness of the model; that is why we add axiom scheme I10.

We shall write $\vdash_{A_\sigma}$ to denote derivability in axiom system A_σ . We shall just write $A, \vdash_A$ when the signature is irrelevant (or even just $\vdash$ if it is clear that we are discussing derivability in A). We will see that A_σ is sound and complete over the class of all singleton models of signature σ .

Theorem 4.2. Axiom system A is sound over singleton models.

Furthermore, all axioms except for I6 and I10 are sound over the class of all models.

Proof. We consider the least intuitive axioms, namely I3 (weak composition), I6 (weak reversibility), I8 (external variables), and I9 (full intervention); the rest is left to the reader.

Axiom I3 (weak composition): $T = (T^-, \mathcal{F}) \models \langle \mathbf{X} = \mathbf{x} \rangle (Z = z \& \mathbf{Y} = \mathbf{y})$ means that there is an $s \in T_{\mathbf{X}=\mathbf{x}}^-$ such that $s(\mathbf{Z}\mathbf{Y}) = z\mathbf{y}$. Now, s is compatible with $\mathcal{F}_{\mathbf{X}=\mathbf{x}, Z=z} \subseteq \mathcal{F}_{\mathbf{X}=\mathbf{x}}$. Furthermore, $s(\mathbf{N}_{\mathbf{X}}) \in T^-(\mathbf{N}_{\mathbf{X}})$, so since $\mathbf{N}_{\mathbf{X}\mathbf{Z}}$ is a subtuple of $\mathbf{N}_{\mathbf{X}}$, $s(\mathbf{N}_{\mathbf{X}\mathbf{Z}}) \in T^-(\mathbf{N}_{\mathbf{X}\mathbf{Z}})$. Then, by definition of intervention, $s \in T_{\mathbf{X}=\mathbf{x}, Z=z}^-$; thus, since $s(\mathbf{Y}) = \mathbf{y}$, $T \models \langle \mathbf{X} = \mathbf{x}, Z = z \rangle \mathbf{Y} = \mathbf{y}$.

Axiom I6 (weak reversibility): Suppose that $T = (\{t\}, \mathcal{F})$ satisfies $\psi : \langle \mathbf{X} = \mathbf{x}, V = v \rangle (Y = y \& \mathbf{Z} = \mathbf{z})$ and $\chi : \langle \mathbf{X} = \mathbf{x}, Y = y \rangle (V = v \& \mathbf{Z} = \mathbf{z})$. By ψ , there is an $s \in T_{\mathbf{X}=\mathbf{x}, V=v}^-$ such that $s(Y) = y$ and $s(\mathbf{Z}) = \mathbf{z}$. But (since $\mathbf{X} \cup \mathbf{Z} \cup \{Y\} \cup \{V\} = \text{Dom}$) χ tells us that s is also in $T_{\mathbf{X}=\mathbf{x}, Y=y}^-$. In particular, if V is internal, s is compatible also with the law $\mathcal{F}_V$; if we prove that $s(\mathbf{N}_{\mathbf{X}}) \in T^-(\mathbf{N}_{\mathbf{X}})$ (i.e., $s(\mathbf{N}_{\mathbf{X}}) = t(\mathbf{N}_{\mathbf{X}})$), then we can conclude that $s \in T_{\mathbf{X}=\mathbf{x}}^-$. But this follows from the fact that $\mathbf{N}_{\mathbf{X}} = \mathbf{N}_{\mathbf{X}Y} \cup \mathbf{N}_{\mathbf{X}V}$, $s(\mathbf{N}_{\mathbf{X}Y}) = t(\mathbf{N}_{\mathbf{X}Y})$ (since $s \in T_{\mathbf{X}=\mathbf{x}, Y=y}^-$ and t is the unique assignment in T^-) and $s(\mathbf{N}_{\mathbf{X}V}) = t(\mathbf{N}_{\mathbf{X}V})$ (for similar reasons). Since $s \in T_{\mathbf{X}=\mathbf{x}}^-$, then, we have $T \models \langle \mathbf{X} = \mathbf{x} \rangle (V = v \& Y = y \& \mathbf{Z} = \mathbf{z})$.

Axiom I8 (external variables): Let Y be an external variable.

Assume further that $T \models \Diamond Y = y$. Then, there is an $s \in T^-$ such that $s(Y) = y$. Since Y is external, by the definition of intervention, the assignment $t(\mathbf{W}_Y) = \mathbf{w}$, $t(Y) = y$ is in $T_{\mathbf{W}_Y=\mathbf{w}}^-$. Thus, $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$.

Vice versa, assume $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$. Then, there is a $t \in T_{\mathbf{W}_Y=\mathbf{w}}^-$ with $t(Y) = y$. But then, since Y is external, by definition of intervention, $y \in T^-(Y)$. Thus, $T \models \Diamond Y = y$.

²³ As shown, e.g., in studies^[14,19,21], in the recursive deterministic case, it is possible to give axiomatizations for simpler signatures that only describe the variable domain and ranges (we describe such signatures in section 10). We leave it as an open question whether this is possible in the indeterministic context, or even in the deterministic context with cycles.

²⁴ It is difficult to establish the correctness of this claim unequivocally, not only because it is difficult to ascertain whether *Full intervention* is derivable from Halpern's system but also because of ambiguities in the definition of intervention in Halpern^[12] that make it difficult to decide whether this axiom is sound or not. However, in a personal communication, Halpern stated that his intended semantics for general deterministic models is the one presented later in Halpern^[13], section 2.7; according to this semantics, the *Full intervention* axiom is clearly sound. Halpern agreed that the axiom should be added, and he himself did so in later publications such as studies^[17,20].

²⁵ Axiom I3 is also analogous to the principle of *Cautious monotonicity* in the field of nonmonotonic logic^[29].

²⁶ It is a special case of the forthcoming soundness theorem 4.2.

Axiom I9 (full intervention): Let $T = (T^-, \mathcal{F})$ be a model. Since $\mathcal{F}_{W=w} = \emptyset$, the assignment s with $s(W) = w$ is compatible with $\mathcal{F}_{W=w}$. Furthermore, in the parenthood graph of $\mathcal{F}_{W=w}$ the set N_W of non-descendants of W is empty; thus, the condition $s(N_W) \in T^-(N_W)$ is trivially satisfied by s . Thus, by the definition of intervention, $s \in (T_{W=w})^-$. Thus, $T_{W=w} \models \Diamond \top$, so $T \models \langle W = w \rangle \top$. $\square$

In general, models can have empty teams, so that explains why I10 does not hold for all models. The following is a minimal example of a model that falsifies weak reversibility (axiom I6).²⁰

Example 4.3. Consider a model T with two binary external variables, L and R , say representing whether a switch on the left (resp. a switch on the right) has been pressed. We know that exactly one switch has been pressed. We thus have the following team:

L	R
1	0
0	1

Intervening to press either the left or the right switch will produce the following teams:

$T_{L=1}:$

L	R
I	O
I	I

$T_{R=1}:$

L	R
I	I
O	I

From the tables, it is immediate to see that $T \models \langle L = 1 \rangle R = 1$, $T \models \langle R = 1 \rangle L = 1$ but $T \not\models \langle L = 1 \ \& \ R = 1 \rangle$. Thus, the simplest instance of axiom I6 is falsified.

The following lemma lists basic properties of $\mathcal{A}$ (proofs can be found in the [Supplementary File 1](#)). We note that these are essentially results in the modal logic K , as their proofs only use MP, NEC, and axioms I0 (instances of tautologies) and I5 (K-axiom).

Lemma 4.4. The following hold in $\mathcal{A}$.

1. (Deduction theorem) If $\Gamma, \psi \vdash \chi$, then $\Gamma \vdash \psi \rightarrow \chi$.
2. ($\Box$ -Monotonicity) If $\Gamma \vdash [X = x]\psi$, $\vdash \psi \rightarrow \psi'$ and $[X = x]\psi' \in \mathcal{H}$, then $\Gamma \vdash [X = x]\psi'$.
3. ($\Diamond$ -Monotonicity) If $\Gamma \vdash \langle X = x \rangle \psi$, $\vdash \psi \rightarrow \psi'$ and $\langle X = x \rangle \psi' \in \mathcal{H}$, then $\Gamma \vdash \langle X = x \rangle \psi'$.
4. $\vdash ([X = x]\psi \ \& \ [X = x]\chi) \leftrightarrow [X = x](\psi \ \& \ \chi)$.
5. (Replacement) Suppose $\Gamma \vdash \theta \leftrightarrow \theta'$. Then $\Gamma \vdash \varphi \leftrightarrow \varphi[\theta'/\theta]$, provided this is an $\mathcal{H}$ formula.
6. $\vdash \sim [X = x]\psi \leftrightarrow \langle X = x \rangle \sim \psi$
7. $\vdash \sim \langle X = x \rangle \psi \leftrightarrow [X = x] \sim \psi$
8. $\vdash (\langle X = x \rangle \psi \ \sqcup \ \langle X = x \rangle \chi) \leftrightarrow \langle X = x \rangle (\psi \ \sqcup \ \chi)$.
9. $\vdash ([X = x]\psi \ \& \ [X = x]\chi) \rightarrow [X = x](\psi \ \& \ \chi)$.
10. $\vdash ([X = x]\psi \ \& \ \langle X = x \rangle \chi) \rightarrow \langle X = x \rangle (\psi \ \& \ \chi)$.
11. $\vdash \langle X = x \rangle (\psi \ \& \ \chi) \rightarrow (\langle X = x \rangle \psi \ \& \ \langle X = x \rangle \chi)$.

We say that a set of $\mathcal{H}_\sigma$ formulas is (A_σ) -consistent if $\Gamma \not\vdash_{A_\sigma} \perp$. Γ is **maximally consistent** if it is consistent and, furthermore, if $\Gamma' \supseteq \Gamma$ is consistent, then $\Gamma' = \Gamma$. We will later consider also a proof system B for the $\mathcal{H}^+$ languages. For both systems, the following classical results hold:

Lemma 4.5. Let Γ be a maximally consistent set of $\mathcal{H}_\sigma$ formulas (or $\mathcal{H}_\sigma^+$ formulas). Then:

1. Closure under $\vdash$: if $\Gamma \vdash \psi$, then $\psi \in \Gamma$.
2. Completeness: for every formula ψ , either ψ or $\sim \psi$ is in Γ .
3. Closure under $\&$: if $\psi, \chi \in \Gamma$, then $\psi \ \& \ \chi \in \Gamma$.
4. Primality: if $\psi \ \sqcup \ \chi \in \Gamma$, then $\psi \in \Gamma$ or $\chi \in \Gamma$.
5. Impossibility of contradiction: $\Diamond \perp \notin \Gamma$.

Lemma 4.6 (Lindenbaum). Any consistent set Δ of $\mathcal{H}_\sigma$ formulas (resp. $\mathcal{H}_\sigma^+$ formulas) can be extended to a maximal consistent set.

The canonical relational causal team associated to a maximal consistent set Γ (of formulas of a fixed signature $\sigma = (\mathcal{E}, \mathcal{I}, \text{Ran})$), which we shall denote as $\mathbb{T}^\Gamma = ((\mathbb{T}^\Gamma)^-, \mathcal{F}^\Gamma)$, is defined as follows:

- We let $(\mathbb{T}^\Gamma)^- := \{s \in A_\sigma \mid \Diamond W = s(W) \in \Gamma\}$.
- For each pair of variables X, Y with $Y \in \mathcal{I}$ we let $X \in PA_Y$ if and only if there are $w, y, x \in \text{Ran}(W_{XY}, Y, X)$ such that either of the following holds:
 1. $\langle W_{XY} = w \rangle Y = y \ \& \ \sim \langle W_{XY} = w, X = x \rangle Y = y \in \Gamma$
 2. $\langle W_{XY} = w, X = x \rangle Y = y \ \& \ \sim \langle W_{XY} = w \rangle Y = y \in \Gamma$
- In case $Y \in \mathcal{I}$, we let $(pa, y) \in \mathcal{F}_Y^\Gamma$ if and only if, for any $w \in \text{Ran}(W_Y)$ such that $w \upharpoonright_{PA_Y} = pa$, $\langle W_Y = w \rangle Y = y \in \Gamma$.

Lemma 4.7 (Canonical team). Let Γ be a maximal consistent set of $\mathcal{H}_\sigma$ formulas that contains all axioms of A_σ . Then $\mathbb{T}^\Gamma$ is a singleton model, i.e.,

1. For all $Y \in \text{Int}(\mathbb{T}^\Gamma)$, $\mathcal{F}_Y^\Gamma$ is well-defined.
2. For all $Y \in \text{Int}(\mathbb{T}^\Gamma)$ and $s \in (\mathbb{T}^\Gamma)^-$, $(s(PA_Y), s(Y)) \in \mathcal{F}_Y^\Gamma$.
3. $(\mathbb{T}^\Gamma)^-$ is a singleton.

Proof. 1) We need to show that the specific choice of a w extending pa does not matter for the definition of $\mathcal{F}_Y^\Gamma$. In other words, we need to prove that, if $w, w' \in \text{Ran}(W_Y)$ are such that $w \upharpoonright_{PA_Y} = w' \upharpoonright_{PA_Y} = pa$, then

$$\langle W_Y = w \rangle Y = y \in \Gamma \text{ if and only if } \langle W_Y = w' \rangle Y = y \in \Gamma.$$

We prove this by induction on the number n of variables on which w, w' differ. If $n = 0$, the statement is trivial. If $n = 1$, w, w' differ over a single variable $X \notin PA_Y$. Write w^* for $w \upharpoonright_{W_{XY}} = w' \upharpoonright_{W_{XY}}$. Assume $\langle W_Y = w \rangle Y = y \in \Gamma$. Since $X \notin PA_Y$, by the definition of PA_Y in the canonical model, clause 2., we then have that $\sim \langle W_{XY} = w^* \rangle Y = y \notin \Gamma$. But then, by Lemma 4.5, 2., $\langle W_{XY} = w^* \rangle Y = y \in \Gamma$. Again, since $X \notin PA_Y$, by clause 1., $\sim \langle W_Y = w' \rangle Y = y \notin \Gamma$; so, similarly as before, we conclude $\langle W_Y = w' \rangle Y = y \in \Gamma$. The converse is analogous.

Now assume that the statement holds for some $n \geq 1$, and suppose w, w' differ on $n+1$ variables $X_1, \dots, X_{n+1}$ (which take values $x_1, \dots, x_{n+1}$ in w and values $x'_1, \dots, x'_{n+1}$ in w'). Write Z for $W_Y \setminus \{X_1, \dots, X_{n+1}\}$ and z for $w \upharpoonright_Z$. By the inductive hypothesis (for case n), we have:

$$\langle W_Y = w \rangle Y = y \in \Gamma \text{ if and only if } \langle ZX_1 \dots X_n X_{n+1} = zx'_1 \dots x'_{n+1} \rangle Y = y \in \Gamma$$

Since $zx'_1 \dots x'_{n+1}$ and w' differ only on one variable (X_{n+1}), by the base case proved above we obtain the following equivalence: $\langle ZX_1 \dots X_n X_{n+1} = zx'_1 \dots x'_{n+1} \rangle Y = y \in \Gamma$ iff $\langle W_Y = w' \rangle Y = y \in \Gamma$.

2) Let $s \in (\mathbb{T}^\Gamma)^-$. By definition of $(\mathbb{T}^\Gamma)^-$, $\Diamond W = s(W) \in \Gamma$. Thus, by axiom I3, MP, and Lemma 4.5, 1., $\langle W_Y = s(W_Y) \rangle Y = s(Y) \in \Gamma$. Since the restriction of $s(W_Y)$ to PA_Y is $s(PA_Y)$, then by definition of $\mathcal{F}^\Gamma$ we have $(s(PA_Y), s(Y)) \in \mathcal{F}_Y^\Gamma$.

3) Let us first prove that $(\mathbb{T}^\Gamma)^-$ is nonempty. By axiom I2b, $\sqcup_{w \in \text{Ran}(W)} W = w \in \Gamma$. By the primality of Γ (Lemma 4.5, 4.), there is a $w^* \in \text{Ran}(W)$ such that $W = w^* \in \Gamma$. By I7, $\Box W = w^* \in \Gamma$. Since, by I10, $\Diamond \top \in \Gamma$, by Lemma 4.4, 9., we obtain $\Diamond W = w^* \in \Gamma$. But then, by definition, the assignment s with $s(W) = w^*$ is in $(\mathbb{T}^\Gamma)^-$.

Suppose for the sake of contradiction that there are two distinct assignments, s_1 and s_2 , both contained in $(\mathbb{T}^\Gamma)^-$. Then, by definition of $(\mathbb{T}^\Gamma)^-$ both $\Diamond W = s_1(W) \in \Gamma$ and $\Diamond W = s_2(W) \in \Gamma$. As proved above, $\Box W = w^* \in \Gamma$ for some $w^* \in \text{Ran}(W)$. Now, $s_1(W)$ and $s_2(W)$ cannot be both equal to w^* ; wlog, suppose $s_2(W) \neq w^*$; in particular, there is a variable $Y \in W$ such that $s_2(Y) \neq y^* = w^* \upharpoonright_Y$. By axiom I1b $\vdash Y = y^* \rightarrow \sim Y = s_2(Y)$, and by I0, $\vdash W = w^* \rightarrow Y = y^*$. Thus, by two applications of monotonicity (Lemma 4.4, 2.), we obtain (*): $\Box \sim Y = s_2(Y) \in \Gamma$. On the other hand, since $\vdash W = s_2(W) \rightarrow Y = s_2(Y)$,

²⁰ We thank an anonymous reviewer for noticing the failure of weak reversibility and illustrating it with a more complex counterexample.

by monotonicity (Lemma 4.4, 3.) we have $\Diamond Y = s_2(Y) \in \Gamma$. Thus, by Lemma 4.4, 10., $\Diamond(Y = s_2(Y) \ \& \ \sim Y = s_2(Y)) \in \Gamma$. By monotonicity again, we obtain $\Diamond \perp \in \Gamma$, contradicting Lemma 4.5, 5. $\square$

In order to prove a truth lemma, we need first to show that $\mathcal{H}$ formulas can be put, up to equivalence, in a simple normal form, similarly as in Halpern^[12].

Lemma 4.8. *Let $w, w' \in \text{Ran}(W)$, $w \neq w'$. Then,*

1. $\vdash \langle \mathbf{X} = \mathbf{x}, W = w \rangle (W = w \ \& \ \chi) \leftrightarrow \langle \mathbf{X} = \mathbf{x}, W = w \rangle \chi$.

2. $\vdash \langle \mathbf{X} = \mathbf{x}, W = w \rangle (W = w' \ \& \ \chi) \leftrightarrow \langle \mathbf{X} = \mathbf{x}, W = w \rangle \perp$.

Proof. 1., $\Rightarrow$ By I0, $\vdash (W = w \ \& \ \chi) \rightarrow \chi$. Thus, by monotonicity (Lemma 4.4, 3.), $\vdash \langle \mathbf{X} = \mathbf{x}, W = w \rangle (W = w \ \& \ \chi) \rightarrow \langle \mathbf{X} = \mathbf{x}, W = w \rangle \chi$.

1., $\Leftarrow$ Assume $\sim \langle \mathbf{X} = \mathbf{x}, W = w \rangle (W = w \ \& \ \chi)$, i.e., $\sim \sim [\mathbf{X} = \mathbf{x}, W = w] \sim (W = w \ \& \ \chi)$. By I0 we get $[\mathbf{X} = \mathbf{x}, W = w] \sim (W = w \ \& \ \chi)$. On the other hand, by I4, $[\mathbf{X} = \mathbf{x}, W = w] W = w$. Thus, by Lemma 4.4, 4., $[\mathbf{X} = \mathbf{x}, W = w] (W = w \ \& \ \sim (W = w \ \& \ \chi))$. Then, by I0 and monotonicity, $[\mathbf{X} = \mathbf{x}, W = w] \sim \chi$. By I0, $\sim \sim [\mathbf{X} = \mathbf{x}, W = w] \sim \chi$, which is the same as $\sim \langle \mathbf{X} = \mathbf{x}, W = w \rangle \chi$.

2., $\Leftarrow$ By I0 and monotonicity.

2., $\Rightarrow$ By axiom I4, $\vdash [\mathbf{X} = \mathbf{x}, W = w] W = w$. Then, by I1, $\vdash [\mathbf{X} = \mathbf{x}, W = w] \sim W = w'$. Together with $\langle \mathbf{X} = \mathbf{x}, W = w \rangle (W = w' \ \& \ \chi)$, by Lemma 4.4, 10., we obtain $\langle \mathbf{X} = \mathbf{x}, W = w \rangle (\sim W = w' \ \& \ (W = w' \ \& \ \chi))$. Then, by I0 and replacement, $\langle \mathbf{X} = \mathbf{x}, W = w \rangle \perp$. $\square$

Lemma 4.9.

$\vdash [\mathbf{X} = \mathbf{x}] \varphi \leftrightarrow [\mathbf{X} = \mathbf{x}] \varphi [\sqcup_{y' \in \text{Ran}(Y) \setminus \{y\}} Y = y' / \sim Y = y]$.

Proof. By replacement, it suffices to prove $\vdash (\sqcup_{y' \in \text{Ran}(Y) \setminus \{y\}} Y = y') \leftrightarrow \sim Y = y$.

For the left-to-right direction, using I0 we have $\&_{y' \in \text{Ran}(Y) \setminus \{y\}} (Y = y' \rightarrow \sim Y = y) \rightarrow [(\sqcup_{y' \in \text{Ran}(Y) \setminus \{y\}} Y = y') \rightarrow \sim Y = y]$. But, by axiom I1b, for each $y' \neq y$ we have $\vdash Y = y' \rightarrow \sim Y = y$; thus, $\vdash \&_{y' \in \text{Ran}(Y) \setminus \{y\}} (Y = y' \rightarrow \sim Y = y)$ by I0. Finally, we obtain $(\sqcup_{y' \in \text{Ran}(Y) \setminus \{y\}} Y = y') \rightarrow \sim Y = y$ by MP.

From right to left, assume $\sim Y = y$. By axiom I2b, we obtain $\sqcup_{y' \in \text{Ran}(Y)} Y = y'$. By I0, then, we have $\sim Y = y \ \& \ \sqcup_{y' \in \text{Ran}(Y)} Y = y'$; but then, again by I0 and MP, we obtain $\sqcup_{y' \in \text{Ran}(Y) \setminus \{y\}} Y = y'$. $\square$

Lemma 4.10. *Consider a formula $\varphi : [\mathbf{X} = \mathbf{x}] \sqcup_{i \in I} \&_{j \in J_i} \chi_i^j$ and let φ' be obtained by replacing a disjunct $\&_{j \in J_{i^*}} \chi_{i^*}^j$ with $(\sqcup_{z \in \text{Ran}(Z)} Z = z) \ \& \ \&_{j \in J_{i^*}} \chi_{i^*}^j$. Then, $\vdash \varphi \leftrightarrow \varphi'$.*

Proof. By I2b, $\vdash \sqcup_{z \in \text{Ran}(Z)} Z = z$. By I0, $\vdash (\sqcup_{z \in \text{Ran}(Z)} Z = z) \rightarrow (\&_{j \in J_{i^*}} \chi_{i^*}^j \leftrightarrow [(\sqcup_{z \in \text{Ran}(Z)} Z = z) \ \& \ \&_{j \in J_{i^*}} \chi_{i^*}^j])$. Thus, $\vdash \&_{j \in J_{i^*}} \chi_{i^*}^j \leftrightarrow [(\sqcup_{z \in \text{Ran}(Z)} Z = z) \ \& \ \&_{j \in J_{i^*}} \chi_{i^*}^j]$ by MP. Then, we obtain the statement by replacement. $\square$

Lemma 4.11 (Normal form). *Every $\mathcal{H}_\sigma$ formula φ is provably equivalent to a Boolean combination of formulas $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = \mathbf{y}$, where $\mathbf{X} \cap \mathbf{Y} = \emptyset$ and $\mathbf{X} \cup \mathbf{Y} = \text{Dom}$.*

Proof. First, observe that all subformulas of the form $\mathbf{Z} = \mathbf{z}$ that do not occur in the scope of a modal operator can be replaced, by axiom I7 and replacement, with $\Box \mathbf{Z} = \mathbf{z}$. Second, if φ has subformulas of the form $[\mathbf{X} = \mathbf{x}] \psi$, by I0 and replacement they can be replaced with $\sim \sim [\mathbf{X} = \mathbf{x}] \sim \sim \psi$, i.e., $\sim \langle \mathbf{X} = \mathbf{x} \rangle \sim \psi$. So, φ is equivalent to a Boolean combination of formulas of the form $\theta : \langle \mathbf{X} = \mathbf{x} \rangle \psi$ (where $\mathbf{X}$ might also be an empty tuple).

Next, by I0 and replacement such ψ can be rewritten in disjunctive normal form, i.e., $\sqcup_{i \in I} \&_{j \in J_i} \chi_i^j$, where each χ_i^j is either of the form $Y = y$ or $\sim Y = y$.

Now, by Lemma 4.9, each conjunct of the form $\sim Y = y$ can be rewritten as $\sqcup_{y' \in \text{Ran}(Y) \setminus \{y\}} Y = y'$. Furthermore, if a disjunct $\&_{j \in J_i} \chi_i^j$ does not mention some variable $Z \in \text{Dom} \setminus \mathbf{X}$, by Lemma 4.10 we can add to it a conjunct of the form $\sqcup_{z \in \text{Ran}(Z)} Z = z$.

Applying distributivity of $\&$ over $\sqcup$ (i.e., I0 plus replacement), we transform θ into a formula of the form $\langle \mathbf{X} = \mathbf{x} \rangle \sqcup_{i \in I'} \mathbf{Y} = \mathbf{y}_i$, where $\mathbf{X} \cup \mathbf{Y} = \text{Dom}$. Then, by Lemma 4.4, 8., we rewrite it as $\sqcup_{i \in I'} \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = \mathbf{y}_i$. We can, finally, eliminate all the conjuncts (within the subformulas $\mathbf{Y} = \mathbf{y}_i$) that contain variables of $\mathbf{X}$ by using Lemma 4.8 (plus replacement). $\square$

Lemma 4.12 (Truth lemma). *Let $\Gamma \supseteq A$ be a maximally consistent set of $\mathcal{H}_\sigma$ formulas, and φ an $\mathcal{H}_\sigma$ formula. Then, $\varphi \in \Gamma \iff \mathbb{T}^\Gamma \models \varphi$.*

Proof. By Lemma 4.11, Lemma 4.5, 1. and soundness, we can assume that φ is a Boolean combination of formulas of the form $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = \mathbf{y}$, with $\mathbf{X} \cap \mathbf{Y} = \emptyset$ and $\mathbf{X} \cup \mathbf{Y} = \text{Dom}$.

We proceed by induction on φ . We shall refer to the inductive hypothesis as "i.h."

Case φ is $\psi \ \& \ \chi$. Then, $\mathbb{T}^\Gamma \models \psi \ \& \ \chi$ iff $\mathbb{T}^\Gamma \models \psi$ and $\mathbb{T}^\Gamma \models \chi$ iff (i.h.) $\psi \in \Gamma$ and $\chi \in \Gamma$, iff (by Lemma 4.5, 3.) $\psi \ \& \ \chi \in \Gamma$.

Case φ is $\sim \psi$. Then, $\mathbb{T}^\Gamma \models \sim \psi$ iff $\mathbb{T}^\Gamma \not\models \psi$ iff (i.h.) $\psi \notin \Gamma$ iff (Lemma 4.5, 2.) $\sim \psi \in \Gamma$.

Case φ is $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = \mathbf{y}$. We proceed by a subinduction on $n = |\text{Dom} \setminus \mathbf{X}| = |\mathbf{Y}|$.

• Case $n = 0$. This is straightforward: we both have $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x} \rangle \top$ by the definition of intervention, and $\langle \mathbf{X} = \mathbf{x} \rangle \top \in \Gamma$ by axiom I9.

• Case $n = 1$. Suppose $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = y \in \Gamma$. If Y is internal in $\mathbb{T}^\Gamma$, by the definition of $\mathcal{F}_Y^\Gamma$, we have $(\mathbf{x} \upharpoonright_{PA_Y}, y) \in \mathcal{F}_Y^\Gamma$. Thus, by definition of intervention, $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = y$.

If instead Y is external in $\mathbb{T}^\Gamma$, axiom I8 yields $\Diamond Y = y \in \Gamma$. By I2 and Lemma 4.4, 10., we obtain $\Diamond(\sqcup_{\mathbf{x}' \in \text{Ran}(\mathbf{X})} \langle \mathbf{X} = \mathbf{x}' \rangle \ \& \ Y = y) \in \Gamma$. By I0 and replacement, $\Diamond \sqcup_{\mathbf{x}' \in \text{Ran}(\mathbf{X})} \langle \mathbf{X} = \mathbf{x}' \rangle \ \& \ Y = y \in \Gamma$. By Lemma 4.5, 4., there is an $\mathbf{x}^* \in \text{Ran}(\mathbf{X})$ such that $\Diamond(\langle \mathbf{X} = \mathbf{x}^* \rangle \ \& \ Y = y) \in \Gamma$. By definition of $\mathbb{T}^\Gamma$, there is an $s \in (\mathbb{T}^\Gamma)^-$ (namely $s(\mathbf{X}\mathbf{Y}) = \mathbf{x}^*y$). Since Y is external, $\mathcal{F}_{\mathbf{X}=\mathbf{x}}$ is empty; thus, by definition of intervention, there is a $t \in \mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma$ with $t(Y) = y$. Thus, $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = y$.

Vice versa, suppose $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = y$. Then, again, we have two cases: Y is internal or external. In the former case, by definition of intervention, $(\mathbf{x} \upharpoonright_{PA_Y}, y) \in \mathcal{F}_Y^\Gamma$. But then, by definition of $\mathcal{F}_Y^\Gamma$, and Lemma 4.7, 1., this means that $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = y \in \Gamma$.

Suppose Y is external; then $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = y$ entails $\mathbb{T}^\Gamma \models \Diamond Y = y$. There is then an $s \in (\mathbb{T}^\Gamma)^-$ such that $s(Y) = y$; write $\mathbf{x}^*$ for $s(\mathbf{X})$. Then $\mathbb{T}^\Gamma \models \Diamond(\langle \mathbf{X} = \mathbf{x}^* \rangle \ \& \ Y = y)$. By definition of $(\mathbb{T}^\Gamma)^-$, $\Diamond(\langle \mathbf{X} = \mathbf{x}^* \rangle \ \& \ Y = y) \in \Gamma$. By I0 and monotonicity, then, $\Diamond Y = y \in \Gamma$. Thus, by axiom I8, $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = y \in \Gamma$.

• Case $n > 1$. Suppose $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = \mathbf{y} \in \Gamma$. Let $Y_1, Y_2 \in \mathbf{Y}$, $Y_1 \neq Y_2$. Let $y_1 = \mathbf{y} \upharpoonright_{Y_1}$ and $y_2 = \mathbf{y} \upharpoonright_{Y_2}$, and define:

$$\mathbf{Y}' = \mathbf{Y} \setminus \{Y_1\} \quad \mathbf{Y}'' = \mathbf{Y} \setminus \{Y_2\} \quad \mathbf{Y}^\circ = \mathbf{Y}' \cap \mathbf{Y}''$$

$$y' = y \setminus \{y_1\} \quad y'' = y \setminus \{y_2\} \quad y^\circ = y' \cap y''$$

From $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = \mathbf{y} \in \Gamma$, by I3 (composition) we obtain $\langle \mathbf{X} = \mathbf{x}, Y_1 = y_1 \rangle \mathbf{Y}' = y' \in \Gamma$ and $\langle \mathbf{X} = \mathbf{x}, Y_2 = y_2 \rangle \mathbf{Y}'' = y'' \in \Gamma$; by the inductive hypothesis, then, $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x}, Y_1 = y_1 \rangle \mathbf{Y}' = y'$ and $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x}, Y_2 = y_2 \rangle \mathbf{Y}'' = y''$. By the soundness of I6, then, $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x} \rangle (Y_1 = y_1 \ \& \ Y_2 = y_2 \ \& \ \mathbf{Y}^\circ = y^\circ)$, i.e., $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = \mathbf{y}$.

Vice versa, assume $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = \mathbf{y}$. With the same notations as above, by the soundness of I3 we obtain $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x}, Y_1 = y_1 \rangle \mathbf{Y}' = y'$ and $\mathbb{T}^\Gamma \models \langle \mathbf{X} = \mathbf{x}, Y_2 = y_2 \rangle \mathbf{Y}'' = y''$. By i.h., $\langle \mathbf{X} = \mathbf{x}, Y_1 = y_1 \rangle \mathbf{Y}' = y' \in \Gamma$ and $\langle \mathbf{X} = \mathbf{x}, Y_2 = y_2 \rangle \mathbf{Y}'' = y'' \in \Gamma$. Then, by axiom I6, $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{Y} = \mathbf{y} \in \Gamma$. $\square$

We write $\Gamma \models_\sigma^1 \varphi$ if every singleton model of signature σ that satisfies Γ also satisfies φ .

Theorem 4.13 (Strong completeness of A). *For $\Gamma \cup \{\varphi\} \subseteq \mathcal{H}_\sigma$,*

$$\Gamma \models_{\sigma}^1 \varphi \iff \Gamma \vdash_{A_{\sigma}} \varphi.$$

Proof. The soundness direction is given by Theorem 4.2. For completeness, suppose $\Gamma \not\models \varphi$. Then, by routine reasoning $\Gamma \cup A \cup \{\sim\varphi\}$ is consistent, so by Lemma 4.6, there is a maximally consistent $\Delta \supseteq \Gamma \cup A \cup \{\sim\varphi\}$. Now, $\mathbb{T}^{\Delta}$ is a singleton model (Lemma 4.7), and by Lemma 4.12, $\mathbb{T}^{\Delta} \models \Gamma \cup \{\sim\varphi\}$. Thus, $\Gamma \not\models_{\sigma}^1 \varphi$. $\square$

5 Definability of law components

Toward the formulation of a complete proof system for $\mathcal{H}^+$ over the general class of models, we need some additional concepts and terminology, which we introduce here.

Note that any fixed signature σ allows finitely many possible law components $\mathcal{F}$; we will denote as $\mathbb{R}_{\sigma}$ the set of these objects. The subtleties of formalization described in section 2 suggest that it might be difficult to talk about the law components *within language* $\mathcal{H}^+$, in particular, it might be difficult to tell apart law components that only disagree about the parent sets of some variables. We introduce a notion of equivalence of law components to take this obstacle into account. When discussing several function components, say $\mathcal{F}, \mathcal{G}$, we will differentiate between their parent sets by using indexed notation such as $PA_V^{\mathcal{F}}$, resp. $PA_V^{\mathcal{G}}$.

Definition 5.1. Let $\mathcal{F}, \mathcal{G}$ be two law components of signature $\sigma = (\mathcal{E}, \mathcal{I}, \text{Ran})$. We say that $\mathcal{F}$ and $\mathcal{G}$ are **equivalent**, $\mathcal{F} \equiv \mathcal{G}$ if, for all $V \in \mathcal{I}$ and all $pa \in \text{Ran}(PA_V^{\mathcal{F}} \cup PA_V^{\mathcal{G}})$, $\mathcal{F}_V(pa_{\upharpoonright PA_V^{\mathcal{F}}}) = \mathcal{G}_V(pa_{\upharpoonright PA_V^{\mathcal{G}}})$.

Two models $S = (S^-, \mathcal{F})$ and $T = (T^-, \mathcal{G})$ of signature σ are **equivalent**, $S \equiv T$, if $S^- = T^-$ and $\mathcal{F} \equiv \mathcal{G}$.

The next theorem shows that this notion of equivalence does indeed characterize all that can be said about law components in language $\mathcal{H}^+$ (as well as in $\mathcal{H}$).

Theorem 5.1. Let $S = (S^-, \mathcal{F}), T = (T^-, \mathcal{G})$ be two models of signature $\sigma = (\mathcal{E}, \mathcal{I}, \text{Ran})$. Then, $S \equiv T$ iff for all $\varphi \in \mathcal{H}_{\sigma}^+$, $S \models \varphi \iff T \models \varphi$ iff for all $\varphi \in \mathcal{H}_{\sigma}$, $S \models \varphi \iff T \models \varphi$.

Proof. 1) First, we prove that $S \equiv T$ implies that S and T agree over all $\mathcal{H}_{\sigma}^+$ formulas. We proceed by induction on φ ; we prove the statement simultaneously for all pairs of equivalent teams of arbitrary signatures. The atomic case (φ is $Y = y$) immediately follows from the fact that S and T have the same team component. The cases for φ of the forms $\sim\psi$ and $\psi \& \chi$ are straightforward.

If $S \models \varphi = [\mathbf{X} = \mathbf{x}]\psi$, then, for all $s \in S_{\mathbf{X}=\mathbf{x}}^-$, $(\{s\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}) \models \psi$. Since $\mathcal{F} \equiv \mathcal{G}$, we have $\mathcal{F}_V(pa_{\upharpoonright PA_V^{\mathcal{F}}}) = \mathcal{G}_V(pa_{\upharpoonright PA_V^{\mathcal{G}}})$ for all $V \in \mathcal{I} \setminus \mathbf{X}$. Since $V \in \mathcal{I} \setminus \mathbf{X}$ is the set of internal variables of $\mathcal{F}_{\mathbf{X}=\mathbf{x}}$ and $\mathcal{G}_{\mathbf{X}=\mathbf{x}}$, and for all such variables $(\mathcal{F}_{\mathbf{X}=\mathbf{x}})_V = \mathcal{F}_V$ and $(\mathcal{G}_{\mathbf{X}=\mathbf{x}})_V = \mathcal{G}_V$, then, $\mathcal{F}_{\mathbf{X}=\mathbf{x}} \equiv \mathcal{G}_{\mathbf{X}=\mathbf{x}}$. Thus, for each $s \in S_{\mathbf{X}=\mathbf{x}}^-$, $(\{s\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}) \equiv (\{s\}, \mathcal{G}_{\mathbf{X}=\mathbf{x}})$. By the inductive hypothesis, then, for all $s \in (S_{\mathbf{X}=\mathbf{x}})^-$ we have $(\{s\}, \mathcal{G}_{\mathbf{X}=\mathbf{x}}) \models \psi$. If we manage to prove that $S_{\mathbf{X}=\mathbf{x}}^- = T_{\mathbf{X}=\mathbf{x}}^-$, then, we are done.

It suffices to show that $S_{\mathbf{X}=\mathbf{x}}^- \subseteq T_{\mathbf{X}=\mathbf{x}}^-$; the converse proof is symmetric. Suppose $s \in S_{\mathbf{X}=\mathbf{x}}^-$; by definition of intervention, this means that $s(\mathbf{N}_{\mathbf{X}}) \in S^-(\mathbf{N}_{\mathbf{X}})$, $s(\mathbf{X}) = \mathbf{x}$ and s satisfies the constraints given by $\mathcal{F}_{\mathbf{X}=\mathbf{x}}$. The latter means that, for all $V \in \mathcal{I} \setminus \mathbf{X}$, $(s(PA_V^{\mathcal{F}}), s(V)) \in \mathcal{F}_V$. In other words, $s(V) \in \mathcal{F}_V(s(PA_V^{\mathcal{F}})) = \mathcal{G}_V(s(PA_V^{\mathcal{G}}))$, where the equality is given by $\mathcal{F} \equiv \mathcal{G}$. Thus $(s(PA_V^{\mathcal{G}}), s(V)) \in \mathcal{G}_V$ for all $V \in \mathcal{I} \setminus \mathbf{X}$; together with $s(\mathbf{N}_{\mathbf{X}}) \in S^-(\mathbf{N}_{\mathbf{X}}) = T^-(\mathbf{N}_{\mathbf{X}})$ and $s(\mathbf{X}) = \mathbf{x}$, this yields $s \in T_{\mathbf{X}=\mathbf{x}}^-$.

2) We now deduce $S \equiv T$ from the assumption that S and T agree on the truth values of $\mathcal{H}_{\sigma}$ formulas. Since S and T satisfy the same formulas of the form $\diamond \mathbf{W} = \mathbf{w}$, we immediately obtain $S^- = T^-$. Now let $pa \in \text{Ran}(PA_V^{\mathcal{F}} \cup PA_V^{\mathcal{G}})$. Suppose $y \in \mathcal{F}_Y(pa_{\upharpoonright PA_V^{\mathcal{F}}})$. By definition of

intervention, for any $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$ with $\mathbf{w}_{\upharpoonright PA_V^{\mathcal{F}} \cup PA_V^{\mathcal{G}}} = pa$ we have $S \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$. But then, by the assumption, $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$. Thus, by the same reasoning, $y \in \mathcal{G}_Y(pa_{\upharpoonright PA_V^{\mathcal{G}}})$. The converse is analogous. Thus, we conclude that $\mathcal{F} \equiv \mathcal{G}$. $\square$

We can improve on this result by showing that each equivalence class of law components can be characterized by a single $\mathcal{H}$ formula. For each signature $\sigma = (\mathcal{E}, \mathcal{I}, \text{Ran})$ and each $\mathcal{F} \in \mathbb{R}_{\sigma}$, we define a formula

$$\Phi^{\mathcal{F}} := \bigwedge_{\substack{Y \in \mathcal{I} \\ \mathbf{w} \in \text{Ran}(\mathbf{W}_Y) \\ y \in \mathcal{F}_Y(\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{F}}})}} \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \ \& \ \bigwedge_{\substack{Y \in \mathcal{I} \\ \mathbf{w} \in \text{Ran}(\mathbf{W}_Y) \\ y \notin \mathcal{F}_Y(\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{F}}})}} \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y.$$

The $\Phi^{\mathcal{F}}$ formulas will be crucial in the formulation of two axioms; the next result shows that they indeed characterize law components up to equivalence.

Theorem 5.2. Let $T = (T^-, \mathcal{G})$ be a model. Then, $T \models \Phi^{\mathcal{F}} \iff \mathcal{G} \equiv \mathcal{F}$.

Proof. $\Rightarrow$) Suppose $T \models \Phi^{\mathcal{F}}$. Let $Y \in \mathcal{I}$, $pa \in \text{Ran}(PA_Y^{\mathcal{F}} \cup PA_Y^{\mathcal{G}})$. Pick a $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$ such that $\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{F}} \cup PA_Y^{\mathcal{G}}} = pa$.

Suppose first that $y \in \mathcal{F}_Y(pa_{\upharpoonright PA_Y^{\mathcal{F}}})$. Thus, we may also write $y \in \mathcal{F}_Y(\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{F}}})$. Since $T \models \Phi^{\mathcal{F}}$, then, in particular $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$. By definition of intervention, then, $y \in \mathcal{G}(\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{G}}}) = \mathcal{G}(pa_{\upharpoonright PA_Y^{\mathcal{G}}})$, as needed.

Vice versa, suppose $y \in \mathcal{G}(pa_{\upharpoonright PA_Y^{\mathcal{G}}})$. Then, by definition of intervention, $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$. Suppose for the sake of contradiction that $y \notin \mathcal{F}_Y(pa_{\upharpoonright PA_Y^{\mathcal{F}}})$. Then, $y \notin \mathcal{F}_Y(\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{F}}})$. So, since $T \models \Phi^{\mathcal{F}}$, we obtain that $T \models \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$, contradicting our earlier statement.

Thus, $\mathcal{G} \equiv \mathcal{F}$.

$\Leftarrow$) Suppose $\mathcal{G} \equiv \mathcal{F}$. Let $Y \in \mathcal{I}$, $pa \in \text{Ran}(PA_Y^{\mathcal{F}} \cup PA_Y^{\mathcal{G}})$. Pick a $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$ such that $\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{F}} \cup PA_Y^{\mathcal{G}}} = pa$.

Assume first that $y \in \mathcal{F}_Y(\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{F}}})$. Since $\mathcal{G} \equiv \mathcal{F}$, we have $y \in \mathcal{G}_Y(\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{G}}})$. Thus, by definition of intervention, $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$, as needed.

If we assume instead that $y \notin \mathcal{F}_Y(\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{F}}})$, since $\mathcal{G} \equiv \mathcal{F}$ we obtain $y \notin \mathcal{G}_Y(\mathbf{w}_{\upharpoonright PA_Y^{\mathcal{G}}})$. By definition of intervention, then, $T \not\models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$, and thus $T \models \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$. $\square$

6 A general axiomatization

In the following, we prove the soundness and completeness of the following system B_{σ} for language $\mathcal{H}^+$ over the class of all models (again, B_{σ} will often just denote the set of axioms of this system).

$$\text{Rule MP. } \frac{\psi \quad \psi \rightarrow \chi}{\chi} \quad \text{Rule NEC. } \frac{\vdash \psi}{\vdash [\mathbf{X} = \mathbf{x}]\psi}$$

J0. $\mathcal{H}_{\sigma}^+$ instances of a classical tautologies in $\sim, \&$.

- J1. $[\mathbf{X} = \mathbf{x}]Y = y \rightarrow [\mathbf{X} = \mathbf{x}]Y \neq y'$ (when $y \neq y'$) [Uniqueness]
- J2. $[\mathbf{X} = \mathbf{x}] \sqcup_{y \in \text{Ran}(Y)} Y = y$ [Definiteness]
- J3. $\langle \mathbf{X} = \mathbf{x} \rangle (Z = z \ \& \ Y = y) \rightarrow \langle \mathbf{X} = \mathbf{x}, Z = z \rangle Y = y$ [Weak composition]
- J4. $[\mathbf{X} = \mathbf{x}, Y = y]Y = y$ [Effectiveness]
- J5. $([\mathbf{X} = \mathbf{x}]\psi \ \& \ [\mathbf{X} = \mathbf{x}](\psi \rightarrow \chi)) \rightarrow [\mathbf{X} = \mathbf{x}]\chi$ [K-axiom]
- J6. $\mathbf{W} = \mathbf{w} \rightarrow \text{I6}$. [Reversibility for small teams]
- J7. $\mathbf{Y} = y \leftrightarrow \square \mathbf{Y} = y$. [Flatness]
- J8. $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \leftrightarrow \diamond Y = y$ (if $Y \in \mathcal{E}$) [External variables]
- J9. $\langle \mathbf{W} = \mathbf{w} \rangle \top$ [Full intervention]
- J10. $\langle \mathbf{X} = \mathbf{x} \rangle \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \leftrightarrow \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$ (if $Y \in \mathcal{I}$ and $Y \notin \mathbf{X}$) [Diamond screen-off]

- J11. $[X = x] \langle W_Y = w \rangle Y = y \leftrightarrow \langle W_Y = w \rangle Y = y$ (if $Y \in \mathcal{I}$ and $Y \notin X$)
[Box screen-off]
J12. $[X = x] \diamond \psi \leftrightarrow [X = x] \psi$ [Conditional diamond]
J13. $\langle X = x \rangle \diamond \psi \leftrightarrow \langle X = x \rangle \psi$ [Mighty diamond]
J14. $\diamond \langle X = x \rangle \psi \leftrightarrow \langle X = x \rangle \psi$ [Double diamond]
J15. $\langle X = x \rangle \psi \rightarrow \langle X = x \rangle (\psi \& \diamond \top)$ [Persistency of diamond]
J16. $\bigsqcup_{\mathcal{F} \in \mathcal{F}_\sigma} \Phi^\mathcal{F}$ [Certainty of laws]
J17. $\Phi^\mathcal{F} \rightarrow ([X = x] \psi \leftrightarrow [X = x] \bigsqcup_{\substack{t \text{ such that} \\ (\{t\}, \mathcal{F}_{X=x}) \models \psi}} \mathbf{W} = t(\mathbf{W}))$ [Modal collapse]

As in system A, we allow $\mathbf{X}$ and $\mathbf{Y}$ to be empty tuples, as a special case.

Eight of the axioms (those marked with a letter I) are shared with system A_σ . Note the absence of I10 and I1b (which only hold in nonempty models) and I2b (which holds iff there is at most one assignment). Axiom J6 says that *Weak reversibility* holds for singleton and empty models (under the same restrictions as for axiom I6, which have not been repeated here). The remaining axioms deal with the nesting of modal operators. Axioms J10–J11 say that an intervention on all variables except an internal Y trumps any previous intervention on variables different from Y . Axioms J12–J13–J14–J15 are various consequences of the fact that the semantics of the modal operators only involves looking at certain singleton submodels. We note that the "empty intervention" versions of these axioms imply that any sequence of $\diamond$ and $\Box$ modalities can always be reduced to the *leftmost* modality occurring in it (contrarily, e.g., to what happens in modal logic S5). J16 says that any model has a fixed set of causal laws. Together with J17, it will be used to show that, when reasoning with maximal consistent sets of formulas, the $\mathcal{H}^+$ formulas can always be "reduced" to $\mathcal{H}$ formulas (note that, differently from $[X = x] \psi$, the rightmost member of J17 is always an unnested formula). We have to admit that axiom scheme J17 is rather unsatisfying due to its complexity and semantic flavour; however, note that membership in this scheme is decidable due to the finiteness of the models. Furthermore, the problem of deciding which formulas are instances of J17 reduces to the proof-theoretic problem of finding proofs of consequences of the forms $\mathbf{W} = s(\mathbf{W}), \Phi^{\mathcal{F}_{X=x}} \vdash \psi$ and $\mathbf{W} = s(\mathbf{W}), \Phi^{\mathcal{F}_{X=x}} \vdash \sim \psi$ in system A, which we already know to be complete.

Theorem 6.1. *System B_σ is sound over the class of all models of signature σ .*

Proof. See the proof of Theorem 4.2 for the soundness of axioms J0, I1–I5, and I7–I9.

Concerning axiom J6, suppose $T \models \mathbf{W} = \mathbf{w}$. Then, T is either a singleton model or empty. If it is a singleton model, then it satisfies I6 (see Theorem 4.2 for a proof). Suppose now that T is empty, and that it satisfies $\psi : \langle X = x, V = v \rangle (Y = y \& Z = z)$ and $\chi : \langle X = x, Y = y \rangle (V = v \& Z = z)$. By ψ , there is an $s \in T_{X=x, V=v}^-$ such that $s(Y) = y$ and $s(Z) = z$. But (since $X \cup Z \cup \{Y\} \cup \{V\} = \text{Dom}$) χ tells us that s is also in $T_{X=x, Y=y}^-$. In particular, if V is internal, s is compatible also with the law $\mathcal{F}_V$; if we prove that (*): $s(\mathbf{N}_X) \in T^-(\mathbf{N}_X)$, then we can conclude that $s \in T_{X=x}^-$. But, since $s \in T_{X=x, V=v}^-$ and T^- is empty, by definition of intervention, we must have $\mathbf{N}_{XV} = \emptyset$; and similarly, we obtain $\mathbf{N}_{XY} = \emptyset$. But then $\mathbf{N}_X = \mathbf{N}_{XY} \cup \mathbf{N}_{XV} = \emptyset$, so that condition (*) is trivially satisfied.

For axiom J10, note that $T \models \langle X = x \rangle \langle W_Y = w \rangle Y = y$ iff, for some $s \in T_{X=x}^-$, $(\{s\}, \mathcal{F}_{X=x}) \models \langle W_Y = w \rangle Y = y$. The latter holds iff there is a $t \in (\{s\}, \mathcal{F}_{X=x})_{W_Y=w}^-$ such that $(\{t\}, (\mathcal{F}_{X=x})_{W_Y=w}) \models Y = y$. But $(\mathcal{F}_{X=x})_{W_Y=w} = \mathcal{F}_{W_Y=w}$ (since $X \subseteq W_Y$). Let us show that $t \in T_{W_Y=w}^-$. Since $t \in (\{s\}, \mathcal{F}_{X=x})_{W_Y=w}^-$, by definition of intervention, $t(\mathbf{W}_Y) = \mathbf{w}$, and, since $Y \in \mathcal{I}$, $t(Y) \in \mathcal{F}_Y(\mathbf{w}_{\uparrow PA_Y})$, as needed. Thus, $t \in T_{W_Y=w}^-$, and therefore $T \models \langle W_Y = w \rangle Y = y$.

Vice versa, if $T \models \langle W_Y = w \rangle Y = y$, there is a $t \in T_{W_Y=w}^-$ such that $(\{t\}, \mathcal{F}_{W_Y=w}) \models Y = y$. As before $(\mathcal{F}_{X=x})_{W_Y=w} = \mathcal{F}_{W_Y=w}$. It suffices then to show that $t \in (T_{X=x})_{W_Y=w}^-$. Since $t \in T_{W_Y=w}^-$, we have $t(\mathbf{W}_Y) = \mathbf{w}$, as needed. Since Y is internal (in all the involved causal teams, since $Y \notin X$), then $t(Y) = y \in (\mathcal{F}_{W_Y=w})_Y(\mathbf{w}_{\uparrow PA_Y}) = ((\mathcal{F}_{X=x})_{W_Y=w})_Y(\mathbf{w}_{\uparrow PA_Y})$; thus, $t \in (T_{X=x})_{W_Y=w}^-$, as needed. The soundness of axiom J11 is proved analogously.

The soundness of axioms J12, J13, and J14 is a straightforward consequence of the semantic clause for the modal operators, which just requires checking whether the consequent is satisfied in singleton models. For example, for axiom J12, note that $T \models [X = x] \diamond \psi$ iff all singleton submodels of $T_{X=x}$ satisfy $\diamond \psi$ iff (being singleton) they satisfy ψ , iff $T \models [X = x] \psi$.

For axiom J15, suppose $T \models \langle X = x \rangle \psi$. Then, there is an $s \in T_{X=x}^-$ such that $(\{s\}, \mathcal{F}_{X=x}) \models \psi$. Obviously $(\{s\}, \mathcal{F}_{X=x}) \models \diamond \top$. Thus, $(\{s\}, \mathcal{F}_{X=x}) \models \psi \& \diamond \top$. But $s \in T_{X=x}^-$; thus, $T \models \langle X = x \rangle (\psi \& \diamond \top)$.

For axiom J16, note that $T = (T^-, \mathcal{F}) \models \Phi^\mathcal{F}$ by Theorem 5.2. Thus, $T \models \bigsqcup_{\mathcal{F} \in \mathcal{F}_\sigma} \Phi^\mathcal{F}$.

For axiom J17, suppose $T \models \Phi^\mathcal{F}$.

$\Rightarrow$) Assume $T \models [X = x] \psi$. Then, for all $s \in (T_{X=x})^-$, (*): $(\{s\}, \mathcal{F}_{X=x}) \models \psi$. But we also have that $(\{s\}, \mathcal{F}_{X=x}) \models \mathbf{W} = s(\mathbf{W})$; thus, by (*), $(\{s\}, \mathcal{F}_{X=x}) \models \bigsqcup_{\substack{t \text{ such that} \\ (\{t\}, \mathcal{F}_{X=x}) \models \psi}} \mathbf{W} = t(\mathbf{W})$. Since this holds for all $s \in T_{X=x}^-$, $T \models [X = x] \bigsqcup_{\substack{t \text{ such that} \\ (\{t\}, \mathcal{F}_{X=x}) \models \psi}} \mathbf{W} = t(\mathbf{W})$.
 $\Leftarrow$) Assume $T \models [X = x] \bigsqcup_{\substack{t \text{ such that} \\ (\{t\}, \mathcal{F}_{X=x}) \models \psi}} \mathbf{W} = t(\mathbf{W})$. Then, for all $s \in (T_{X=x})^-$, $(\{s\}, \mathcal{F}_{X=x}) \models \bigsqcup_{\substack{t \text{ such that} \\ (\{t\}, \mathcal{F}_{X=x}) \models \psi}} \mathbf{W} = t(\mathbf{W})$. Thus, by the semantic clauses, $s = t$ for some t such that $(\{t\}, \mathcal{F}) \models \psi$. In other words, $(\{s\}, \mathcal{F}) \models \psi$. Thus, $T \models [X = x] \psi$. $\square$

Since this system includes modal logic K (MP, NEC, axioms J0 and I5), we automatically obtain the same theorems as in Lemma 4.4. But note that points 2., 3. and 5. have more relaxed statements.

Lemma 6.2. *The following hold in B.*

1. (Deduction theorem) If $\Gamma, \psi \vdash \chi$, then $\Gamma \vdash \psi \rightarrow \chi$.
2. ($\Box$ -Monotonicity) If $\Gamma \vdash [X = x] \psi$ and, $\vdash \psi \rightarrow \psi'$, then $\Gamma \vdash [X = x] \psi'$.
3. ($\Diamond$ -Monotonicity) If $\Gamma \vdash \langle X = x \rangle \psi$ and $\vdash \psi \rightarrow \psi'$, then $\Gamma \vdash \langle X = x \rangle \psi'$.
4. $\vdash ([X = x] \psi \& [X = x] \chi) \leftrightarrow [X = x] (\psi \& \chi)$.
5. (Replacement) Suppose $\Gamma \vdash \theta \leftrightarrow \theta'$. Then, $\Gamma \vdash \varphi \leftrightarrow \varphi[\theta'/\theta]$.
6. $\vdash \sim [X = x] \psi \leftrightarrow \langle X = x \rangle \sim \psi$
7. $\vdash \sim \langle X = x \rangle \psi \leftrightarrow [X = x] \sim \psi$
8. $\vdash (\langle X = x \rangle \psi \sqcup \langle X = x \rangle \chi) \leftrightarrow \langle X = x \rangle (\psi \sqcup \chi)$.
9. $\vdash ([X = x] \psi \& \langle X = x \rangle \top) \rightarrow \langle X = x \rangle \psi$.
10. $\vdash ([X = x] \psi \& \langle X = x \rangle \chi) \rightarrow \langle X = x \rangle (\psi \& \chi)$.
11. $\vdash \langle X = x \rangle (\psi \& \chi) \rightarrow (\langle X = x \rangle \psi \& \langle X = x \rangle \chi)$.

Let us clarify the relationship between systems A_σ and B_σ . The intuitive picture is that system B_σ is conservative over A_σ when restricting our attention to singleton models. Formally, this means:

Theorem 6.3. *Let $\Gamma \cup \{\varphi\} \subseteq \mathcal{H}_\sigma$. Then, $\Gamma \vdash_{A_\sigma} \varphi$ iff $\diamond \top, \bigsqcup_{\mathbf{w} \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w}, \Gamma \vdash_{B_\sigma} \varphi$.*

Proof. For brevity, let us write Δ for the set of assumptions $\{\diamond \top\} \cup \{\bigsqcup_{\mathbf{w} \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w}\} \cup \Gamma$. Let us first observe that, if θ is an axiom of A_σ , then $\Delta \vdash_{B_\sigma} \theta$. Indeed, most of the axioms of A_σ are included in B_σ . For axiom I10, obviously $\Delta \vdash_{B_\sigma} \diamond \top$. For I2b, note first that $\Delta \vdash_{B_\sigma} \bigsqcup_{\mathbf{w} \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w}$ trivially. But then $\Delta \vdash_{B_\sigma} \bigsqcup_{y \in \text{Ran}(Y)} Y = y$ follows by J0 and MP.

For axiom I1b, write Ξ for $\Delta \cup \{Y = y\} \cup \{Y = y'\}$. By axiom I7, $\Xi \vdash_{B_\sigma} \Box Y = y$ and $\Xi \vdash_{B_\sigma} \Box Y = y'$. On the other hand, by I1, $\Box Y = y \rightarrow \Box Y \neq y'$; thus, by MP, $\Xi \vdash_{B_\sigma} \Box Y \neq y'$. By J0 and Lemma 6.2, 4., $\Xi \vdash_{B_\sigma} \Box(Y = y' \& Y \neq y')$. Thus, by J0 and replacement, $\Xi \vdash_{B_\sigma} \Box \perp$. Since $\Diamond \top \in \Delta$, by Lemma 6.2, 9., $\Xi \vdash_{B_\sigma} \Diamond \perp$, i.e., $\Xi \vdash_{B_\sigma} \sim \Box \top$. On the other hand, $\Xi \vdash_{B_\sigma} \Box \top$ by J0 and NEC. So, by J0, $\Xi \vdash_{B_\sigma} \perp$. By the deduction theorem, then, $\Delta \vdash_{B_\sigma} (Y = y \& Y = y') \rightarrow \perp$. By J0, then, $\Delta \vdash_{B_\sigma} Y = y \rightarrow Y \neq y'$.

Finally, for axiom I6, just note that, by J6, J0 and MP, $\Delta \vdash_{B_\sigma} \bigsqcup_{w \in \text{Ran}(\mathbf{W})} \text{I6}$. Thus, by J0 again, $\Delta \vdash_{B_\sigma} \text{I6}$.

Since all axioms of A_σ are $\vdash_{B_\sigma}$ -derivable from Δ , and B includes all the rule instances of A, it is straightforward to transform any proof of the form $\Gamma \vdash_{A_\sigma} \varphi$ into a proof of $\Delta \vdash_{B_\sigma} \varphi$.

Vice versa, if $\Delta \vdash_{B_\sigma} \varphi$, by the soundness of B (Theorem 6.1), $\Delta \models \varphi$. But this (by Lemma 4.1) means that every singleton model that satisfies Γ also satisfies φ ; i.e., $\Gamma \models \varphi$. But then $\Gamma \vdash_{A_\sigma} \varphi$ by the completeness theorem for A_σ (Theorem 4.13). $\square$

We will now consider maximally B_σ -consistent sets and their relationships to canonical teams. Note that, for sets Γ of $\mathcal{H}_\sigma^+$ formulas that are maximally consistent according to B_σ , we can define their associate canonical causal team $\mathbb{T}^\Gamma$ exactly as in section 4; we may write $\mathbb{T}_B^\Gamma$ where there is the risk of confusion. As before, we can prove:

Lemma 6.4. *Let $\Gamma \supseteq B_\sigma$ be a maximal consistent set of $\mathcal{H}_\sigma^+$ formulas. Then $\mathbb{T}^\Gamma$ is a model, i.e.,*

1. *For all $Y \in \text{Int}(\mathbb{T}^\Gamma)$, $\mathcal{F}_Y^\Gamma$ is well-defined.*
2. *For all $Y \in \text{Int}(\mathbb{T}^\Gamma)$ and $s \in (\mathbb{T}^\Gamma)^-$, $(s(PA_Y), s(Y)) \in \mathcal{F}_Y^\Gamma$.*

The proof is as for A, since it only uses axiom I3 and classical logic.

The following corollary of Theorem 6.3 shows that the notational abuse of writing $\mathbb{T}^\Gamma$ for $\mathbb{T}_B^\Gamma$ is not too dangerous, at least when $\mathbb{T}_B^\Gamma$ is a singleton model.

Corollary 6.5. *Let $\Gamma \supseteq B_\sigma$ be a maximal B_σ -consistent set of $\mathcal{H}_\sigma^+$ formulas such that $\mathbb{T}_B^\Gamma$ is a singleton model, and let Γ_0 be the set of $\mathcal{H}_\sigma$ formulas in Γ . Then, Γ_0 is maximally A_σ -consistent, and $\mathbb{T}^{\Gamma_0} = \mathbb{T}_B^\Gamma$.*

Proof. First, we show that $\Diamond \top, \bigsqcup_{w \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w} \in \Gamma$.

For the former, note that since $\mathbb{T}_B^\Gamma$ is a singleton model, by its definition there is a $\mathbf{w} \in \text{Ran}(\mathbf{W})$ such that $\Diamond \mathbf{W} = \mathbf{w} \in \Gamma$. But $\vdash_{B_\sigma} \mathbf{W} = \mathbf{w} \rightarrow \top$ by J0, so $\Diamond \top \in \Gamma$ by monotonicity.

For the latter, note that, by the definition of $\mathbb{T}_B^\Gamma$ and the fact that it is a singleton model, we also have that, for all $\mathbf{w}' \neq \mathbf{w}$, $\Diamond \mathbf{W} = \mathbf{w}' \notin \Gamma$. Since Γ is maximal, $\sim \Diamond \mathbf{W} = \mathbf{w}' \in \Gamma$. By Lemmas 6.2, 7. and Lemma 4.5, 3., we obtain $\bigwedge_{\mathbf{w}' \neq \mathbf{w}} \Box \sim \mathbf{W} = \mathbf{w}' \in \Gamma$. By Lemma 6.2, 4., $\Box \bigwedge_{\mathbf{w}' \neq \mathbf{w}} \sim \mathbf{W} = \mathbf{w}' \in \Gamma$. Thus, since we have $\Box \bigsqcup_{w \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w} \in \Gamma$ based on I2, from Lemma 4.5, 3. and Lemma 6.2, 4. we obtain $\Box (\bigwedge_{\mathbf{w}' \neq \mathbf{w}} \sim \mathbf{W} = \mathbf{w}' \& \bigsqcup_{w \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w}) \in \Gamma$. Thus, based on J0 and replacement, $\Box \mathbf{W} = \mathbf{w} \in \Gamma$. By I7, $\mathbf{W} = \mathbf{w} \in \Gamma$. Thus, by I0, $\bigsqcup_{w \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w} \in \Gamma$.

We use these facts to prove that Γ_0 is closed under A-derivations. Suppose $\Gamma_0 \vdash_A \varphi$. Then, from Theorem 6.3, $\Diamond \top, \bigsqcup_{w \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w}, \Gamma_0 \vdash_B \varphi$; but then, based on what we proved above, $\Gamma \vdash_B \varphi$. Since Γ is maximal, then, $\varphi \in \Gamma$; since $\varphi \in \mathcal{H}$, we conclude $\varphi \in \Gamma_0$.

Now suppose for the sake of contradiction that Γ_0 is A-inconsistent; then $\Gamma_0 \vdash_{A_\sigma} \perp$, but since Γ_0 is closed under A-derivations, $\perp \in \Gamma_0 \subseteq \Gamma$, contradicting the B-consistency of Γ .

Finally, suppose for the sake of contradiction that $\Gamma_1 \supset \Gamma_0$ is an A_σ -consistent set of $\mathcal{H}_\sigma$ formulas; say, $\varphi \in \Gamma_1 \setminus \Gamma_0$. Then, $\varphi \notin \Gamma$, so, by the maximality of Γ , $\sim \varphi \in \Gamma$. But $\sim \varphi$ is an $\mathcal{H}$ formula, so $\sim \varphi \in \Gamma_0 \subseteq \Gamma_1$. Thus, after all, Γ_1 is inconsistent.

Now, since Γ_0 is maximally A_σ -consistent, it defines the canonical team $\mathbb{T}^{\Gamma_0}$, which by definition is the same as $\mathbb{T}_B^\Gamma$. $\square$

Theorem 6.3 also ensures that the truth lemma will work correctly at least for the $\mathcal{H}$ formulas in Γ , provided $\mathbb{T}^\Gamma$ is a singleton model.

Corollary 6.6. *Let $\Gamma \supseteq B_\sigma$ be a set of $\mathcal{H}_\sigma^+$ formulas that is maximally consistent, and φ be an $\mathcal{H}_\sigma$ formula.*

1. *If φ is an atomic formula, then $\varphi \in \Gamma \iff \mathbb{T}^\Gamma \models \varphi$.*
2. *If $\mathbb{T}^\Gamma$ is a singleton model, then $\varphi \in \Gamma \iff \mathbb{T}^\Gamma \models \varphi$.*

Proof. 1) Suppose $Y = y \in \Gamma$. Then, $\Box Y = y \in \Gamma$ by I7. Suppose for the sake of contradiction that there is a $t \in (\mathbb{T}^\Gamma)^-$ such that $t(Y) = y' \neq y$. Then, by definition of $\mathbb{T}^\Gamma$, $\Diamond \mathbf{W} = t(\mathbf{W}) \in \Gamma$. Thus, $\Diamond Y = y' \in \Gamma$ by Lemma 6.2, 11. Thus, by Lemma 6.2, 10., $\Diamond(Y = y \& Y = y') \in \Gamma$. On the other hand, by I1 and MP we have $\Box Y \neq y' \in \Gamma$. Again, by Lemma 6.2, 10., we obtain $\Diamond(Y = y \& Y = y' \& Y \neq y') \in \Gamma$, so, by I0 and monotonicity, $\Diamond \perp \in \Gamma$. This contradicts Lemma 4.5, 5. Thus, all assignments in $(\mathbb{T}^\Gamma)^-$ satisfy $Y = y$, and so $\mathbb{T}^\Gamma \models Y = y$.

Vice versa, suppose $\mathbb{T}^\Gamma \models Y = y$. Suppose for the sake of contradiction that $Y = y \notin \Gamma$. By axiom I7, $\Box Y = y \notin \Gamma$. By maximality, $\sim \Box Y = y \in \Gamma$, and by Lemma 6.2, 6., $\Diamond \sim Y = y \in \Gamma$. On the other hand, by axioms I2, $\Box \bigsqcup_{w \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w} \in \Gamma$. Thus, by Lemma 6.2, 10., $\Diamond(\sim Y = y \& \bigsqcup_{w \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w}) \in \Gamma$. By I0 and monotonicity, $\Diamond \bigsqcup_{w \in \text{Ran}(\mathbf{W}), w \models Y \neq y} \mathbf{W} = \mathbf{w} \in \Gamma$. By Lemma 6.2, 8., $\bigsqcup_{w \in \text{Ran}(\mathbf{W}), w \models Y \neq y} \Diamond \mathbf{W} = \mathbf{w} \in \Gamma$. Thus, by primality of Γ , there is a $\mathbf{w} \in \text{Ran}(\mathbf{W})$ such that $w \models Y \neq y$ and $\Diamond \mathbf{W} = \mathbf{w} \in \Gamma$. Thus, the assignment $s(\mathbf{W}) = \mathbf{w}$ is in $(\mathbb{T}^\Gamma)^-$. But $s(Y) \neq y$; thus, $\mathbb{T}^\Gamma \not\models Y = y$, contradicting our initial assumption.

2) Let Γ_0 be the set of all $\mathcal{H}$ formulas in Γ . From right to left, suppose $\mathbb{T}^\Gamma \models \varphi$; by corollary 6.5, this gives $\mathbb{T}^{\Gamma_0} \models \varphi$. But then, since $\varphi \in \mathcal{H}$, by the truth lemma 4.12 for Γ_0 , $\varphi \in \Gamma_0 \subseteq \Gamma$.

From left to right, note that, since $\mathbb{T}^{\Gamma_0} = \mathbb{T}^\Gamma$ is a singleton model, $\mathbb{T}^{\Gamma_0} \models \Diamond \top, \bigsqcup_{w \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w}$. Now, if $\varphi \in \Gamma$, then $\varphi \in \Gamma_0$ (since it is an $\mathcal{H}$ formula). Then, by the maximality of Γ_0 , $\Gamma_0 \vdash_{A_\sigma} \varphi$. Thus, by Theorem 6.3, we have $\Diamond \top, \bigsqcup_{w \in \text{Ran}(\mathbf{W})} \mathbf{W} = \mathbf{w}, \Gamma_0 \vdash_{B_\sigma} \varphi$. Furthermore, $\mathbb{T}^{\Gamma_0} \models \Gamma_0$ by the truth lemma for $\mathbb{T}^{\Gamma_0}$. Thus, by the soundness of B_σ , $\mathbb{T}^\Gamma = \mathbb{T}^{\Gamma_0} \models \varphi$. $\square$

This corollary gives us some control on the $\mathcal{H}_\sigma$ -formulas that are B-provable for singleton models. To extend our grasp to $\mathcal{H}^+$ formulas, we aim to show that the $\mathcal{H}_\sigma$ formulas that hold in a singleton model completely determine the truth or falsity of $\mathcal{H}_\sigma^+$ formulas. We already know that this holds semantically (even for non-singleton models) by Theorem 5.1; we seek here a proof-theoretic counterpart of this result. We need a couple of intermediate lemmas; the first one shows that, if $\Gamma \supseteq B_\sigma$ is a maximal consistent set of formulas, then there is exactly one $\mathcal{F}$ (up to equivalence) such that $\Phi^\mathcal{F} \in \Gamma$.

Lemma 6.7. *If $\Gamma \supseteq B_\sigma$ is a maximal consistent set of $\mathcal{H}^+$ formulas, then there is an $\mathcal{F} \in \mathbb{F}_\sigma$ such that $\Phi^\mathcal{F} \in \Gamma$; if $\Phi^\mathcal{F}, \Phi^\mathcal{G} \in \Gamma$, then $\mathcal{F} \equiv \mathcal{G}$.*

Proof. Existence: Since $\Gamma \supseteq B_\sigma$, by axiom J16 we have $\bigsqcup_{\mathcal{F} \in \mathbb{F}_\sigma} \Phi^\mathcal{F} \in \Gamma$. Thus, by the primality of Γ , there is an $\mathcal{F}$ such that $\Phi^\mathcal{F} \in \Gamma$.

Uniqueness: Suppose for the sake of contradiction that $\Phi^\mathcal{F}, \Phi^\mathcal{G} \in \Gamma$, where $\mathcal{F} \neq \mathcal{G}$. The statement $\mathcal{F} \neq \mathcal{G}$ implies that there is a $Y \in \mathcal{I}$ and a tuple $pa \in \text{Ran}(PA_Y^\mathcal{F} \cup PA_Y^\mathcal{G})$ such that $\mathcal{F}_Y(pa_{\upharpoonright PA_Y^\mathcal{F}}) \neq \mathcal{G}_Y(pa_{\upharpoonright PA_Y^\mathcal{G}})$. Suppose wlog that there is a $y \in \mathcal{F}_Y(pa_{\upharpoonright PA_Y^\mathcal{F}}) \setminus \mathcal{G}_Y(pa_{\upharpoonright PA_Y^\mathcal{G}})$. Then, by J0, $\Phi^\mathcal{F} \in \Gamma$ implies $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$ and $\Phi^\mathcal{G} \in \Gamma$ implies $\sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$ for some $\mathbf{w}$ such that $\mathbf{w}_{\upharpoonright PA_Y^\mathcal{F} \cup PA_Y^\mathcal{G}} = pa$, contradicting the consistency of Γ . $\square$

Lemma 6.8. *Let $\Gamma \supseteq B_\sigma$ be a maximal consistent set of $\mathcal{H}_\sigma^+$ formulas. Then, for each $\varphi \in \mathcal{H}_\sigma^+$ there is a $\varphi_\Gamma^* \in \mathcal{H}_\sigma$ such that $\Gamma \vdash_{B_\sigma} \varphi \leftrightarrow \varphi_\Gamma^*$.*

Furthermore, φ_Γ^ is determined just by the causal laws encoded in Γ , i.e., if $\Phi^\mathcal{F} \in \Gamma$ and $\Phi^\mathcal{G} \in \Gamma'$, then $\varphi_\Gamma^* = \varphi_{\Gamma'}^*$.*

Proof. We define φ_Γ^* and prove $\Gamma \vdash \varphi \leftrightarrow \varphi_\Gamma^*$ by induction on φ .

If φ is atomic, we let $\varphi_\Gamma^* := \varphi$; then, $\Gamma \vdash \varphi \leftrightarrow \varphi_\Gamma^*$ follows from J0 (since $B_\sigma \subseteq \Gamma$).

If φ is $\psi \& \chi$, then we let $\varphi_\Gamma^* := \psi_\Gamma^* \& \chi_\Gamma^*$. By the i.h., $\Gamma \vdash \psi \leftrightarrow \psi_\Gamma^*$ and $\Gamma \vdash \chi \leftrightarrow \chi_\Gamma^*$; thus, $\Gamma \vdash \varphi \leftrightarrow \varphi_\Gamma^*$ follows by I0 and MP. The case for φ of the form $\sim \psi$ is similar (letting $\varphi_\Gamma^* := \sim \psi_\Gamma^*$).

If φ is $[\mathbf{X} = \mathbf{x}]\psi$, then observe that by Lemma 6.7, there is one function component $\mathcal{F}$ (up to equivalence) such that $\Phi^\mathcal{F} \in \Gamma$. We let $\varphi_\Gamma^* := [\mathbf{X} = \mathbf{x}] \bigsqcup_{\substack{t \text{ such that} \\ (t, \mathcal{F}_{\mathbf{X}=\mathbf{x}}) \models \psi}} \mathbf{W} = t(\mathbf{W})$. We note that, by Theorem 5.1, if $\mathcal{F} \equiv \mathcal{G}$ then this formula is the same as $[\mathbf{X} = \mathbf{x}] \bigsqcup_{\substack{t \text{ such that} \\ (t, \mathcal{G}_{\mathbf{X}=\mathbf{x}}) \models \psi}} \mathbf{W} = t(\mathbf{W})$ (as obviously $\mathcal{F} \equiv \mathcal{G}$ implies $\mathcal{F}_{\mathbf{X}=\mathbf{x}} \equiv \mathcal{G}_{\mathbf{X}=\mathbf{x}}$); thus, the definition of φ_Γ^* is determined by the equivalence class of $\mathcal{F}$. Now, since $\Phi^\mathcal{F} \in \Gamma$, by axiom J17 $\Gamma \vdash \varphi \leftrightarrow \varphi_\Gamma^*$. $\square$

Upon closer inspection, this lemma also proves that, if an $\mathcal{H}^+$ formula completely determines the causal laws (i.e., it entails $\Phi^\mathcal{F}$ for some $\mathcal{F}$) then it is equivalent to an $\mathcal{H}$ formula. We will not need this fact in the following.

Theorem 6.9. *For every $\Gamma_0 \supseteq A_\sigma$ maximal A_σ -consistent set of $\mathcal{H}_\sigma$ formulas there is a unique maximal B_σ -consistent set $\Gamma \supseteq \Gamma_0 \cup B_\sigma$ of $\mathcal{H}^+$ formulas.*

Proof. Existence. Since Γ_0 includes A_σ and is maximally A_σ -consistent, by Lemma 4.12 Γ_0 has a (singleton) model T . Let $\Gamma := \{\varphi \in \mathcal{H}_\sigma^+ \mid T \models \varphi\}$. Since $T \models \Gamma$, by the soundness of B_σ (Lemma 6.1), we have $B_\sigma \subseteq \Gamma$. And since it is the set of $\mathcal{H}_\sigma^+$ formulas satisfied by a model, Γ is maximally consistent.

Uniqueness. Suppose for the sake of contradiction that there are two distinct Γ, Γ' satisfying the statement. Say, wlog, that there is a $\varphi \in \Gamma \setminus \Gamma'$. Thus, by maximality of Γ' , $\sim \varphi \in \Gamma'$. By Lemma 6.8, there is a $\varphi_\Gamma^* \in \mathcal{H}_\sigma$ such that $\Gamma \vdash \varphi \leftrightarrow \varphi_\Gamma^*$; thus, by MP, (*): $\varphi_\Gamma^* \in \Gamma_0 \subseteq \Gamma$. By the same lemma, we also have $\Gamma' \vdash \sim \varphi \leftrightarrow (\sim \varphi)_\Gamma^*$. But the proof of the lemma shows that $(\sim \varphi)_\Gamma^* = \sim \varphi_\Gamma^*$. Furthermore, note that, since $\Phi^\mathcal{F} \in \mathcal{H}$, we have $\Phi^\mathcal{F} \in \Gamma \iff \Phi^\mathcal{F} \in \Gamma_0 \iff \Phi^\mathcal{F} \in \Gamma'$. Since the φ_Γ^* are determined by the $\mathcal{F}$ such that $\Phi^\mathcal{F} \in \Gamma$, we conclude that $\varphi_\Gamma^* = \varphi_{\Gamma'}^*$. Therefore, $\Gamma' \vdash \sim \varphi \leftrightarrow \sim \varphi_\Gamma^*$. So, $\sim \varphi_\Gamma^* \in \Gamma_0 \subseteq \Gamma'$. This, together with (*), contradicts the consistency of Γ_0 . $\square$

We now move toward a completeness proof. The failure of *Weak reversibility* over general models (and the apparent lack of an adequate substitute axiom) prevents us from applying the typical proof strategy used for logics of causal reasoning; *weak reversibility* seems essential for the usual proof of the truth lemma (cp. Lemma 4.12). Our strategy will be to introduce an (artificial) modal semantics for $\mathcal{H}^+$, so that we can use the standard techniques for proving completeness in modal logics, see e.g. Blackburn et al.^[32]; this was also the reason for switching from $\mathcal{H}$ to $\mathcal{H}^+$. In particular, we will build a canonical modal model whose worlds can be identified (modulo equivalence) with our canonical teams, and prove a truth lemma for it.

Let us fix some notation. We assume that there is a fixed alphabetical order of the variables. For a given signature σ , we write I_σ for the set of all conjunctions of the form $\mathbf{X} = \mathbf{x}$, where the variables in $\mathbf{X}$ are listed alphabetically and without repetitions.

Definition 6.1. *An unintended model (umodel) for σ is a tuple $M = (W, R_\top, (R_{\mathbf{X}=\mathbf{x}})_{\mathbf{X}=\mathbf{x} \in I_\sigma}, V)$, where W is an arbitrary set (of "worlds"), $R_\top$ and the $R_{\mathbf{X}=\mathbf{x}}$ are binary relations on W , and V is a function $W \times \text{Atom}_\sigma \rightarrow \{0, 1\}$.*

We can then define the notion of truth at a world. Given a umodel M , a world u of M and an $\mathcal{H}^+$ formula φ , the relation $M, u \models \varphi$ is given by the following inductive clauses:

- $M, u \models Y = y$ iff $V(u, Y = y) = 1$.
- $M, u \models \psi \& \chi$ iff $M, u \models \psi$ and $M, u \models \chi$.
- $M, u \models \sim \psi$ iff $M, u \not\models \psi$.
- $M, u \models \Box \psi$ iff for all v such that $uR_\top v$, $M, v \models \psi$.
- $M, u \models [\mathbf{X} = \mathbf{x}]\psi$ iff for all v such that $uR_{\mathbf{X}=\mathbf{x}} v$, $M, v \models \psi$.

Umodel–world pairs M, w might in general not be equivalent to any causal team (in the sense that they might not satisfy the same set of $\mathcal{H}^+$ formulas as a causal team). However, we introduce a canonical umodel whose worlds are maximal consistent sets²⁸, and they do behave like causal teams. We need to introduce a bit of extra notation. If $\sigma = (\mathcal{E}, \mathcal{I}, \text{Ran})$ is a signature and $\mathbf{X} \subseteq \mathcal{E} \cup \mathcal{I}$, we denote as $\sigma_\mathbf{X}$ the signature $(\mathcal{E} \cup \mathbf{X}, \mathcal{I} \setminus \mathbf{X}, \text{Ran})$ and call it a **derived signature** of σ . Note that, conveniently, $\mathcal{H}_{\sigma_\mathbf{X}}^+ = \mathcal{H}_\sigma^+$ and $\mathcal{H}_{\sigma_\mathbf{X}} = \mathcal{H}_\sigma$; on the other hand, $B_{\sigma_\mathbf{X}}$ may differ from B_σ (they may contain different instances of axioms I8, J10, J11, J16, and J17). Thus, in general, $\mathbb{T}_{\sigma_\mathbf{X}}^\Gamma$ might differ from $\mathbb{T}_\sigma^\Gamma$, even if Γ is both a set of $\mathcal{H}_{\sigma_\mathbf{X}}^+$ and $\mathcal{H}_\sigma^+$ formulas. We denote by $\hat{\sigma}$ the set of all derived signatures of σ ; note that $\sigma \in \hat{\sigma}$.

The **canonical umodel** $\mathbb{M}_\sigma$ has the following components:

- W is the set of all maximal consistent sets $\Gamma \supseteq B_\tau$ of $\mathcal{H}_\tau^+$ formulas, where $\tau \in \hat{\sigma}$.
- $\Gamma R_\top \Delta$ iff, for some $\tau \in \hat{\sigma}$, $\Gamma, \Delta \supseteq B_\tau$ and:
 1. for some $\mathbf{w} \in \text{Ran}(\mathbf{W})$, $\Diamond \top, \Box \mathbf{W} = \mathbf{w} \in \Delta$
 2. for all $Y \in \mathcal{I}, y \in \text{Ran}(Y)$ and $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$, if $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$, then it is also in Δ
 3. for all $\chi \in \mathcal{H}_\sigma^+$, if $\chi \in \Delta$, then $\Diamond \chi \in \Gamma$.
- $\Gamma R_{\mathbf{X}=\mathbf{x}} \Delta$ iff for some $\tau \in \hat{\sigma}$, we have $\Gamma \supseteq B_\tau$, $\Delta \supseteq B_{\tau_\mathbf{X}}$ and:
 1. for some $\mathbf{w} \in \text{Ran}(\mathbf{W}_\mathbf{X})$, $\Diamond \top, \Box (\mathbf{XW}_\mathbf{X} = \mathbf{xw}) \in \Delta$
 2. for all $Y \in \mathcal{I} \setminus \mathbf{X}, y \in \text{Ran}(Y)$ and $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$, if $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$, then it is also in Δ
 3. for all $\chi \in \mathcal{H}_\sigma^+$, if $\chi \in \Delta$, then $\langle \mathbf{X} = \mathbf{x} \rangle \chi \in \Gamma$.
- $V(\Gamma, Y = y) = 1$ iff $Y = y \in \Gamma$

We will simply write $\mathbb{M}$ when the intended signature is not ambiguous. In the definition of the accessibility relations for $\mathbb{M}$, clause 3 is the standard one used in the definition of canonical models in modal logic. The purpose of the additional clause 1 is to guarantee that, if a world Δ is accessible from some other world, then it behaves analogously as a singleton causal model. This is to mimic the fact that the evaluation of a formula $[\mathbf{X} = \mathbf{x}]\psi$ on a causal model amounts to checking whether ψ holds on *singleton* submodels of the intervened model. Finally, clause 2 (together with 3) guarantees that such singleton models have the same causal laws as the initial causal model—except for the variables $\mathbf{X}$ that have been intervened upon. Concerning these variables, the fact that Δ has signature $\tau_\mathbf{X}$ guarantees that they behave as external variables.

Lemma 6.10. *With notations as above:*

1. In the definition of $\Gamma R_\top \Delta$, clause 3 can be replaced with: for all $\varphi \in \mathcal{H}_\sigma^+$, $\Box \varphi \in \Gamma$ implies $\varphi \in \Delta$.
2. In the definition of $\Gamma R_{\mathbf{X}=\mathbf{x}} \Delta$, clause 3 can be replaced with: for all $\varphi \in \mathcal{H}_\sigma^+$, $[\mathbf{X} = \mathbf{x}]\varphi \in \Gamma$ implies $\varphi \in \Delta$.

Proof. We only prove point 2., as the proof of point 1 is analogous.

From left to right, suppose $\Gamma R_{\mathbf{X}=\mathbf{x}} \Delta$ and $\varphi \notin \Delta$. By Lemma 4.5, 2., $\sim \varphi \in \Delta$. By clause 3 of the definition of $\mathbb{M}$, then, $\langle \mathbf{X} = \mathbf{x} \rangle \sim \varphi \in \Gamma$. By Lemma 6.2, 6., $\sim [\mathbf{X} = \mathbf{x}]\varphi \in \Gamma$. Thus, by the consistency of Γ , $[\mathbf{X} = \mathbf{x}]\varphi \notin \Gamma$.

²⁸ The overall approach used here is inspired by Wang & Cao^[33].

²⁹ Throughout the rest of this section, the notions of consistency and maximal consistency are relative to system B unless otherwise specified.

From right to left, suppose $[X = x]\varphi \in \Gamma$ implies $\varphi \in \Delta$ for all formulas $\varphi \in \mathcal{H}_\sigma^+$. By contraposition, $\varphi \notin \Delta$ (which is equivalent to $\sim\varphi \in \Delta$ by Lemma 4.5, 2.) implies $[X = x]\varphi \notin \Gamma$, thus $\sim[X = x]\varphi \in \Gamma$ by Lemma 4.5, 2., and finally $\langle X = x \rangle \sim\varphi \in \Gamma$ by Lemma 6.2, 6. Since every formula χ is equivalent to a negated formula (say, $\chi \equiv \sim\sim\chi$), this is just a reformulation of clause 3 in the definition of $\mathbb{M}$. $\square$

Lemma 6.11. *Let ψ be an $\mathcal{H}_\sigma^+$ formula and $\Gamma \supseteq B_\sigma$ be a maximal consistent set of $\mathcal{H}_\sigma^+$ formulas.*

1. *Suppose $\diamond\psi \in \Gamma$. Then, there is a $w \in \text{Ran}(W)$ such that $\diamond(\psi \& \Box W = w) \in \Gamma$.*
2. *Suppose $\langle X = x \rangle \psi \in \Gamma$. Then, there is a $w \in \text{Ran}(W_X)$ such that $\langle X = x \rangle(\psi \& \Box XW_X = xw) \in \Gamma$.*

Proof. We prove only statement 2., as the proof of statement 1. is analogous.

Suppose $\langle X = x \rangle \psi \in \Gamma$. By axiom J2, $[X = x] \bigsqcup_{xw \in \text{Ran}(XW_X)} XW_X = xw \in \Gamma$. By axiom I7 and replacement, $[X = x] \bigsqcup_{xw \in \text{Ran}(XW_X)} \Box XW_X = xw \in \Gamma$. Thus, by Lemma 6.2, 10., $\langle X = x \rangle(\psi \& \bigsqcup_{xw \in \text{Ran}(XW_X)} \Box XW_X = xw) \in \Gamma$. By J0 and replacement, $\langle X = x \rangle \bigsqcup_{xw \in \text{Ran}(XW_X)} (\psi \& \Box XW_X = xw) \in \Gamma$. By Lemma 6.2, 8., $\bigsqcup_{xw \in \text{Ran}(XW_X)} \langle X = x \rangle(\psi \& \Box XW_X = xw) \in \Gamma$. By primality of Γ (Lemma 4.5, 4.), there is an $x'w \in \text{Ran}(XW_X)$ such that $\langle X = x \rangle(\psi \& \Box XW_X = x'w) \in \Gamma$. Now, suppose for the sake of contradiction that $x' \neq x$. Note that, from the previous statement, by Lemma 6.2, 11., and J0 we obtain $\langle X = x \rangle \Box XW_X = x'w \in \Gamma$. By I7 and replacement, $\langle X = x \rangle XW_X = x'w \in \Gamma$. Thus, $\langle X = x \rangle X = x' \in \Gamma$ by J0 and monotonicity. But $[X = x]X = x \in \Gamma$ by axiom I4; thus, by Lemma 6.2, 10., $\langle X = x \rangle(X = x \& X = x') \in \Gamma$. By the same lemma, J1 and J0, we obtain $\langle X = x \rangle \perp \in \Gamma$, i.e. $\sim[X = x] \sim \perp \in \Gamma$. On the other hand, $[X = x] \sim \perp \in \Gamma$ by J0 and NEC, contradicting the consistency of Γ . Thus, $x' = x$. $\square$

Lemma 6.12 (Existence lemma). *Let $\Gamma \supseteq B_\sigma$ be a maximal consistent set of $\mathcal{H}_\sigma^+$ formulas, φ an $\mathcal{H}_\sigma^+$ formula. Then:*

1. *if $\diamond\varphi \in \Gamma$, then there is a maximal B_σ -consistent set Δ of $\mathcal{H}_\sigma^+$ formulas such that $\Gamma \cup \Delta$ and $\varphi \in \Delta$.*
2. *if $\langle X = x \rangle \varphi \in \Gamma$, then there is a maximal consistent B_{σ_X} set Δ of $\mathcal{H}_\sigma^+$ formulas such that $\Gamma_{X=x} \cup \Delta$ and $\varphi \in \Delta$.*

Proof. We only prove point 2., as the proof of point 1. is analogous.

Suppose $\langle X = x \rangle \varphi \in \Gamma$. We first observe that, since $\langle X = x \rangle \varphi \in \Gamma$, by Lemma 6.11 there is a $w \in \text{Ran}(W)$ such that $\langle X = x \rangle(\varphi \& \Box W = w) \in \Gamma$ and $w \upharpoonright_X = x$. Let $\Delta_1 = \{\psi \mid [X = x]\psi \in \Gamma\}$, $\Delta_2 = \{\langle W_Y = w \rangle Y = y \in \Gamma \mid Y \in \mathcal{I} \setminus X, wy \in \text{Ran}(W_Y Y)\}$ and $\Delta_0 = \Delta_1 \cup \Delta_2 \cup \{\varphi, \diamond\top, \Box W = w\}$. Suppose for the sake of contradiction that Δ_0 is not consistent in B_{σ_X} . Since B_{σ_X} is finitary, this means that there are formulas $\psi_1, \dots, \psi_n \in \Delta_1$ and $\chi_1 : \langle W_{Y_1} = w_1 \rangle Y_1 = y_1, \dots, \chi_m : \langle W_{Y_m} = w_m \rangle Y_m = y_m \in \Delta_2$ such that $\psi_1, \dots, \psi_n, \chi_1, \dots, \chi_m, \varphi, \diamond\top, \Box W = w \vdash \perp$. By some applications of the deduction theorem and J0, we obtain $\vdash (\psi_1 \& \dots \& \psi_n \& \chi_1 \& \dots \& \chi_m \& \diamond\top) \rightarrow \sim(\varphi \& \Box W = w)$. By NEC, I5, Lemma 6.2, 4. and replacement, $\vdash ([X = x]\psi_1 \& \dots \& [X = x]\psi_n \& [X = x]\chi_1 \& \dots \& [X = x]\chi_m \& [X = x]\diamond\top) \rightarrow [X = x] \sim(\varphi \& \Box W = w)$. We can remove the diamond using axiom J12 and replacement, and then by m applications of J11, we obtain $\vdash ([X = x]\psi_1 \& \dots \& [X = x]\psi_n \& \chi_1 \& \dots \& \chi_m \& [X = x]\top) \rightarrow [X = x] \sim(\varphi \& \Box W = w)$ (applying J11 is correct because $Y_1, \dots, Y_m$ are internal and not in X). But now, $[X = x]\psi_1, \dots, [X = x]\psi_n \in \Gamma$ (by

definition of Δ_1), $\chi_1, \dots, \chi_m \in \Gamma$ (by definition of Δ_2) and $[X = x]\top \in \Gamma$ (by J0 and NEC). Thus, by MP, $[X = x] \sim(\varphi \& \Box W = w) \in \Gamma$. Then, by Lemma 6.2, 7., $\sim\langle X = x \rangle(\varphi \& \Box W = w) \in \Gamma$. But we already know that $\langle X = x \rangle(\varphi \& \Box W = w) \in \Gamma$, so we contradict the consistency of Γ .

Thus, Δ_0 is consistent. By the Lindenbaum lemma, there is a maximal B_{σ_X} -consistent set $\Delta \supseteq \Delta_0$. Note that $[X = x]\psi \in \Gamma$ implies $\psi \in \Delta$ for all ψ . By Lemma 6.10, then, we obtain clause 3 of the definition of $R_{X=x}$. Furthermore, we have that $\diamond\top \in \Delta$ and $\Box W = w \in \Delta$ with $w \upharpoonright_X = x$; this is clause 1 of the definition of $R_{X=x}$. Finally, $\Delta_2 \subseteq \Delta$; this is clause 2. Thus, $\Gamma_{X=x} \cup \Delta$. Finally, we observe that $\varphi \in \Delta$, as needed. $\square$

Lemma 6.13 (Truth lemma for $\mathbb{M}$). *Let $\Gamma \supseteq B_\sigma$ be a maximally consistent set of $\mathcal{H}_\sigma^+$ formulas, and φ an $\mathcal{H}_\sigma^+$ formula. Then, $\mathbb{M}, \Gamma \models \varphi$ iff $\varphi \in \Gamma$.*

Proof. By some applications of I0 and replacement, we can assume up to equivalence that all modal operators $[X = x]$ occurring in φ are replaced with $\sim\sim[X = x]\sim\sim$, i.e., $\sim\langle X = x \rangle\sim$. We can then proceed by induction on the subformulas ψ of φ , thought as combinations of atoms, $\sim$, $\&$ and $\langle X = x \rangle$ operators; we prove the statement simultaneously for all Γ .

If ψ is an atom $Y = y$, then $\mathbb{M}, \Gamma \models Y = y$ iff (by definition of $\models$) $V(\Gamma, Y = y) = 1$ iff (by definition of $\mathbb{M}$) $Y = y \in \Gamma$. The cases for $\psi = \chi_1 \& \chi_2$ or $\psi = \sim\chi$ are also straightforward.

If ψ is of the form $\langle X = x \rangle \chi$, then $\mathbb{M}, \Gamma \models \langle X = x \rangle \chi$ iff (by definition of $\models$) there is a Δ such that $\Gamma_{X=x} \cup \Delta$ and $\mathbb{M}, \Delta \models \chi$, iff (induction hypothesis) there is a Δ such that $\Gamma_{X=x} \cup \Delta$ and $\chi \in \Delta$. By clause 3 of the definition of $R_{X=x}$, then, the latter implies $\langle X = x \rangle \chi \in \Gamma$. Vice versa, if $\langle X = x \rangle \chi \in \Gamma$, by the existence Lemma 6.12 there is a Δ such that $\Gamma_{X=x} \cup \Delta$ and $\chi \in \Delta$. $\square$

Corollary 6.14. *If Γ is a world of $\mathbb{M}_\sigma$ of signature σ_X , then it satisfies the axioms of B_{σ_X} .*

Lemma 6.15 (Canonical singletons). *Let $\Delta \supseteq B_\sigma$ be a maximally consistent set of $\mathcal{H}_\sigma^+$ formulas, and let Δ_0 be the set of $\mathcal{H}$ formulas in Δ . If $\diamond\top, \Box W = s(W) \in \Delta$, then $\Delta_0 = \{\chi \in \mathcal{H}_\sigma \mid (\{s\}, \mathcal{F}^\Delta) \models \chi\}$.³⁰*

Proof. By Lemma 6.2, 9., $\diamond\top, \Box W = s(W) \in \Delta_0 \subseteq \Delta$ implies that $\diamond W = s(W) \in \Delta_0$. Furthermore, by axiom I1, $\vdash \Box W = s(W) \rightarrow \Box W \neq t(W)$ if $s \neq t$; thus, $\Box W \neq t(W) \in \Delta$. If we had $\diamond W = t(W) \in \Delta_0$ for some $t \neq s$, then, by Lemma 6.2, 10., we would obtain $\diamond(W = t(W) \& W \neq t(W))$; thus, by J0 and replacement, $\diamond\perp \in \Delta_0 \subseteq \Delta$; this contradicts Lemma 4.5, 5. Thus, s is the unique assignment such that $\diamond W = s(W) \in \Delta$. By definition of $(\mathbb{T}_B^\Delta)^-$, then, we have $(\mathbb{T}_B^\Delta)^- = \{s\}$. Thus, $(\{s\}, \mathcal{F}^\Delta) = (\{s\}, \mathcal{F}^\Delta) = \mathbb{T}^\Delta$. By corollary 6.6, then, $\Delta = \{\chi \in \mathcal{H}_\sigma^+ \mid \mathbb{T}^\Delta \models \chi\}$; thus, $\Delta_0 = \{\chi \in \mathcal{H}_\sigma \mid \mathbb{T}^\Delta \models \chi\} = \{\chi \in \mathcal{H}_\sigma \mid (\{s\}, \mathcal{F}^\Delta) \models \chi\}$. $\square$

As we said early on, the canonical model $\mathbb{M}$ is designed so that each of its worlds Γ would behave as the canonical causal team $\mathbb{T}^\Gamma$, and each world Δ with $\Gamma_{X=x} \cup \Delta$ would behave as one of the singleton models that are produced by the intervention $do(X = x)$ applied to $\mathbb{T}^\Gamma$. The second of these claims, unfortunately, will not in general be correct; the law component $\mathcal{F}^\Delta$ of $\mathbb{T}^\Delta$ might disagree with the law component $\mathcal{F}_{X=x}^\Gamma$ of $\mathbb{T}_{X=x}^\Gamma$ about the parent sets of some variables.³¹ We will see that, nonetheless, $\mathbb{T}^\Delta$ is $\equiv$ -equivalent to a corresponding singleton submodel of $\mathbb{T}_{X=x}^\Gamma$, and thus it satisfies the same $\mathcal{H}^+$ formulas.

Lemma 6.16. *Let $\tau \in \hat{\sigma}$. Let $\Gamma \supseteq B_\tau$ and $\Delta \supseteq B_{\tau_X}$ be maximal consistent sets of $\mathcal{H}_\sigma^+$ formulas. If $\Gamma_{X=x} \cup \Delta$, then $\mathcal{F}^\Delta \equiv \mathcal{F}_{X=x}^\Gamma$.*

³⁰ Remember that $\mathcal{F}^\Delta$ denotes the law component of the canonical causal team $\mathbb{T}^\Delta$.

³¹ Intuitively, the reason is that, after fixing the value of X to x , information is lost, and relationships of direct cause that were "visible" in Γ are not detectable anymore in Δ .

Proof. The causal laws for an internal variable Y do not necessarily have the same parent set in $\mathcal{F}^\Delta$ as in $\mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma$; so, we shall write PA_Y^Δ , resp. PA_Y^Γ , to differentiate them. Note that $\Gamma R_{\mathbf{X}=\mathbf{x}}\Delta$ entails that $\mathcal{F}^\Delta$ and $\mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma$ have the same signature (namely, $\tau_{\mathbf{X}}$), so that the definition of $\equiv$ may apply. Let us consider a tuple of values $pa \in \text{Ran}(PA_Y^\Delta \cup PA_Y^\Gamma)$; to conclude that $\mathcal{F}^\Delta \equiv \mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma$ we need to prove that $\mathcal{F}_Y^\Delta(pa_{\upharpoonright PA_Y^\Delta}) = \mathcal{F}_Y^\Gamma(pa_{\upharpoonright PA_Y^\Gamma})$ when $Y \in I \setminus \mathbf{X}$.^②

Suppose $y \in \mathcal{F}_Y^\Delta(pa_{\upharpoonright PA_Y^\Delta})$. By definition of $\mathcal{F}^\Delta$, this means that, for any $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$ with $\mathbf{w}_{\upharpoonright PA_Y^\Delta} = pa_{\upharpoonright PA_Y^\Delta}$, $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Delta$. In particular, this holds for any $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$ with $\mathbf{w}_{\upharpoonright PA_Y^\Delta \cup PA_Y^\Gamma} = pa$; fix one such tuple $\mathbf{w}$. By clause 3 of the definition of $R_{\mathbf{X}=\mathbf{x}}$, then, $\langle \mathbf{X} = \mathbf{x} \rangle \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$. Since $Y \notin \mathbf{X}$, by axiom J10 we have $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$. Thus, by definition of $\mathcal{F}_Y^\Gamma$, $y \in \mathcal{F}_Y^\Gamma(pa_{\upharpoonright PA_Y^\Gamma})$.

Vice versa, suppose $y \in \mathcal{F}_Y^\Gamma(pa_{\upharpoonright PA_Y^\Gamma})$. Reasoning as before, there is a $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$ with $\mathbf{w}_{\upharpoonright PA_Y^\Delta \cup PA_Y^\Gamma} = pa$ such that $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$. But then, $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Delta$ by clause 2 of the definition of $R_{\mathbf{X}=\mathbf{x}}$. $\square$

The following two lemmas show that the singleton subteams of a canonical model $\mathbb{T}^\Gamma$, resp. those of $\mathbb{T}_{\mathbf{X}=\mathbf{x}}^\Gamma$, can be identified (up to $\equiv$ -equivalence) with well-chosen canonical teams.

Lemma 6.17. *Let $\Gamma \supseteq B_\sigma$ be a maximal consistent set of $\mathcal{H}_\sigma^+$ formulas. If $s \in (\mathbb{T}^\Gamma)^-$, then $(\{s\}, \mathcal{F}^\Gamma) \equiv \mathbb{T}^\Delta$, where $\Delta = \{\varphi \in \mathcal{H}_\sigma^+ \mid \Diamond(\mathbf{W} = s(\mathbf{W}) \ \& \ \varphi) \in \Gamma\}$. Furthermore, $\Gamma R_\Delta \Delta$.*

Proof. First of all, note that it is easy to show that, for any formula φ , $(\{s\}, \mathcal{F}^\Gamma) \models \varphi$ iff $\mathbb{T}^\Gamma \models \Diamond(\mathbf{W} = s(\mathbf{W}) \ \& \ \varphi)$. Therefore, $(\{s\}, \mathcal{F}^\Gamma) \models \Delta$ and it satisfies no $\mathcal{H}_\sigma^+$ formula outside of Δ . Thus, Δ is a maximal B_σ -consistent set of $\mathcal{H}_\sigma^+$ formulas.

Thus, $\mathbb{T}^\Delta$ is well-defined. Since Δ is the set of *all* formulas satisfied by $(\{s\}, \mathcal{F}^\Gamma)$, and $(\{s\}, \mathcal{F}^\Gamma)$ is a model of signature σ , by the soundness of B_σ we have $\Delta \supseteq B_\sigma$. Let Δ_0 be the set of all $\mathcal{H}$ formulas in Δ . If we can prove that $\mathbb{T}^\Delta$ is a singleton model, then by corollary 6.6 $\mathbb{T}^\Delta \models \Delta_0$, i.e., $\mathbb{T}^\Delta$ and $(\{s\}, \mathcal{F}^\Gamma)$ satisfy the same $\mathcal{H}_\sigma$ formulas; thus, by Theorem 5.1, $\mathbb{T}^\Delta \equiv (\{s\}, \mathcal{F}^\Gamma)$.

Let us then prove that $\mathbb{T}^\Delta$ is a singleton model. Since $s \in (\mathbb{T}^\Gamma)^-$, by definition of $\mathbb{T}^\Gamma$ we have $\Diamond \mathbf{W} = s(\mathbf{W}) \in \Gamma$. Thus, $\Diamond(\mathbf{W} = s(\mathbf{W}) \ \& \ \mathbf{W} = s(\mathbf{W})) \in \Gamma$ by I0 and replacement; so, by the definition of Δ , $\mathbf{W} = s(\mathbf{W}) \in \Delta$. By axiom I7, (*): $\Box \mathbf{W} = s(\mathbf{W}) \in \Delta$.

On the other hand, since $\Diamond \mathbf{W} = s(\mathbf{W}) \in \Gamma$, by axiom J15 $\Diamond(\mathbf{W} = s(\mathbf{W}) \ \& \ \Diamond \top) \in \Gamma$. Thus, $\Diamond \top \in \Delta$. Together with (*), this gives (by Lemma 6.2, 10.) $\Diamond \mathbf{W} = s(\mathbf{W}) \in \Delta$. Thus, by definition of $\mathbb{T}^\Delta$, $s \in (\mathbb{T}^\Delta)^-$.

If $t \neq s$, then $\Diamond \mathbf{W} = t(\mathbf{W}) \in \Delta$, together with (*), would imply (Lemma 6.2, 10.) that $\Diamond(\mathbf{W} = s(\mathbf{W}) \ \& \ \mathbf{W} = t(\mathbf{W})) \in \Delta$. Together with I1 and J0, this yields $\Diamond \perp$, contradicting Lemma 4.5, 5. Thus, for each $t \neq s$, $\Diamond \mathbf{W} = t(\mathbf{W}) \notin \Delta$. So, by definition of $\mathbb{T}^\Delta$, $(\mathbb{T}^\Delta)^-$ is the singleton $\{s\}$.

Lastly, let us prove that $\Gamma R_\Delta \Delta$. We already know that $\Gamma, \Delta \supseteq B_\sigma$. We need to verify the three clauses of the definition of R_Δ .

Clause 1. Since obviously $(\{s\}, \mathcal{F}^\Gamma) \models \Diamond \top$, $\Box \mathbf{W} = s(\mathbf{W})$ and we have shown $(\{s\}, \mathcal{F}^\Gamma) \equiv \mathbb{T}^\Delta$, we have $\mathbb{T}^\Delta \models \Diamond \top$, $\Box \mathbf{W} = s(\mathbf{W})$. Since $\Diamond \top$, $\Box \mathbf{W} = s(\mathbf{W})$ are $\mathcal{H}$ formulas, and furthermore $\mathbb{T}^{\Delta_0} = \mathbb{T}^\Delta$ is a singleton, by corollary 6.6 we obtain $\Diamond \top$, $\Box \mathbf{W} = s(\mathbf{W}) \in \Delta_0 \subseteq \Delta$, as needed.

Clause 2. Let $Y \in I$ (where $\sigma = (\mathcal{E}, I, \text{Ran})$), and suppose $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$. Since we have $\Diamond \mathbf{W} = s(\mathbf{W}) \in \Delta$, by definition of Δ , $\Diamond(\mathbf{W} = s(\mathbf{W}) \ \& \ \mathbf{W} = s(\mathbf{W})) \in \Gamma$. By J0 and replacement, $\Diamond \mathbf{W} = s(\mathbf{W}) \in \Gamma$. On the other hand, by axiom J11, from the initial assumption we get $\Box \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$. So, by Lemma 6.2, 10., we

obtain $\Diamond(\mathbf{W} = s(\mathbf{W}) \ \& \ \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y) \in \Gamma$. Thus, by definition of Δ , $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Delta$.

Clause 3. Suppose $\chi \in \Delta$. By definition of Δ , then, $\Diamond(\mathbf{W} = s(\mathbf{W}) \ \& \ \chi) \in \Gamma$. Thus, $\Diamond \chi \in \Gamma$ by Lemma 6.2, 11. $\square$

Lemma 6.18. *Let $\tau \in \hat{\sigma}$, and $\mathbf{X}$ be a tuple of variables in the domain of σ . Let $\Gamma \supseteq B_\tau$ be a maximal B -consistent set of $\mathcal{H}_\sigma^+$ formulas. If $s \in (\mathbb{T}_{\mathbf{X}=\mathbf{x}}^\Gamma)^-$, then there is a maximal consistent set $\Delta \supseteq B_{\tau_{\mathbf{X}}}$ of $\mathcal{H}_\sigma^+$ formulas such that 1) $\Gamma R_{\mathbf{X}=\mathbf{x}}\Delta$ and 2) $\mathbb{T}_{\mathbf{X}=\mathbf{x}}^\Delta \equiv (\{s\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma)$.*

Proof. In case $\mathbf{X} = \emptyset$, this is just Lemma 6.17. Let us assume henceforth that $\mathbf{X} \neq \emptyset$.

We let $\Delta := \{\varphi \in \mathcal{H}_\sigma^+ \mid \langle \mathbf{X} = \mathbf{x} \rangle (\mathbf{W} = s(\mathbf{W}) \ \& \ \varphi) \in \Gamma\}$. Let us verify that Δ is maximally consistent and $B_{\tau_{\mathbf{X}}} \subseteq \Delta$ (so that $\mathbb{T}^\Delta$ is well-defined), and that it satisfies 1) and 2).

Observe first that, for any formula φ , $(\{s\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma) \models \varphi$ iff $\mathbb{T}_\tau^\Gamma \models \langle \mathbf{X} = \mathbf{x} \rangle (\mathbf{W} = s(\mathbf{W}) \ \& \ \varphi)$. Therefore, $(\{s\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma)$ is a model for Δ , and it satisfies no $\mathcal{H}_\sigma^+$ formulas outside Δ . From this we obtain, exactly as in the proof of Lemma 6.17, that Δ is maximally consistent and (since $(\{s\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma)$ is a model of signature $\tau_{\mathbf{X}}$) $B_{\tau_{\mathbf{X}}} \subseteq \Delta$.

1) We need to show that clauses 1.–3. of the definition of $R_{\mathbf{X}=\mathbf{x}}$ are satisfied.

Clause 1. Since $s \in (\mathbb{T}_{\mathbf{X}=\mathbf{x}}^\Gamma)^-$, we have $\mathbb{T}_\tau^\Gamma \models \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{W} = s(\mathbf{W})$. But then, there is a $t \in (\mathbb{T}^\Gamma)^-$ such that $(\{t\}, \mathcal{F}^\Gamma) \models \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{W} = s(\mathbf{W})$. Now, by Lemma 6.17, $(\{t\}, \mathcal{F}^\Gamma) \equiv \mathbb{T}^\Delta$ for the maximal consistent set $\Delta := \{\varphi \in \mathcal{H}_\sigma^+ \mid \Diamond(\mathbf{W} = t(\mathbf{W}) \ \& \ \varphi) \in \Gamma\}$; thus, by Theorem 5.1, $\mathbb{T}^\Delta \models \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{W} = s(\mathbf{W})$. Since this is an $\mathcal{H}$ formula, and $\mathbb{T}^\Delta$ is a singleton model, by corollary 6.6 we have $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{W} = s(\mathbf{W}) \in \Delta$. Therefore, $\Diamond(\mathbf{W} = t(\mathbf{W}) \ \& \ \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{W} = s(\mathbf{W})) \in \Gamma$. By Lemma 6.2, 11., $\Diamond \langle \mathbf{X} = \mathbf{x} \rangle \mathbf{W} = s(\mathbf{W}) \in \Gamma$. Thus, by axiom J14, $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{W} = s(\mathbf{W}) \in \Gamma$. By axiom J15, $\langle \mathbf{X} = \mathbf{x} \rangle (\mathbf{W} = s(\mathbf{W}) \ \& \ \Diamond \top) \in \Gamma$. Thus, by definition of Δ , we have $\Diamond \top \in \Delta$, as needed.

Note also that, from $\langle \mathbf{X} = \mathbf{x} \rangle (\mathbf{W} = s(\mathbf{W})) \in \Gamma$, by J0 and replacement we obtain $\langle \mathbf{X} = \mathbf{x} \rangle (\mathbf{W} = s(\mathbf{W}) \ \& \ \mathbf{W} = s(\mathbf{W})) \in \Gamma$. Thus, by definition of Δ , $\mathbf{W} = s(\mathbf{W}) \in \Delta$. Since we have already shown $B_{\tau_{\mathbf{X}}} \subseteq \Delta$, and Δ , being maximally consistent, is closed under $\vdash_{B_{\tau_{\mathbf{X}}}}$ (Lemma 4.5, 1.) by axiom I7, we conclude $\Box \mathbf{W} = s(\mathbf{W}) \in \Delta$. If we can prove that $s(\mathbf{X}) = \mathbf{x}$, then, we are done. But that immediately follows from the assumption $s \in (\mathbb{T}_{\mathbf{X}=\mathbf{x}}^\Gamma)^-$.

Clause 2. Suppose $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$ for an internal $Y \notin \mathbf{X}$. From this we obtain, by axiom J11, that (*): $[\mathbf{X} = \mathbf{x}] \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$. On the other hand, note that, since $\Diamond \top$, $\Box \mathbf{W} = s(\mathbf{W}) \in \Delta$, by Lemma 6.2, 10, we have $\Diamond \mathbf{W} = s(\mathbf{W}) \in \Delta$. By definition of Δ , then, $\langle \mathbf{X} = \mathbf{x} \rangle (\mathbf{W} = s(\mathbf{W}) \ \& \ \Diamond \mathbf{W} = s(\mathbf{W})) \in \Gamma$. By Lemma 6.2, 11., $\langle \mathbf{X} = \mathbf{x} \rangle \Diamond \mathbf{W} = s(\mathbf{W}) \in \Gamma$. By axiom J13, then, $\langle \mathbf{X} = \mathbf{x} \rangle \mathbf{W} = s(\mathbf{W}) \in \Gamma$. This, together with (*), yields $\langle \mathbf{X} = \mathbf{x} \rangle (\mathbf{W} = s(\mathbf{W}) \ \& \ \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y) \in \Gamma$ by Lemma 6.2, 10. Thus, by definition of Δ , $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Delta$.

Clause 3. Suppose $\chi \in \Delta$. By definition of Δ , $\langle \mathbf{X} = \mathbf{x} \rangle (\mathbf{W} = s(\mathbf{W}) \ \& \ \chi) \in \Gamma$. By Lemma 6.2, 11., then, $\langle \mathbf{X} = \mathbf{x} \rangle \chi \in \Gamma$.

2) Let Δ_0 be the set of $\mathcal{H}$ formulas in Δ . Since $\mathbb{T}^\Delta$ is a singleton, by corollary 6.6, we have $\mathbb{T}^\Delta \models \Delta_0$. Thus, $\mathbb{T}^\Delta$ and $(\{s\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma)$ satisfy the same $\mathcal{H}_\sigma$ formulas. By Theorem 5.1, then, $\mathbb{T}^\Delta \equiv (\{s\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma)$. $\square$

Lemma 6.19 (Representation lemma). *Let $\tau \in \hat{\sigma}$, and let $\Gamma \supseteq B_\tau$ be a maximally consistent set of $\mathcal{H}_\sigma^+$ formulas, φ an $\mathcal{H}_\sigma^+$ formula. Then, $\mathbb{M}, \Gamma \models \varphi$ iff $\mathbb{T}^\Gamma \models \varphi$.*

Proof. By induction on φ : The cases for $\&$ and $\sim$ are straightforward.

Atomic case: φ is $Y = y$. Since $Y = y$ is atomic, corollary 6.6 tells us that $\mathbb{T}^\Gamma \models Y = y$ iff $Y = y \in \Gamma$, iff (by definition of $\mathbb{M}$) $V(\Gamma, Y = y) = 1$ iff (by definition of $\models$) $\mathbb{M}, \Gamma \models Y = y$.

^② Note that $\mathcal{F}_Y^\Gamma = (\mathcal{F}_{\mathbf{X}=\mathbf{x}}^\Gamma)_Y$ when $Y \in I \setminus \mathbf{X}$.

Case φ is $[X = x]\psi$.

$\Rightarrow$) Suppose $\mathbb{M}, \Gamma \Vdash [X = x]\psi$. Since $\mathbb{M}, \Gamma \Vdash [X = x]\psi$, we have that, for all Δ such that $\Gamma R_{X=x}\Delta$, $\mathbb{M}, \Delta \Vdash \psi$. By the i.h., $\mathbb{T}^\Delta \models \psi$. Therefore, if we show that 1) $\mathbb{T}^\Delta$ is a singleton model, 2) every singleton submodel of $\mathbb{T}_{X=x}^\Gamma$ is $\equiv \mathbb{T}^\Delta$ for some Δ with $\Gamma R_{X=x}\Delta$, we will have $\mathbb{T}^\Gamma \models [X = x]\psi$ by the semantic clauses for $\models$.

2) This point is given by Lemma 6.18 (or Lemma 6.17 in the special case for $X = \emptyset$).

1) Since $\Gamma R_{X=x}\Delta$, by definition, $\Diamond T \in \Delta$ and $\Box W = w \in \Delta$ for some w such that $w|_X = x$. By Lemma 6.2, 9., $\Diamond W = w \in \Delta$; thus, $s \in (\mathbb{T}^\Delta)^-$, where s is the assignment $s(W) = w$. Suppose, for the sake of contradiction, that $t \in (\mathbb{T}^\Delta)^-$, where $t(W) = w' \neq w \in \Delta$. By definition of $\mathbb{T}^\Delta$, then, $\Diamond W = w' \in \Delta$. On the other hand, $\Box W = w \in \Delta$ together with axiom I1 yields $\Box W \neq w'$. Thus, by Lemma 6.2, 10., $\Diamond(W = w' \ \& \ W \neq w') \in \Delta$. By I0 and monotonicity, then, $\Diamond \perp \in \Delta$. Since Δ is consistent, we contradict Lemma 4.5, 5.

$\Leftarrow$) Assume $\mathbb{T}^\Gamma \models [X = x]\psi$. Then, for all $s \in (\mathbb{T}_{X=x}^\Gamma)^-$, $(\{s\}, \mathcal{F}_{X=x}^\Gamma) \models \psi$. Suppose we manage to prove the following, for all maximally consistent sets of $\mathcal{H}_\sigma^+$ formulas $\Delta \supseteq B_{\tau_X}$:

(*) : if $\Gamma R_{X=x}\Delta$, then for some $s \in (\mathbb{T}_{X=x}^\Gamma)^-$, $\Delta = \{\chi \in \mathcal{H}_\sigma^+ \mid (\{s\}, \mathcal{F}_{X=x}^\Gamma) \models \chi\}$.

If so, then by the above $\psi \in \Delta$ if $\Gamma R_{X=x}\Delta$. But then, by the truth lemma for $\mathbb{M}$ (Lemma 6.13), $\mathbb{M}, \Delta \Vdash \psi$ if $\Gamma R_{X=x}\Delta$. Thus, $\mathbb{M}, \Gamma \Vdash [X = x]\psi$, as needed.

Let us then prove (*). If $\Gamma R_{X=x}\Delta$, then by definition of $R_{X=x}$ we have $\Diamond T, \Box W = w^* \in \Delta$ for some w^* ; let s be the assignment such that $s(W) = w^*$. Let Δ_0 be the set of $\mathcal{H}$ formulas in Δ . By Lemma 6.15, we have that $\Delta_0 = \{\chi \in \mathcal{H}_\sigma \mid (\{s\}, \mathcal{F}^\Delta) \models \chi\}$. By Lemma 6.16, $\mathcal{F}^\Delta \equiv \mathcal{F}_{X=x}^\Gamma$; but then, by Theorem 5.1, $(\{s\}, \mathcal{F}^\Delta)$ and $(\{s\}, \mathcal{F}_{X=x}^\Gamma)$ satisfy the same $\mathcal{H}_\sigma^+$ formulas. Thus, $\Delta_0 = \{\chi \in \mathcal{H}_\sigma \mid (\{s\}, \mathcal{F}_{X=x}^\Gamma) \models \chi\}$. Since Δ_0 has a singleton model of signature τ_X , by the soundness of A we have $A_{\tau_X} \subseteq \Delta_0$, and Δ_0 is maximally A_{τ_X} -consistent. Thus, by Theorem 6.9, Δ is the only maximal consistent set of $\mathcal{H}^+$ formulas that extends both Δ_0 and B_{τ_X} . Now, $\Delta' := \{\chi \in \mathcal{H}_\sigma^+ \mid (\{s\}, \mathcal{F}_{X=x}^\Gamma) \models \chi\}$ also extends Δ_0 . Since $(\{s\}, \mathcal{F}_{X=x}^\Gamma)$ is a model of Δ' , we also have $B_{\tau_X} \subseteq \Delta'$. Thus, $\Delta = \Delta'$, as needed. $\square$

Theorem 6.20 (Strong completeness of B). For $\Gamma \cup \{\varphi\} \subseteq \mathcal{H}_\sigma^+$,

$$\Gamma \models_\sigma \varphi \iff \Gamma \vdash_{B_\sigma} \varphi.$$

Proof. The right-to-left direction is given by the soundness theorem 6.1. In the other direction, suppose $\Gamma \not\models_{B_\sigma} \varphi$. By the usual reasoning, then, $\Gamma \cup B_\sigma \cup \{\sim\varphi\}$ is B_σ -consistent. By the Lindenbaum lemma, there is a maximal consistent set of $\mathcal{H}_\sigma^+$ formulas $\Delta \supseteq \Gamma \cup B_\sigma \cup \{\sim\varphi\}$. By the truth lemma for $\mathbb{M}$ (Lemma 6.13), $\mathbb{M}, \Delta \Vdash \Gamma, \sim\varphi$. But then, by the representation lemma (Lemma 6.19), $\mathbb{T}^\Delta \models \Gamma, \sim\varphi$. Since $\mathbb{T}^\Delta$ is a model (by Lemma 6.4), we conclude that $\Gamma \not\models \varphi$. $\square$

As a final remark, we note that now we may obtain an axiomatization of the $\mathcal{H}_\sigma^+$ languages over *singleton* models as an extension of B_σ . Consider the following axiom:

Sing. $\Diamond T \ \& \ \bigwedge_{w \in \text{Ran}(W)} W = w$.

Theorem 6.21. For $\Gamma \cup \{\varphi\} \subseteq \mathcal{H}_\sigma^+$, we have $\Gamma \models_\sigma^1 \varphi \iff \Gamma, \text{Sing} \vdash_{B_\sigma} \varphi$.

Proof. This is an immediate consequence of Theorem 6.3 and Theorem 4.13. $\square$

We are left instead with the open issue of axiomatizing language $\mathcal{H}_\sigma$ over the class of all models (of signature σ). We might conjecture that system B_σ could be conservative over the subsystem of axioms involving only unnested formulas (axioms J0, I1-I5, J6, I7-I9, and J16); the completeness for $\mathcal{H}_\sigma$ formulas of this smaller set of axioms

would immediately follow. But nothing of what has been done here supports this conjecture.

7 Parents, direct causes, and dummy arguments

The literature on causal inference defines a multitude of notions of causation in terms of interventionist counterfactuals; see, e.g., Woodward^[34]. Among these, a key foundational role is played by the notion of *direct cause*; roughly speaking, X is a direct cause of Y when there is some intervention on X that alters the value of Y *while all other variables in the system are held fixed* (at some tuple of values). In most discussions of causal models, saying that X is a direct cause of Y is equivalent to saying that X is a parent of Y ; furthermore, it is equivalent to saying that X is a non-dummy argument of the causal law for Y . As hinted at in section 2, these three perspectives diverge in the indeterministic case. In the present section, we give precise definitions of these notions and prove some basic connections between them. In particular, we show that, for any model T ,

$$PA_V \supseteq^* \{\text{direct causes of } V\} \supseteq \{\text{non-dummy arguments of } V\}$$

and there are models that make these inclusions strict. We marked the first inclusion with a star because it holds only under the assumption that the model is *total*.

We will also consider a fourth relationship, that of *visible direct cause*, which, while not being properly causal in nature, will be useful for the purposes of axiomatization. We will prove the following important inclusion:

$$\{\text{direct causes of } V\} \supseteq \{\text{visible direct causes of } V\}.$$

7.1 Definitions

1. Whether X is a parent of Y in a model T is arbitrarily decided by the model (this relation determines the **parenthood graph**). As we have seen, the effect of this stipulation is that, if X is updated by an intervention, then Y is also updated. If the parenthood graph of T is acyclic, we say that T is **strictly recursive**. As we already remarked, this class of models has many analogies with the recursive class of (the usual, deterministic) causal models. While it is no more true that the values for the external variables uniquely determine the values of the internal variables, it is still the case that the effects of interventions can be calculated by a straightforward recursive procedure.

Another analogy is that, if furthermore the causal laws are total, then intervening on a nonempty model always produces solutions.

Proposition 7.1. Let $T = (T^-, \mathcal{F})$ be a nonempty, strictly recursive total model, $\mathbf{X}$ variables in its domain, $\mathbf{x} \in \text{Ran}(\mathbf{X})$, $s \in T^-$. Then, $s_{\mathbf{X}=\mathbf{x}}^\mathcal{F}$ is nonempty. In particular, $T_{\mathbf{X}=\mathbf{x}}$ is nonempty.

Proof. As already mentioned in section 3, the acyclicity of the parenthood graph ensures that, if $Z \in PA_Y$, then $d(\mathbf{X}, Z) < d(\mathbf{X}, Y)$. Furthermore, acyclicity and finiteness ensure that each variable has a (finite) distance from $\mathbf{X}$.

Fix an $s \in T^-$. We define a $t \in s_{\mathbf{X}=\mathbf{x}}^\mathcal{F}$ by recursion on the distance of variables from $\mathbf{X}$. First, we let $t(\mathbf{X}) := \mathbf{x}$ and $t(\mathbf{N}_\mathbf{X}) := s(\mathbf{N}_\mathbf{X})$. Each of the remaining variables, say Y , is internal and not in $\mathbf{X}$. Since $d(\mathbf{X}, Z) < d(\mathbf{X}, Y)$ for each $Z \in PA_Y$, we can assume by i.h. that $t(Z)$ is already defined for each $Z \in PA_Y$. Then, we can let $t(Y) := y$ for an arbitrary $y \in \mathcal{F}_Y(t(PA_Y)) \neq \emptyset$. Since there is a maximum distance, this defining procedure ends. $\square$

2A. We define the notion of direct cause following the suggestion of Wysocki^[3]: that X should count as a direct cause of Y if intervening

on X (while all the remaining variables are fixed at some values) will change the range of available values of Y . More precisely, let $T = (T^-, \mathcal{F})$ be a model of signature σ . We will say that X is a **direct cause** of Y (in T) if, for some assignment s of signature σ compatible with $\mathcal{F}$, the following formula $X \rightsquigarrow Y$ is satisfied by $(\{s\}, \mathcal{F})$:

$$X \rightsquigarrow Y : \bigcup_{(z, x, y) \in \text{Ran}(ZY)} \sim(\langle Zx = zx \rangle Y = y \leftrightarrow \langle Z = z \rangle Y = y),$$

(where, Z stands for $\text{Dom} \setminus \{XY\}$).³³ Note that this notion of direct cause, in agreement with previous literature on causation, only depends on the law component, and not the team component of the model. If there being no variable is a direct cause of Y , we say that Y is **exogenous**; otherwise, it is **endogenous**.

We call **causal graph** the graph of the direct cause relation, and say that a model is **recursive** if its causal graph is acyclic. In section 10.1 we prove that direct cause and the related notions are not definable in $\mathcal{H}$. We do not know if they may be definable in some stronger language such as $\mathcal{H}^+$.

2B. We will say that X is a **visible direct cause** of Y (in T) if $T \models X \rightsquigarrow Y$. This is not equivalent to X being a direct cause of Y , and there are reasons to think that such a definition is not causal. In the example of two-coin Bob (Example 2.2), B turns out to be both a direct cause of O according to the definition in 2A and a visible direct cause. But imagine a variant T' of the two-coin Bob model T in which, instead of one assignment, we have two:

B	O
1	heads
1	tails

Intervening on B does not change the available values of O in T' ; thus, B is not a visible direct cause of O in T' . Then we have two models (T and T') that have the same set of causal laws, but disagree on whether B is a visible direct cause of O or not. We believe that a causal notion should not depend on the team component (which is typically used to describe *epistemic uncertainty*). Thus, "visible direct cause" does not fit the bill. On the other hand, clearly, the definition of direct cause does not depend on the team component (in particular, T and T' agree about it).

We consider the notion of visible direct cause because it is a useful tool for the sake of axiomatization. Differently from direct cause, it is easily definable in $\mathcal{H}$ (by the formula $X \rightsquigarrow Y$); we can then define related notions such as Y being **visibly exogenous**, resp. **visibly endogenous**, by saying that the model satisfies the following formulas:

$$\bullet \varphi_{\text{VExo}(Y)} : \&_{X \in \text{Dom} \setminus \{Y\}} \sim X \rightsquigarrow Y \quad \bullet \varphi_{\text{VEnd}(Y)} : \sim \varphi_{\text{VExo}(Y)}.$$

We may call **visible causal graph** the graph of the relation $\rightsquigarrow$, and say that a model is **visibly recursive** if its visible causal graph is acyclic.

A key property of an external variable Y is the fact that it is not affected when intervening on the set of all other variables in the system.³⁴ For a visibly exogenous variable, we do not get that much (We will soon show that such an intervention may make some new values for the visibly exogenous variable appear), but at least we can prove that such an intervention does not make values disappear. Remember that we write $T^-(Y)$ for the set of values that Y takes in T^- , i.e., $T^-(Y) := \{y \in \text{Ran}(Y) \mid \text{there is an } s \in T^- \text{ such that } s(Y) = y\}$.

Proposition 7.2. *Let $T = (T^-, \mathcal{F})$ be a model such that $T \models \varphi_{\text{VExo}(Y)}$; furthermore, let $y \in T^-(Y)$. Let $\mathbf{w}$ be a tuple of values for $\mathbf{W}_Y$ (a list of all variables except Y). Then, $y \in T_{\mathbf{W}_Y=\mathbf{w}}^-(Y)$.*

³³ In a 2-variable model, the formula on the right-hand side of the equivalence would read $\Diamond Y = y$.

³⁴ More generally, interventions on smaller sets of variables will not produce new values for Y — but some values may disappear, at least in the non-strictly recursive case.

Proof. Let $\mathbf{w}, \mathbf{w}'$ be two tuples of values for $\mathbf{W}_Y$. We first show that $y \in T_{\mathbf{W}_Y=\mathbf{w}}^-(Y)$ iff $y \in T_{\mathbf{W}_Y=\mathbf{w}'}^-(Y)$, by induction on the number n of variables on which $\mathbf{w}, \mathbf{w}'$ disagree. If $n = 1$, this follows immediately from $T \models \varphi_{\text{VExo}(Y)}$. Suppose that it holds for n , and assume that $\mathbf{w}, \mathbf{w}'$ differ on $n+1$ variables $X_1, \dots, X_{n+1}$; say, these take values $x_1, \dots, x_{n+1}$ in $\mathbf{w}$ and $x'_1, \dots, x'_{n+1}$ in $\mathbf{w}'$. Let also Z list the variables in $\mathbf{W}_Y \setminus \{X_1, \dots, X_{n+1}\}$ and $\mathbf{z}$ be the restriction of $\mathbf{w}$ to Z . By the inductive hypothesis (for n), we have $y \in T_{\mathbf{W}_Y=\mathbf{w}'}^-(Y)$ iff $y \in T_{\mathbf{z}, X_1, \dots, X_{n+1}=\mathbf{z}, x'_1, \dots, x'_{n+1}}^-(Y)$. By the base case ($n = 1$), the latter is equivalent to $y \in T_{\mathbf{W}_Y=\mathbf{w}}^-(Y)$.

Now let $s \in T^-$ such that $s(Y) = y$, and let $\mathbf{w}^* := s(\mathbf{W}_Y)$. We can show that $s \in T_{\mathbf{W}_Y=\mathbf{w}^*}^-$. This is by definition in case Y is external; if it is internal, from $s \in T^-$ we know that $y = s(Y) \in \mathcal{F}_Y(\mathbf{w}_{\uparrow PA_Y}^*)$; thus again, by definition of intervention, $s \in T_{\mathbf{W}_Y=\mathbf{w}^*}^-$. Thus, $y \in T_{\mathbf{W}_Y=\mathbf{w}^*}^-(Y)$. By the above, then, $y \in T_{\mathbf{W}_Y=\mathbf{w}}^-(Y)$ for any $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$. $\square$

In other words, any model T satisfies the formula $\varphi_{\text{VExo}(Y)} \rightarrow (\Diamond Y = y \rightarrow \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y)$. As we have seen in section 4, if Y is an external variable, the stronger principle $\Diamond Y = y \leftrightarrow \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$ holds. Let us see an example showing that the right-to-left implication may fail if Y is visibly exogenous, but not external.

Example 7.3 (Refereed two-coin Bob). *We have an external variable A for whether Arbiter chooses coin 1 or coin 2; after Arbiter makes their choice, Bob will toss the corresponding coin ($B = 1$ or 2); variable C registers the outcome of the toss (heads or tails). Let T be the model with $PA_B = A$, $PA_C = B$, deterministic law $\mathcal{F}_B(x) = x$ and indeterministic law $\mathcal{F}_C(0) = \mathcal{F}_C(1) = \{\text{heads}, \text{tails}\}$, and containing only the assignment $s(A, B, C) = (1, 1, \text{heads})$ (the scenario in which Arbiter chose coin 1, and Bob tossed coin 1 and it came heads). Now, it can be checked that in this model, neither A nor B is a visible direct cause of C (A is not even a direct cause), the reason being that, once we intervene on A (resp. B), all possible values of C become available, and they will not disappear when intervening also on the other variable. Thus, C is visibly exogenous. On the other hand, intervening on A alone makes the value tails possible, even if that was not possible in the initial model. In other words, we have $T \models \langle A = 1 \rangle C = \text{tails}$ but $T \not\models \Diamond C = \text{tails}$.*

3. Finally, we say $X(\neq Y)$ is a **non-dummy argument** of Y if $X \in PA_Y$ and there are $pa, pa' \in \text{Ran}(PA_Y)$ that coincide on all variables except X , such that $\mathcal{F}_Y(pa) \neq \mathcal{F}_Y(pa')$. If $X \in PA_Y$ but the other condition is not met, X is a **dummy argument** of Y . Note that, by definition, if a variable is not a parent of Y , then it is neither a dummy nor a non-dummy argument of Y . If the graph of the non-dummy argument relation is acyclic, we may say that the model is **weakly recursive**.

7.2 Parents vs. direct causes

Theorem 7.4. 1. *Let T be a model, with a variable X that is either external or internal with a total causal law. If X is a direct cause of Y in T , then $X \in PA_Y$.*

2. *There are (total) models T such that $X \in PA_Y$ but X is not a direct cause of Y in T .*

Proof. 1) Let $T = (T^-, \mathcal{F})$. Assume $X \notin PA_Y$, and that $\mathcal{F}_X$ is a total multivalued function. We show that then, for any $\mathbf{w} \in \text{Ran}(\mathbf{W}_{XY})$, $x \in \text{Ran}(X)$, $y \in \text{Ran}(Y)$, and any $s \in T^-$, $(\{s\}, \mathcal{F}) \models \langle \mathbf{W}_{XY} = \mathbf{w} \rangle Y = y \leftrightarrow \langle \mathbf{W}_{XY} = \mathbf{w} \rangle Y = y$ (thus, X is not a direct cause of Y).

Suppose first that $(\{s\}, \mathcal{F}) \models \langle \mathbf{W}_{XY} = \mathbf{w} \rangle Y = y$. In case Y is internal, since we assumed $X \notin PA_Y$, by definition of intervention, we have

$y \in \mathcal{F}_Y(\mathbf{w}_{\uparrow PA_Y})$. Thus, $(\{s\}, \mathcal{F})_{\mathbf{w}_{XY}X=\mathbf{w}_x}$ is nonempty (it contains the assignment $s(\mathbf{w}_{XY}XY = \mathbf{w}_x)$); thus, we again have, by definition of intervention, that $(\{s\}, \mathcal{F}) \models \langle \mathbf{w}_{XY}X = \mathbf{w}_x \rangle Y = y$. In case Y is external, the reasoning is similar (since $(\{s\}, \mathcal{F}) \models \langle \mathbf{w}_{XY} = \mathbf{w} \rangle Y = y$ then entails $s(Y) = y$).

Vice versa, suppose $(\{s\}, \mathcal{F}) \models \langle \mathbf{w}_{XY}X = \mathbf{w}_x \rangle Y = y$. Since we assumed $X \notin PA_Y$, we can observe that the parenthood graph $G_{\mathbf{w}_{XY}}$ is acyclic. Thus, since $y \in \mathcal{F}_Y(\mathbf{w}_{\uparrow PA_Y}) = \mathcal{F}_Y(\mathbf{w}_{\uparrow PA_Y})$ (so that, in particular, $\mathcal{F}_Y(\mathbf{w}_{\uparrow PA_Y}) \neq \emptyset$), and by the totality assumption also $\mathcal{F}_X(\mathbf{w}_{\uparrow PA_X}) \neq \emptyset$ in case X is internal, by definition of intervention, we then have $(\{s\}, \mathcal{F})_{\mathbf{w}_{XY}=\mathbf{w}} \neq \emptyset$. Thus, since $y \in \mathcal{F}_Y(\mathbf{w}_{\uparrow PA_Y})$, we conclude, by definition of intervention, that $(\{s\}, \mathcal{F}) \models \langle \mathbf{w}_{XY} = \mathbf{w} \rangle Y = y$.

2) Consider a model with internal variable Y and external variables Z, X ; $PA_Y = \{Z, X\}$; $\mathcal{F}_Y$ being the full ternary relation on $\{0, 1\}$; and team component $\{s\} = \{(Z, z^*), (X, x^*), (Y, y^*)\}$, for any $z^*, x^*, y^* \in \{0, 1\}$. It can be checked systematically that any intervention on Z alone, and any intervention on both Z and X lead to a model in which both values 0 and 1 are available for Y . Thus, for any $z, x, y \in \{0, 1\}$, $(\{s\}, \mathcal{F}) \models \langle ZX = z, x \rangle Y = y \leftrightarrow \langle Z = z \rangle Y = y$. Since this holds for all s allowed by the signature, X is not a direct cause of Y . $\square$

To see that the assumption that X has a total causal law is needed for part 1, we may consider the following model:

$Z \rightarrow X \leftarrow Y$
1 1 1

where, the causal law $\mathcal{F}_X(Z, Y)$ for X is such that $\mathcal{F}_X(1, 1) = \{1\}$ and $\mathcal{F}_X(2, 1) = \emptyset$. Then, $T \models \langle Z = 2, X = 1 \rangle Y \neq 2$ (since $T_{Z=2, X=1}^-$ contains the assignment $s(Z, X, Y) = (2, 1, 1)$) while it is not the case that $T \models \langle Z = 2 \rangle Y \neq 2$ (since $T_{Z=2}$ is empty). Thus, X is a direct cause of Y even if X is not a parent of Y .

7.3 Direct causes vs. non-dummy arguments

If $\mathbf{z}$ is a tuple of values for $\mathbf{Z}$, and $X \in \mathbf{Z}$, we write $\mathbf{z}_{-X}$ for the restriction of $\mathbf{z}$ to $\mathbf{Z} \setminus \{X\}$.

Theorem 7.5. 1. If X is a non-dummy argument of Y , then it is a direct cause of Y .

2. There are models T such that X is a direct cause of Y but also a dummy argument of Y .

Proof. 1) Let T be a model, and suppose X is a non-dummy argument of Y in T . Then, there are $pa, pa' \in \text{Ran}(PA_Y)$ such that $pa_{-X} = pa'_{-X}$ but $\mathcal{F}_Y(pa) \neq \mathcal{F}_Y(pa')$. Let $x = pa_{\uparrow \{X\}}$ and $x' = pa'_{\uparrow \{X\}}$; pick a $\mathbf{w} \in \text{Ran}(\mathbf{w}_{XY})$ such that $\mathbf{w}_{\uparrow PA_Y \setminus \{X\}} = pa_{-X}$. Let $s \in T^-$ and $S = \{y \in \text{Ran}(Y) \mid (\{s\}, \mathcal{F}) \models \langle \mathbf{w}_{XY} = \mathbf{w} \rangle Y = y\}$. S cannot be equal to both $\mathcal{F}_Y(pa)$ and $\mathcal{F}_Y(pa')$; wlog suppose $S \neq \mathcal{F}_Y(pa)$.

We note first that there cannot be a $y \in S \setminus \mathcal{F}_Y(pa)$. Indeed, $(\{s\}, \mathcal{F}) \models \langle \mathbf{w}_{XY} = \mathbf{w} \rangle Y = y$ would imply the existence of an assignment in $\mathcal{F}_{\mathbf{w}_{XY}=\mathbf{w}}$ that is not compatible with $\mathcal{F}_Y$.

Thus, there is instead a $y \in \mathcal{F}_Y(pa) \setminus S$. Then $(\{s\}, \mathcal{F}) \models \langle \mathbf{w}_{XY}X = \mathbf{w}_x \rangle Y = y$, but $(\{s\}, \mathcal{F}) \not\models \langle \mathbf{w}_{XY} = \mathbf{w} \rangle Y = y$. So, X is a direct cause of Y .

2) In the two-coin Bob example, B is a direct cause and a dummy argument of O . $\square$

7.4 Direct causes vs. visible direct causes

Theorem 7.6. If X is a visible direct cause of Y in T , then it is a direct cause of Y in T .

Proof. Suppose $T = (T^-, \mathcal{F}) \models X \rightsquigarrow Y$. Then, there are $\mathbf{w}, x, y \in \text{Ran}(\mathbf{w}_{XY}XY)$ such that either:

- $T \models \langle \mathbf{w}_{XY} = \mathbf{w} \rangle Y = y$ but $T \not\models \langle \mathbf{w}_{XY}X = \mathbf{w}_x \rangle Y = y$. By the definition of intervention (plus the related observations in section 3.1), $T \models \langle \mathbf{w}_{XY} = \mathbf{w} \rangle Y = y$ entails that there is an $\hat{s} \in T^-$ and a $t \in \mathcal{F}_{\mathbf{w}_{XY}=\mathbf{w}}$

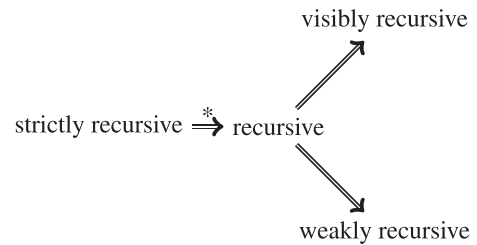
such that $t(Y) = y$. But then $(\{\hat{s}\}, \mathcal{F}) \models \langle \mathbf{w}_{XY} = \mathbf{w} \rangle Y = y$. On the other hand, by a similar argument $T \not\models \langle \mathbf{w}_{XY}X = \mathbf{w}_x \rangle Y = y$ entails that, for all $s \in T^-$ (in particular $\hat{s}$), $(\{s\}, \mathcal{F}) \not\models \langle \mathbf{w}_{XY}X = \mathbf{w}_x \rangle Y = y$. Thus, $(\{\hat{s}\}, \mathcal{F}) \models X \rightsquigarrow Y$. So, X is a direct cause of Y .

- $T \models \langle \mathbf{w}_{XY}X = \mathbf{w}_x \rangle Y = y$ but $T \not\models \langle \mathbf{w}_{XY} = \mathbf{w} \rangle Y = y$. The proof for this case is analogous. $\square$

8 Acyclicity of causal laws

8.1 Axiomatization

Besides the full class of models, we want to characterize axiomatically the class of models in which the causal laws are acyclic. In the deterministic context, such models are called *recursive*. It is not straightforward how to identify the analogue of this notion in the indeterministic context; in the previous section, we have identified four possible candidates: the *strictly recursive*, *recursive*, *visibly recursive*, and *weakly recursive* indeterministic models. We also identified a few logical connections between them, as illustrated in the following picture:



where the asterisked entailment only holds for total models.

We show here how to axiomatize the *strictly* and the *visibly* recursive classes. We do not know how to axiomatize the remaining two classes, but in the next section we will see that the class of recursive, total models can be straightforwardly axiomatized.

In order to turn an axiomatization of the general class of models into one for the strictly (resp. visibly) recursive class, it will suffice to add a recursivity axiom of the form

$$R. (X_1 \rightsquigarrow X_2 \& \dots \& X_{n-1} \rightsquigarrow X_n) \rightarrow \sim X_n \rightsquigarrow X_1. \quad [\text{Generalized recursivity}]$$

Note that $\rightsquigarrow$ does not capture the notion of direct cause but rather the stricter notion of visible direct cause. Thus, this axiom explicitly forbids cycles of visible direct causes, but might be insufficient by itself to forbid cycles in the causal graph. Axioms of this form, using different notions of causal dependence in place of $X \rightsquigarrow Y$, are common in the literature; e.g., Barbero & al.^[18] uses the deterministic notion of direct cause, while Halpern^[12] uses a weaker dependence called "causally affecting".

From the definition of canonical model, we immediately obtain the following facts.

Proposition 8.1. Let Γ be a maximally consistent set of formulas. Then:

1. X is a parent of Y in T^Γ iff X is a visible direct cause of Y in T^Γ .
2. T^Γ is strictly recursive iff it is visibly recursive.

Lemma 8.2. Suppose $\Gamma \supseteq A \cup \{R\}$ (or $\Gamma \supseteq B \cup \{R\}$) is a maximally consistent set of $\mathcal{H}_\sigma$ (resp. $\mathcal{H}_\sigma^+$) formulas. Then, T^Γ is strictly recursive and visibly recursive.

Proof. By Proposition 8.1, it suffices to prove that T^Γ is visibly recursive. Suppose the visible causal graph of T^Γ has a cycle $X_1, \dots, X_n$. Then, by the definition of T^Γ , for all $i = 1, \dots, n$, $X_i \rightsquigarrow X_{i+1} \in \Gamma$ and $X_n \rightsquigarrow X_1 \in \Gamma$. But, since Γ contains all the

instances of axiom R, we also obtain $\sim X_n \rightsquigarrow X_1 \in \Gamma$. This contradicts the consistency of Γ . $\square$

Fix a signature σ . We write $\Gamma \models^{\text{SR}} \varphi$ (resp. $\Gamma \models^{\text{VR}} \varphi$) if every strictly recursive (resp. visibly recursive) model of signature σ that satisfies Γ also satisfies φ . We write $\Gamma \models_1^{\text{SR}} \varphi$ (resp. $\Gamma \models_1^{\text{VR}} \varphi$) if every strictly recursive (or equivalently, visibly recursive) *singleton* model of signature σ that satisfies Γ also satisfies φ .

Theorem 8.3 (Strong completeness for strictly/visibly recursive models)

1. For $\Gamma \cup \{\varphi\} \subseteq \mathcal{H}_\sigma$, $\Gamma \models_1^{\text{SR}} \varphi \Leftrightarrow \Gamma \models_1^{\text{VR}} \varphi \Leftrightarrow \Gamma, R \vdash_A \varphi$.

2. For $\Gamma \cup \{\varphi\} \subseteq \mathcal{H}_\sigma^+$, $\Gamma \models^{\text{SR}} \varphi \Leftrightarrow \Gamma \models^{\text{VR}} \varphi \Leftrightarrow \Gamma, R \vdash_B \varphi$.

Proof. As for Theorem 4.13 (resp. 6.20), using the fact that $\mathbb{T}^\Delta$ is strictly recursive if $A^R \subseteq \Delta$ (Lemma 8.2). $\square$

8.2 Failure of composition

There are further significant differences between the deterministic and indeterministic recursive cases that are not evident from the proposed axiomatization. In the deterministic case, intervening on a single state of affairs again produces a single state of affairs (the unique solution of a certain system of equations). It then follows that counterfactuals $[\mathbf{X} = \mathbf{x}] \psi$ and *might*-counterfactuals $\langle \mathbf{X} = \mathbf{x} \rangle \psi$ are equivalent, and so the latter are redundant. If indeterministic laws are involved, instead, interventions on a single state of affairs may produce multiple possible states of affairs even if there are no cyclic causal laws. The operator $\langle \mathbf{X} = \mathbf{x} \rangle$, then, albeit definable as $\sim[\mathbf{X} = \mathbf{x}] \sim$, seems to be vital for expressing the properties of the solution sets in a natural way.

Another important difference is the failure of one of Galles and Pearl's principles for recursive models, the law of *Composition*. This can be expressed as

$$([\mathbf{X} = \mathbf{x}]W = w \ \& \ [\mathbf{X} = \mathbf{x}]Y = y) \rightarrow [\mathbf{X} = \mathbf{x}, W = w]Y = y,$$

which can sometimes³⁵ be replaced, in axiomatizations for recursive models, by the more intuitive *Conjunction conditionalization*

$$(\mathbf{X} = \mathbf{x} \ \& \ \eta) \rightarrow [\mathbf{X} = \mathbf{x}] \eta,$$

where, crucially, η is a formula without counterfactuals. The example given in section 2.1 shows that both laws fail for indeterministic (even weakly) recursive models. Indeed, in it $A = 1$ and $C = \text{heads}$ hold, but $[A = 1]C = \text{heads}$ does not.

As we have seen, it turns out that the weakened form of *Composition*,

$$\langle \mathbf{X} = \mathbf{x} \rangle (W = w \ \& \ Y = y) \rightarrow \langle \mathbf{X} = \mathbf{x}, W = w \rangle Y = y,$$

which was proposed by Halpern^[12] for the axiomatization of the (possibly) *cyclic* case, is also sound for indeterministic models (both in the recursive and nonrecursive case); it is just axiom I3. This law illustrates the importance of *might*-counterfactuals in the indeterministic context.

8.3 Strong reversibility

Besides *Reversivity* and *Composition*, the third principle proposed by Galles and Pearl^[11] for recursive models is the (strong) *Reversibility* axiom:

$$([\mathbf{X} = \mathbf{x}, W = w]Y = y \ \& \ [\mathbf{X} = \mathbf{x}, Y = y]W = w) \rightarrow [\mathbf{X} = \mathbf{x}]Y = y \text{ (when } Y \neq W).$$

In Galles & Pearl^[11], this axiom was thought of as a characterization of the *unique solution* property³⁶, which (in the deterministic case)

is more general than recursivity. The axiom is indeed part of Halpern's axiomatization of unique solution causal models in Halpern^[12]; it does not feature in the axiomatization of recursive causal models because it is derivable from the other axioms including *Reversivity*. The insights from Zhang & al.^[35] and Fang & Zhang^[22], however, clearly show that this hypothesized connection between *Reversibility* and unicity of solutions is incorrect.

It turns out that strong *Reversibility* is also valid on indeterministic *strictly* recursive models. Since multiplicity of solutions is the norm when the laws are indeterministic, by looking at indeterministic, strictly recursive models, one may observe even more clearly that *Reversibility* has nothing to do with the unicity of solutions.

Theorem 8.4 (Strong reversibility). *Let $T = (T^-, \mathcal{F})$ be a strictly recursive model. Then, $T \models ([\mathbf{X} = \mathbf{x}, W = w]Y = y \ \& \ [\mathbf{X} = \mathbf{x}, Y = y]W = w) \rightarrow [\mathbf{X} = \mathbf{x}]Y = y$.*

Proof. Suppose $T \models [\mathbf{X} = \mathbf{x}, W = w]Y = y$ and $T \models [\mathbf{X} = \mathbf{x}, Y = y]W = w$. Since T is strictly recursive, in the parenthesis graph either W is not an ancestor of Y or Y is not an ancestor of W .

Case 1: W is not an ancestor of Y . Then, by definition of intervention, intervening on W does not affect Y (at most, some assignments may disappear). Thus, from the assumption $T \models [\mathbf{X} = \mathbf{x}, W = w]Y = y$ we immediately obtain $T \models [\mathbf{X} = \mathbf{x}]Y = y$.

Case 2: Y is not an ancestor of W . By the assumption that $T \models [\mathbf{X} = \mathbf{x}, Y = y]W = w$, we get $T_{\mathbf{X}=\mathbf{x}, Y=y} \models W = w$. Since Y is not an ancestor of W , we get $T_{\mathbf{X}=\mathbf{x}} \models W = w$. But then the intervention $do(W = w)$ does not modify the team component of $T_{\mathbf{X}=\mathbf{x}}$, i.e., $T_{\mathbf{X}=\mathbf{x}}^- = T_{\mathbf{X}=\mathbf{x}, W=w}^-$. Now, for each $s \in T_{\mathbf{X}=\mathbf{x}, W=w}^-$ we have $s(Y) = y$ by the assumption that $T \models [\mathbf{X} = \mathbf{x}, W = w]Y = y$; so, $T \models [\mathbf{X} = \mathbf{x}]Y = y$. $\square$

9 Further axiomatizations

9.1 The deterministic case

We now consider the issues of definability and axiomatizability for the subclass of deterministic models. Remember that $T = (T^-, \mathcal{F})$ is deterministic if, for each internal variable Y , $\mathcal{F}_Y$ assigns at most one value to each tuple of values for the parents of Y ($|\mathcal{F}_Y(pa)| \leq 1$ for each such tuple pa). We characterize this property by the following axiom:

$$\text{Det. } \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y' \text{ (when } y \neq y' \text{ and } Y \in \mathcal{I})$$

We remark that this axiom scheme (like axiom schemes I8, J10, J11, J16, and J17) depends, for its formulation, on having the internal/external distinction in the signature; the formulas in this scheme are *not* sound in deterministic models when Y is an external variable. We will show in section 10.1 that this dependency is ineliminable.

Lemma 9.1. *Let T be a model. Then, $T \models \text{Det}$ iff T is deterministic.*

Proof. $\Rightarrow$) Suppose, for the sake of contradiction, that T is not deterministic. Then, there is a variable $Y \in \text{Int}(T)$, a value $pa \in \text{Ran}(PA_Y)$, and two values $y \neq y' \in \text{Ran}(Y)$ such that $y, y' \in \mathcal{F}_Y(pa)$. Pick a $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$ such that $\mathbf{w}_{\uparrow PA_Y} = pa$. By the definition of intervention, $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$ and $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$. From the latter, we obtain $T \models \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$. But then, T falsifies an instance of Det.

$\Leftarrow$) Suppose T is deterministic, and assume that $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$ for some $Y, y, \mathbf{w}$, with $Y \in \text{Int}(T)$.

³⁵ E.g., in the systems considered in Barbero & Yang^[19], the two are provably equivalent. Nested conditionals seem to be needed in the equivalence proof.

³⁶ A structural equation model has this property if 1) its system of equations is satisfied by a unique assignment once values for the exogenous variables are fixed, and 2) the same holds after any intervention.

Since $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$ {and Y is internal}, then, by the semantic clause and the assumption of determinism, $T \not\models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$ when $y' \neq y$. Thus, $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$ whenever $y \neq y'$ and Y is internal, i.e., $T \models \text{Det}$. $\square$

Lemma 9.2. *Let $\Gamma \supseteq A \cup \{\text{Det}\}$ (resp. $\Gamma \supseteq B \cup \{\text{Det}\}$) be a maximal consistent set of $\mathcal{H}_\sigma$ (resp. $\mathcal{H}_\sigma^+$) formulas. Then, $\mathbb{T}^\Gamma$ is deterministic.*

Proof. Suppose for the sake of contradiction that there is an internal variable Y , values $y \neq y' \in \text{Ran}(Y)$, and a tuple $pa \in \text{Ran}(PA_Y)$ such that $\{y, y'\} \subseteq \mathcal{F}_Y^\Gamma(pa)$. Then, by the definition of canonical model, there is a $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$ such that $\mathbf{w}_{\uparrow PA_Y} = pa$ and $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \in \Gamma$, $\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y' \in \Gamma$. However, $\{\text{Det}\} \subseteq \Gamma$, so in particular $\sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y' \in \Gamma$. Thus Γ is inconsistent, contradicting the assumption. $\square$

By the usual argument, then, we obtain the following completeness result. Write $\models^D$ for semantic entailment in the class of deterministic models, and $\models_1^D$ for entailment in the class deterministic singleton models.

Theorem 9.3. $\Gamma \models_1^D \varphi$ iff $\Gamma, \text{Det} \vdash_A \varphi$, and $\Gamma \models^D \varphi$ iff $\Gamma, \text{Det} \vdash_B \varphi$.

9.2 The total and recursive cases

We consider here the axiomatization of total causal teams (of a given fixed signature). Remember that a causal team $T = (T^-, \mathcal{F})$ is total if, for each of its internal variables Y , $\mathcal{F}_Y$ is a total multivalued function; i.e, for each tuple $pa \in PA_Y$, $\mathcal{F}_Y(pa) \neq \emptyset$. We claim that a complete axiomatization for the class of total models is obtained by adding to B the following axiom scheme:³⁷

$\text{Tot. } \langle \mathbf{W}_Y = \mathbf{w} \rangle \top$ (when $Y \in I$),

where $\mathbf{W}_Y$ denotes a list of all variables in the domain, except Y . In the following two lemmas, we will see that axiom Tot captures the notion of totality from a semantic point of view.

Proposition 9.4. *Let T be a model. $T \models \text{Tot}$ iff T is total.*

Proof. $\Leftarrow$) Let $T = (T^-, \mathcal{F})$ be total. Since Y is internal, then, we may pick a $y \in \mathcal{F}_Y(\mathbf{w}_{\uparrow PA_Y})$. We note then that the assignment $s(\mathbf{W}_Y) = \mathbf{w}$, $s(Y) = y$ is in $T_{\mathbf{W}_Y = \mathbf{w}}^-$. Thus, $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle \top$.

$\Rightarrow$) Suppose T satisfies all instances of Tot. Let $Y \in \text{Int}(T)$, $pa \in \text{Ran}(PA_Y)$, and $\mathbf{w} \in \text{Ran}(\mathbf{W}_Y)$ such that $\mathbf{w}_{\uparrow PA_Y} = pa$. Then, $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle \top$. Thus, there is an $s \in T_{\mathbf{W}_Y = \mathbf{w}}^-$, and $s(PA_Y) = pa$. Then, by the definition of intervention, $(pa, s(Y)) \in \mathcal{F}_Y$. Since the choice of $pa \in \text{Ran}(PA_Y)$ was arbitrary, $\mathcal{F}_Y$ is a total function. Since the choice of Y among internal variables was arbitrary, T is total. $\square$

Lemma 9.5. *Let Γ be a maximal consistent set of $\mathcal{H}$ (resp. $\mathcal{H}^+$) formulas containing all instances of A (resp. B) and Tot. Then, $\mathbb{T}^\Gamma$ is total.*

Proof. In the A case, since axiom Tot is an $\mathcal{H}$ formula, by the truth lemma 4.12, $\mathbb{T}^\Gamma \models \text{Tot}$. Thus, $\mathbb{T}^\Gamma$ is total by Proposition 9.4.

In the B case, we use Lemmas 6.13 and 6.19 instead of Lemma 4.12. $\square$

Let us fix a signature σ . We write $\Gamma \models^T \varphi$ if, for all total causal teams T of signature σ , $T \models \Gamma$ implies $T \models \varphi$. Similarly, we write $\Gamma \models_1^T \varphi$ if this holds for all singleton total causal teams of signature σ .

Theorem 9.6.

1. $\Gamma \models_1^T \varphi$ iff $\Gamma, \text{Tot} \vdash_A \varphi$.
2. $\Gamma \models^T \varphi$ iff $\Gamma, \text{Tot} \vdash_B \varphi$.

³⁷ This axiom is analogous to one of the conjuncts of axiom D9 in Halpern^[12].

³⁸ The reader should refer to Halpern^[13] for the clearest description of the semantics.

³⁹ This issue does not arise in Halpern's later book^[13], where formulas are evaluated on assignment over the set of all variables.

Proof. We prove the left-to-right directions. Suppose $\Gamma, \text{Tot} \not\models \varphi$. Then, by the usual argument, $\Gamma \cup \text{Tot} \cup \{\sim \varphi\} \cup A$ (resp. $\Gamma \cup \text{Tot} \cup \{\sim \varphi\} \cup B$) is consistent, and by the Lindenbaum lemma, it is included in a maximally consistent set Δ . By Lemma 9.5, $\mathbb{T}^\Delta$ is total. Furthermore, $\mathbb{T}^\Delta \models \Delta$ (by the truth lemma 4.12, in the A case; by Lemmas 6.13 and 6.19 in the B case). In particular, $\mathbb{T}^\Delta \models \Gamma$ but $\mathbb{T}^\Delta \not\models \varphi$, and thus $\Gamma \not\models_{(1)}^T \varphi$. $\square$

Moreover, we can see that the class of total recursive models can be axiomatized by further adding axiom R (it turns out that, within the class of total models, the logic of recursivity coincides with that of strict/visible recursivity).

Lemma 9.7. *Let $\Gamma \supseteq A \cup \{\text{Tot}, R\}$ be a maximally consistent set of $\mathcal{H}_\sigma$ formulas. Then, $\mathbb{T}^\Gamma$ is total and recursive.*

Proof. Totality is given by Lemma 9.5. Let us prove recursivity. Suppose for the sake of contradiction that there is a cycle $X_1, \dots, X_n$ in the causal graph of $\mathbb{T}^\Gamma$. Since $\mathbb{T}^\Gamma$ is total, then, we may apply Theorem 7.4 and conclude that $X_1, \dots, X_n$ is also a cycle in the parenthood graph. By Proposition 8.1, it is a cycle in the visible causal graph. The proof then proceeds as for Lemma 8.2. $\square$

Then, the usual line of reasoning tells us that $B \cup \{\text{Tot}, R\}$ is a sound and complete axiomatization for the class of total recursive models, while $A \cup \{\text{Tot}, R\}$ is a sound and complete axiomatization for the class of total recursive singleton models.

Finally, a moment of thought shows that the singleton causal teams that are total, deterministic, and recursive essentially describe the same semantics as the recursive causal models of Galles & Pearl^[11] and Halpern^[12–13].³⁸ Indeed, first of all, while Halpern's causal settings only include an assignment u over the *exogenous* variables, the recursivity and determinism constraints make so that u determines a unique assignment s over all variables that complies with the causal laws. Second, even though our definition of intervention in the general case differs from that of Halpern, it was shown in Barbero & Galliani^[30] that in the (deterministic and) recursive case the two notions of interventions coincide. On the side of the of syntax, we note that our axioms are sound in Halpern's models if we follow him in interpreting the formulas ψ without modal operators as abbreviations for $\Box \psi$. We can thus conclude that:

- The axiom systems $A + \text{Rec} + \text{Tot} + \text{Det}$ and $B + \text{Rec} + \text{Tot} + \text{Det} + \text{Sing}$ are alternative axiomatizations for the class of Halpern's recursive causal models.

One might be tempted to surmise that the axiom systems $A + \text{Tot} + \text{Det}$ and $B + \text{Tot} + \text{Det} + \text{Sing}$ are then axiomatizations of Halpern's general causal models, but that is incorrect. One reason, already hinted above, is that without the restriction of recursivity, our notion of intervention produces distinct results from both the ref.^[12] and ref.^[13] versions of Halpern's definition. For example, Halpern's empty interventions can produce new assignments, so that axiom I7 ($\mathbf{Y} = \mathbf{y} \leftrightarrow \Box \mathbf{Y} = \mathbf{y}$) can be falsified. Note furthermore that, in our framework, if we apply the same intervention to two assignments s, s' that agree on the values of the external variables, we might still get two distinct sets of post-intervention assignments, in case s and s' disagree on the values of some variables that are not descendants of the intervened variables. Thus, it is not even clear, in the nonrecursive case, to what singleton causal team should a causal setting (M, u) of Halpern's^[12] semantics correspond to.³⁹ Most likely, such a causal setting should correspond to the *maximal* causal team containing *all* the assignments that assign u to the exogenous variables and respect the constraint of compatibility with the causal laws (this shows that,

implicitly, causal team semantics was already used in Halpern^[12]). Interestingly, it can be shown that this class of model satisfies axiom I6 (*Weak reversibility*), and thus plausibly it can be axiomatized by some appropriate extension of system A \ {I10}. We will not pursue this line of thought any further in this paper, as in any case the language from Halpern^[12] requires further complications in the semantics.^④

10 Simplified signatures

We have been working with signatures that explicitly encode the distinction between internal and external variables. An alternative, closer to the tradition in the philosophical literature, is to take signatures to be pairs of the form $s = (\text{Dom}, \text{Ran})$ (where Dom and Ran are as before); we will call these **simplified signatures**. Models of simplified signature s and interventions on them can be redefined with minimal changes to our earlier definitions; each model T has its own set $\text{Int}(T)$ of internal variables, which can be an arbitrary subset of the variable domain given by the signature. The result is a bit more elegant, as intervening on a model of simplified signature s produces again a model of simplified signature s . In the same vein, we can define languages $\mathcal{H}_s$. A number of completeness results have been proved in the literature^[14,19,21] relative to simplified signatures; such results are more general than those concerning the other type of signatures, as they cover larger classes of models. The main reason why we do not adopt simplified signatures in our official framework is that the internal/external distinction is essential in order to formulate axiom I8 of our axiomatizations (and, for system B, axioms J10, J11, J16, and J17); we will see that the internal/external distinction is *undefinable* in $\mathcal{H}_s$. It may be, of course, that a different, still reasonable axiomatization exists for classes of models with simplified signature, but this needs remain an open problem for now. A second reason for favouring the ordinary signatures was the discovery that the deterministic class of models is undefinable if one uses simplified signatures. In this section, we first explore these kinds of undefinability results. After that, we see three strategies for recovering the definability of determinism while keeping simplified signatures: by restricting attention to singleton models, by extending the language $\mathcal{H}$ with right-nested counterfactuals, or by extending it with a learning operator.

10.1 Undefinability results

In this section, we see some limitations in the expressivity of language $\mathcal{H}$ in the context of simplified signatures; a few of them extend to the language with full signatures. We first identify a crucial closure condition for the classes of models definable in $\mathcal{H}_s$. The following definition describes a type of internal variables that may be particularly difficult to distinguish from external variables.

Definition 10.1. Let $T = (T^-, \mathcal{F})$ be a model and $V \in \text{Int}(T)$. We say that:

1. V is **quasi-external** for T if $\mathcal{F}_V(pa) = \mathcal{F}_V(pa')$ for all $pa, pa' \in \text{Ran}(PA_V)$.
2. T is **generic** for V if, for each value $v \in \bigcup_{pa \in \text{Ran}(PA_V)} \mathcal{F}_V(pa)$, there is an $s \in T^-$ such that $s(V) = v$.

If Γ is a set of formulas of (simplified) signature s , we denote as $\mathcal{K}_T^s$ the set of models of signature s which satisfy all formulas of Γ . We are now ready to formulate the closure condition.

Theorem 10.1. Let $T = (T^-, \mathcal{F})$ be a model of signature s and $V \in \text{Int}(T)$, and suppose further that V is quasi-external for T and T is generic for V . Consider the model $T^* = (T^-, \mathcal{G})$, where $\mathcal{G} = \mathcal{F}_{\text{Int}(T) \setminus \{V\}}$ (i.e., the only difference between T and T^* is that V is internal in T and external in T^*). For all $\Gamma \subseteq \mathcal{H}_s$,

$$T \in \mathcal{K}_T^s \iff T^* \in \mathcal{K}_{T^*}^s.$$

For the proof, it may be useful to review here some earlier notational conventions. Remember that, if G is the parenthood graph of a model T , we write G_X for the graph obtained from G by omitting all arrows that point to variables in X ; we write G_V for $G_{\{V\}}$. If G is the parenthood graph of T , we will write N_X^G for the set of nondescendants of X in G , and PA_V^G for the set of parents of V in G . And remember that, if T^- is a team, $T^-(X)$ denotes the set $\{s(X) \mid s \in T^-\}$.

Proof. Let $\varphi \in \Gamma$. It suffices to prove that $T \models \varphi$ iff $T^* \models \varphi$; we proceed by induction on φ .

For the base case, we just observe that, since T and T^* have the same team component, $T \models Y = y$ iff $T^* \models Y = y$. The inductive steps for φ of the form $\psi \& \chi$ or $\sim \psi$ are straightforward. In case φ is $\Box \psi$, we observe that, since ψ contains no modal operators, we obviously have $(\{s\}, \mathcal{F}) \models \psi$ iff $(\{s\}, \mathcal{G}) \models \psi$ for each $s \in T^-$; thus, $T \models \Box \psi$ iff $T^* \models \Box \psi$.

Lastly, suppose φ is equal to $[X = x]\psi$, with $X \neq \emptyset$. Since ψ has no occurrences of modal operators, the desired conclusion follows immediately if we prove that $(T_{X=x})^- = (T_{X=x}^*)^-$. This is obvious if $V \in X$. Let us then assume $V \notin X$.

⊆ Let $t \in (T_{X=x})^-$. From this it follows that: 1) $t(X) = x$, 2) t is compatible with $\mathcal{F}_{X=x}$, and 3) $t(N_X^G) \in T^-(N_X^G)$ (where G is the parenthood graph of T). From 2), it follows a fortiori that t is compatible with $\mathcal{G}_{X=x}$. It remains then to check that $t(V) \in T^-(V)$ (as $N_X^{G_V} = N_X^G \cup \{V\}$). If $V \in N_X^G$, this is immediate. Suppose then that $V \notin N_X^G$. Note that, since $t \in (T_{X=x})^-$ and $V \notin X$, $t(V) \in \mathcal{F}_V(t(PA_V))$; thus, since T is generic for V , there is an $s \in T^-$ such that $s(V) = t(V)$. Thus, $t(V) \in T^-(V)$, and we may conclude that $t \in (T_{X=x}^*)^-$.

⊇ Let $t \in (T_{X=x}^*)^-$. Then, we have that 1) $t(X) = x$, 2) t is compatible with $\mathcal{G}_{X=x}$, and 3) $t(N_X^{G_V}) \in T^-(N_X^{G_V})$. Since $N_X^G \subseteq N_X^{G_V}$, from 3) we obtain $t(N_X^G) \in T^-(N_X^G)$. We only need to show, then, that t is compatible with $(\mathcal{F}_{X=x})_V = \mathcal{F}_V$, i.e., $t(V) \in \mathcal{F}_V(t(PA_V^G))$. For this purpose, note that, since V is external in T^* , by definition of intervention, $t(V) \in T^-(V)$. Thus, there is an $s \in T^-$ such that $s(V) = t(V)$. By the compatibility constraints of T , $t(V) \in \mathcal{F}_V(s(PA_V^G)) = \mathcal{F}_V(t(PA_V^G))$, where the equality holds because V is quasi-external in T . Thus, t is compatible with $\mathcal{F}_{X=x}$, and we conclude that $t \in (T_{X=x})^-$. □

Corollary 10.2. Suppose the simplified signature s has at least two variables, each of them with at least two distinct possible values.^④ Then:

1. The set of deterministic models of signature s is not definable in $\mathcal{H}_s$ (not even using an infinite set of formulas). Neither is its complement.
2. The same holds for its intersections with the classes of total models, recursive models, nonempty models, or any of their intersections.

Proof. Let $s = (\text{Dom}, \text{Ran})$, where $\text{Dom} = \{B, C\}$, $\text{Ran}(B) = \text{Ran}(C) = \{1, 2\}$, and let Γ be an arbitrary set of $\mathcal{H}_s$ formulas. We consider a small modification of the model given in Example 2.2. Let $T = (T^-, \mathcal{F})$, where $T^- = \{(B, 1), (C, 1)\}, \{(B, 1), (C, 2)\}$, $\text{Int}(T) = \{C\}$,

^④ We are referring to the fact that the formulas used in Halpern^[12] may involve multiple contexts—e.g., the formula $Y(u) = y \& Y(u') = y'$ involves two different assignments u, u' over the exogenous variables. Such formulas require a *family* of maximal causal teams for their evaluation.

^④ The same argument can also be easily adapted to the case that one of the two variables is single-valued, but we prefer to illustrate it in a less pathological context.

$PA_C = \{B\}$, and $\mathcal{F}_C(1) = \mathcal{F}_C(2) = \{1, 2\}$. This last condition says that C is quasi-external in T ; the fact that $T^-(C) = \mathcal{F}_C(1) = \bigcup_{b=1,2} \mathcal{F}_C(b)$ says that T is generic for C . Thus, by Theorem 10.1, $T \models \Gamma$ iff $T^* = (T^-, \mathcal{F}_{\text{Int}(T) \setminus \{C\}}) = (T^-, \emptyset) \models \Gamma$. But T^* is deterministic, while T is not. Thus, $\mathcal{K}_\Gamma^s$ is not the set of deterministic models of signature s (nor the set of nondeterministic ones). This proves point 1. To prove point 2, it suffices to note that both T and T^* are nonempty, total, and (strictly) recursive models. $\square$

The undefinability of the deterministic class of models does not necessarily entail the impossibility of a reasonable axiomatization—it might simply be that this class shares the same set of validities with a larger, definable class of models, and that such a class admits a recursive axiomatization. We leave it as an open problem whether such a class exists and whether it has a defining set of formulas that is canonical for the deterministic class.

A possible reaction to this kind of result is to think that the class of deterministic models, as described here, is an unnatural one. After all, if we allow the external variables to take multiple values, aren't these models just forbidding indeterminism for the internal variables while allowing it for the external ones? If we forbid indeterminism at both levels, we obtain the singleton deterministic models; so, this should be the correct notion of a deterministic model. We find this attitude to be untenable, for a multitude of reasons. First, we find it doubtful that the multiplicity of values for an external variable U in a team may always be interpreted as the existence of an indeterministic causal law producing U in the same sense as indeterminacy for the causal laws; it may simply encode epistemic uncertainty. Thus, we take it that the teams may encode forms of uncertainty that are not reducible to indeterminism in laws. Second, even the 1-assignment models exhibit forms of uncertainty that are not due to indeterminism in the laws—the uncertainty that is triggered by the multiplicity of solutions in cyclic models. Should we then also forbid multiple-solution models, in order to properly characterize determinism? But the final blow for this attitude comes from the fact that the class of deterministic models (in our sense) is definable when we use the ordinary signatures (see section 9.1). Furthermore, even with simplified signatures, determinism becomes definable within very natural extensions of language $\mathcal{H}_s$, as we shall see in a moment. Thus, we take it that the problem lies not in the definition of the deterministic class, but rather in the limited expressivity of language $\mathcal{H}_s$.

As a second application of the closure condition outlined in Theorem 10.1, we show that the distinction between internal and external variables cannot be captured in $\mathcal{H}_s$ (nor can the notion of strict recursivity). If we had a formula $\varphi_{\text{Ext}(Y)}$ characterizing the fact that Y is external, then we could reformulate axiom I8 in the form $\varphi_{\text{Ext}(Y)} \rightarrow (\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \leftrightarrow \diamond Y = y)$ and obtain a complete system; the undefinability result shows that this simple trick is unavailable. A quick argument to show this undefinability consists in observing that, if we had formulas $\varphi_{\text{Ext}(Y)}$, then we would be able to characterize the deterministic models of signature s by the formula $\varphi_{\text{Ext}(Y)} \rightarrow \text{Det}$ (cp. section 9.1), contrary to Corollary 10.2. Our alternative argument will also show that the notion of direct cause and related concepts such as exogeneity/endogeneity and recursivity are undefinable in $\mathcal{H}$; the same goes for the notion of a dummy argument.

Corollary 10.3. *Let $s = (\text{Dom}, \text{Ran})$ be a simplified signature.*

1. *The property of being internal (resp. external) is not definable in $\mathcal{H}_s$, i.e., for each Y and for all models of signature s , there is no $\Gamma \subseteq \mathcal{H}_s$ such that $T \models \Gamma \iff Y$ is internal (resp. external) in T .*

2. *The notion of direct cause is not definable in $\mathcal{H}_s$, i.e., there is no $\Gamma \subseteq \mathcal{H}_s$ such that, for all models of signature s , $T \models \Gamma \iff X$ is a direct cause of Y in T . The same for not being a direct cause.*

3. *The notion of recursivity is not definable in $\mathcal{H}_s$, i.e., there is no $\Gamma \subseteq \mathcal{H}_s$ such that, for all models of signature s , $T \models \Gamma \iff T$ is recursive. The same for its negation.*

4. *The notion of endogeneity is not definable in $\mathcal{H}_s$, i.e., there is no $\Gamma \subseteq \mathcal{H}_s$ such that, for all models of signature s , $T \models \Gamma \iff X$ is endogenous in T . The same for exogeneity.*

5. *The notion of being a dummy argument is not definable in $\mathcal{H}_s$, i.e., for all variables $X, Y \in \text{Dom}$, there is no $\Gamma \subseteq \mathcal{H}_s$ such that, for all models of signature s , $T \models \Gamma \iff X$ is a dummy argument for Y .*

Proof. Consider the simplified signature $s = (\text{Dom}, \text{Ran})$ such that $\text{Dom} = \{X, Y\}$, $\text{Ran}(X) = \text{Ran}(Y) = \{1, 2\}$. Let T^- be the full team for this signature:

X	Y
1	1
1	2
2	1
2	2

2) This team is compatible with the following causal laws: $PA_X = \{Y\}$, $PA_Y = \{X\}$, $\mathcal{F}_X(1) = \mathcal{F}_X(2) = \{1, 2\}$, and $\mathcal{F}_Y(1) = \mathcal{F}_Y(2) = \{1, 2\}$. Thus, in particular, X is a direct cause of Y in $T = (T^-, \mathcal{F})$ (just look at the effect of any intervention on a single assignment). Instead, in $T^* = (T^-, \mathcal{F}_{\{X\}})$, $X \notin PA_Y$, so, by Theorem 7.4 (whose proof is unchanged if we use simplified signatures), X is not a direct cause of Y in T^* . On the other hand, Y is quasi-external in T , and T is generic for Y ; thus, by Theorem 10.1, T and T^* satisfy the same formulas.

1) In the example above, Y is internal in T but external in T^* .

But, by Theorem 10.1, T and T^* satisfy the same $\mathcal{H}$ formulas.

3) Just observe that T^* is recursive while T is not.

4) Observe that Y is endogenous in T but exogenous in T^* .

5) Observe that X is a dummy argument for Y in T but not in T^* (as, in T^* , Y is external). $\square$

We remark that the notions of direct cause and those defined in terms of it do not really really depend on having the distinction between external and internal variables in the signature. Thus, an $\mathcal{H}_\sigma$ formula defining, say, " X is a direct cause of Y " (where $\sigma = (\mathcal{E}, \mathcal{I}, \text{Ran})$) does so also over models of signature s (where $s = (\mathcal{E} \cup \mathcal{I}, \text{Ran})$), so it would be an (impossible) $\mathcal{H}_s$ definition of " X is a direct cause of Y ". Thus, some of the definability results also extend to full signatures.

Corollary 10.4. *The notions of direct cause, recursivity, endogeneity and their negations are not definable in $\mathcal{H}_\sigma$.*

The previous shower of undefinability results suggests that, if we use simplified signatures, we should either modify the notion of model or consider a stronger language. We now turn to these two possibilities.

10.2 Capturing determinism in the singleton case

We have seen that the class of deterministic models is undefinable if we adopt the simplified signatures. We show here that, however, determinism can be characterized if we restrict attention to the class of singleton models. Consider the following axiom:

$\text{Det}_s. \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$.

This is just axiom Det, without Y being restricted to internal variables (we cannot make this distinction with simplified signatures). As already observed, such an axiom is not sound over the class of deterministic models (an external Y may violate it); but it is sound for singleton deterministic models and characterizes them.

Lemma 10.5. *Let $T = (\{s\}, \mathcal{F})$ be a singleton relational causal team with simplified signature. Then, $T \models \text{Det}_s$ iff T is deterministic.*

Proof. $\Rightarrow$) This direction of the proof is identical to the left-to-right implication in the proof of Lemma 9.1.

$\Leftarrow$) Suppose T is a deterministic singleton model, and assume that $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$ for some $Y, y, \mathbf{w}$. In case Y is not in $\text{Int}(T)$, and thus Y is not a descendant of any other variable, by definition of intervention, and the fact that T is a singleton model, $T^-(Y) = T_{\mathbf{W}_Y = \mathbf{w}}^-(Y) = \{y\}$. Thus, $T \not\models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$ if $y' \neq y$, from which we immediately obtain $T \models \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$.

If instead $Y \in \text{Int}(T)$, the proof proceeds as in Lemma 9.1. $\square$

We also note that generic⁴² singleton models do not exist except for trivial signatures (which allow single-valued variables). Thus, for singleton models, we cannot prove undefinability results via Theorem 10.1. Nevertheless, we are not aware of sets of formulas that may be used to define the relativizations to singleton models of notions like being internal/external, being a direct cause, and so on.

10.3 Capturing determinism via nested counterfactuals

A natural way of extending language $\mathcal{H}$ is to allow for right-nested counterfactuals—e.g., formulas of the form $[\mathbf{X} = \mathbf{x}][\mathbf{Y} = \mathbf{y}]\psi$. We have already introduced such languages; when discussing a simplified signature s , the corresponding language with right-nested counterfactuals will be denoted as $\mathcal{H}_s^+$. We immediately see that $\mathcal{H}_s^+$ is more expressive than $\mathcal{H}_s$, precisely because it captures the deterministic class.

Theorem 10.6. *Let T be a model with simplified signature. Then, T is deterministic iff it satisfies the following axiom scheme:*

$$\text{Det}^\square. \quad \square(\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y') \& \\ (\sim \Diamond T \rightarrow (\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y')) \quad (\text{when } y \neq y').$$

Proof. $\Rightarrow$) Suppose $T = (T^-, \mathcal{F})$ is deterministic. Then, for all $s \in T^-$, $(\{s\}, \mathcal{F})$ is deterministic. Thus, by Lemma 10.5, $(\{s\}, \mathcal{F}) \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$ (for any choice of $Y, \mathbf{w}, y \neq y'$). Since this holds for all $s \in T^-$, $T \models \square(\langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y')$.

Now suppose $T \models \sim \Diamond T$. Then, T is empty. If Y is external, then, by definition of intervention, $T \not\models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$, thus $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$. If Y is internal, then, since T is deterministic, $T_{\mathbf{W}_Y = \mathbf{w}}$ has at most one assignment. If it is empty, then $T \not\models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$, and we conclude as before. Suppose instead $T_{\mathbf{W}_Y = \mathbf{w}}^- = \{s\}$, with $s(Y) = y^*$. If $y^* \neq y$, then $T \not\models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y$ and we conclude as before. In case $y^* = y$, then $T \not\models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$, so $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$.

$\Leftarrow$) Suppose $T \models \text{Det}^\square$. Then, for any choice of $Y, \mathbf{w}, y \neq y'$, and any $s \in T^-$, we have $(\{s\}, \mathcal{F}) \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$. Thus, by Lemma 10.5, $(\{s\}, \mathcal{F})$ is deterministic. Suppose T is nonempty, i.e., there is such an s ; then, since T and $(\{s\}, \mathcal{F})$ have the same causal laws, also T is deterministic. In case T is empty, by the second conjunct of axiom Det , we obtain $T \models \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w} \rangle Y = y'$ (for each pair of distinct y, y'). Thus, if Y is internal in T , by definition of intervention, $|\mathcal{F}_Y(\mathbf{w}_{\uparrow PA_Y})| \leq 1$. $\square$

Thus, the axiom system $\text{B} + \text{Det}^\square$ is sound and complete for $\mathcal{H}^+$ over the class of deterministic models.

10.4 Capturing determinism via learning

A second natural extension of language $\mathcal{H}$ is obtained by adding to it a second conditional operator $\triangleright$; the intended interpretation of $\psi \triangleright \chi$ is that after *learning* ψ (e.g., after observing a feature of the world or

the outcome of an experiment) one can conclude that χ holds. This operator is introduced in Barbero & Sandu^[28] under the name of *selective implication* and is a sort of qualitative counterpart of probabilistic conditionalization; see Barbero & Sandu^[37].

$$\text{Language } \mathcal{HO}_s: \quad X = x \mid \sim \psi \mid \psi \& \chi \mid \alpha \triangleright \chi \mid [\mathbf{X} = \mathbf{x}] \eta$$

where η has no occurrences of modal operators, and α is restricted to formulas whose occurrences of $\sim$ (if any) are exclusively inside the scope of modal operators. We will say that such an α *has the $\sim$ -restriction*. The semantic clause for the additional operator is:

$$\bullet \quad T \models \alpha \triangleright \chi \text{ iff } T^\alpha \models \chi, \text{ where } T^\alpha = \{s \text{ of signature } s \mid (\{s\}, \mathcal{F}) \models \alpha\}.$$

The restriction on α guarantees that α is *flat*, i.e., it holds of a causal team iff it holds of all its singleton subteams (this will be proved below), and, consequently, that the conditional can really be interpreted as saying that "upon learning α , χ holds". Without this restriction, unintended meanings can easily arise. For example, this operator should satisfy the classical principle $A \triangleright A$. But using $\sim$ we can define $\sqcup$, and a formula of the form $(X = 0 \sqcup X = 1) \triangleright (X = 0 \sqcup X = 1)$ is not satisfied by teams which feature both values 0 and 1 for X .

We may also write $\alpha \triangleright \beta$ as an abbreviation for $(\alpha \triangleright \beta) \& (\beta \triangleright \alpha)$. Since both α and β need to be flat, $T = (T^-, \mathcal{F}) \models \alpha \triangleright \beta$ means that, for all $s \in T^-$, $(\{s\}, \mathcal{F}) \models \alpha$ iff $(\{s\}, \mathcal{F}) \models \beta$; or, equivalently, that $T^\alpha = T^\beta$.

Also $\mathcal{HO}$ captures determinacy. Consider the following axiom scheme:

$$\text{Det}^O. \quad \mathbf{W} = \mathbf{w} \triangleright [\langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y'] \quad (\text{when } y \neq y')$$

Theorem 10.7. *Let T be a model with simplified signature. $T \models \text{Det}^O$ iff T is deterministic.*

Proof. $\Rightarrow$) Suppose $T \models \text{Det}^O$. We need to prove that, if Y is internal, then for all $pa \in \text{Ran}(PA_Y)$, $\mathcal{F}_Y(pa)$ is contained in a singleton set. Pick an arbitrary $\mathbf{w} \in \text{Ran}(\mathbf{W})$ and let $\mathbf{w}' \in \text{Ran}(\mathbf{W}_Y)$ be such that $\mathbf{w}'_{\uparrow PA_Y} = pa$. Since $T \models \text{Det}^O$, we have $T^{\mathbf{W}=\mathbf{w}} \models \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y'$ whenever $y \neq y'$; thus, $(T^{\mathbf{W}=\mathbf{w}})_{\mathbf{W}_Y = \mathbf{w}'}$ contains at most one assignment. If it is empty, then, by definition of intervention, $\mathcal{F}_Y(pa) = \emptyset$, and if it has an assignment s , then $\mathcal{F}_Y(pa) = \{s(Y)\}$, as needed.

$\Leftarrow$) Suppose $T = (T^-, \mathcal{F})$ is deterministic. Suppose first that Y is internal, and assume that $T^{\mathbf{W}=\mathbf{w}} \models \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y$. Let $y' \in \text{Ran}(Y)$ be such that $y' \neq y$. Since $\mathcal{F}_Y$ is deterministic, $T^{\mathbf{W}=\mathbf{w}} \models \sim \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y'$, as needed.

Suppose instead Y is external. If $T^{\mathbf{W}=\mathbf{w}}$ is a singleton model, then, by the definitions of intervention and determinism also $(T^{\mathbf{W}=\mathbf{w}})_{\mathbf{W}_Y = \mathbf{w}'}$ is a singleton model, say with team $\{t\}$. If $t(Y) \neq y$, then $T^{\mathbf{W}=\mathbf{w}} \models \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y'$ because the antecedent is false. If $t(Y) = y$, then $T^{\mathbf{W}=\mathbf{w}} \models \sim \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y'$, and so again $T^{\mathbf{W}=\mathbf{w}} \models \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y'$.

If instead $T^{\mathbf{W}=\mathbf{w}}$ is empty, since Y is external, by definition of intervention also $(T^{\mathbf{W}=\mathbf{w}})_{\mathbf{W}_Y = \mathbf{w}'}$ is empty; so, it is not the case that $T^{\mathbf{W}=\mathbf{w}} \models \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y$. Thus, we may conclude that $T \models \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y \rightarrow \sim \langle \mathbf{W}_Y = \mathbf{w}' \rangle Y = y'$. $\square$

Again, we lack a complete axiomatization for $\mathcal{HO}$, either on ordinary or on simplified signatures; and we remark that the Henkin method via canonical models seems ill-suited for treating the operator $\triangleright$.⁴³ We think that, in any case, the additional operator is perfectly natural in the context of causal inference; thus, we use the rest of this section to understand language $\mathcal{HO}$ better and suggest a partial axiomatization.

⁴² In the sense of definition 10.1.

⁴³ Actually, an axiomatization via canonical models could be obtained for a probabilistic language featuring $\triangleright$ in Barbero & Virtama^[21], but the proof strategy does not apply to our case.

The following definitions are (analogues of) standard ones from the literature on team semantics; we assume a fixed simplified signature s . $S = (S^-, \mathcal{F})$ is a **causal subteam** of $T = (T^-, \mathcal{G})$ ($S \leq T$) if $S^- \subseteq T^-$ and $\mathcal{F} = \mathcal{G}$.

- φ is **flat** if, for all causal teams $T = (T^-, \mathcal{F})$, $T \models \varphi$ iff for every $t \in T^-$, $(\{t\}, \mathcal{F}) \models \varphi$.
 - φ has the **empty team property** if it is satisfied by all empty causal teams of signature s .
 - φ is **downward closed** if, whenever $T \models \varphi$ and $S \leq T$, then $S \models \varphi$.
 - φ is **upwards closed** if, whenever $S \models \varphi$ and $S \leq T$, then $T \models \varphi$.
- Note in particular that flatness implies downward closure and the empty team property.

Lemma 10.8. 1. If $\alpha \in \mathcal{HO}$ has the $\sim$ -restriction, then it is flat (and therefore is downward closed and has the empty team property).

2. $\mathcal{HO}$ formulas of the form $\langle \mathbf{X} = \mathbf{x} \rangle \psi$ are upwards closed.

3. If ψ, χ are downward closed, then $\psi \sqcup \chi$ is downward closed.

Proof. 1) By induction on α , simultaneously for all models. The cases for $\&$ and $Y = y$ are straightforward.

• Case φ is $\beta \supset \gamma$. Suppose first that $T = (T^-, \mathcal{F}) \models \beta \supset \gamma$. Then, $T^\beta \models \gamma$. Since γ has no occurrence of $\sim$ except in the scope of modal operators, by inductive hypothesis $(\{s\}, \mathcal{F}) \models \gamma$ for each $s \in (T^\beta)^-$. Similarly, by applying the i.h. for β we obtain that $(\{s\}, \mathcal{F}) \models \beta$ for each $s \in (T^\beta)^-$. Thus, if $s \in T^-$ is such that $(\{s\}, \mathcal{F}) \models \beta$, then we also have $(\{s\}, \mathcal{F})^\beta \models \gamma$, and so $(\{s\}, \mathcal{F}) \models \beta \supset \gamma$. If instead $s \in T^-$ is such that $(\{s\}, \mathcal{F}) \not\models \beta$, then $(\{s\}, \mathcal{F})^\beta = (\emptyset, \mathcal{F})$; since γ is flat by the i.h., it has the empty team property, and so $(\{s\}, \mathcal{F})^\beta \models \gamma$. Thus, again, $(\{s\}, \mathcal{F}) \models \beta \supset \gamma$.

Vice versa, suppose $(\{s\}, \mathcal{F}) \models \beta \supset \gamma$ for all $s \in T^-$. Then, for all such s , $(\{s\}, \mathcal{F})^\beta \models \gamma$. By the i.h., $T^\beta \models \gamma$. Thus, $T \models \beta \supset \gamma$.

• Case φ is $[\mathbf{X} = \mathbf{x}] \beta$. Suppose $T \models [\mathbf{X} = \mathbf{x}] \beta$. Then, for all $t \in T_{\mathbf{X}=\mathbf{x}}^-$, $(\{t\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}) \models \beta$. Note that in case T is empty, flatness just amounts to the assumption that $T \models [\mathbf{X} = \mathbf{x}] \beta$; thus, we may henceforth assume that T is nonempty. We then have $T_{\mathbf{X}=\mathbf{x}}^- = \bigcup_{s \in T^-} s_{\mathbf{X}=\mathbf{x}}^\mathcal{F}$, and so, for all $s \in T^-$ and all $t \in s_{\mathbf{X}=\mathbf{x}}^\mathcal{F}$, $(\{t\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}) \models \beta$. Thus, for all $s \in T^-$, $(\{s\}, \mathcal{F}) \models [\mathbf{X} = \mathbf{x}] \beta$. All the steps in this argument are reversible.

2) Let $S = (S^-, \mathcal{F}) \leq T = (T^-, \mathcal{F})$ and $S \models (\mathbf{X} = \mathbf{x}) \psi$. Then, there is a $t \in S_{\mathbf{X}=\mathbf{x}}^-$ such that $(\{t\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}) \models \psi$. Since $S_{\mathbf{X}=\mathbf{x}} \leq T_{\mathbf{X}=\mathbf{x}}$, such t is in $T_{\mathbf{X}=\mathbf{x}}^-$. Thus, $T \models (\mathbf{X} = \mathbf{x}) \psi$.

3) This is a well-known result. $\square$

We remark that $\supset$ does not coincide with the previously defined material implication $\rightarrow$. The subtle relationship between the two operators is described in the following proposition.

Proposition 10.9. 1. $\psi \rightarrow \chi \not\models \psi \supset \chi$, even when ψ, χ are atomic formulas.

2. If α has the $\sim$ -restriction, then $\alpha \supset \chi \models \alpha \rightarrow \chi$. So, this holds for $\mathcal{HO}$ formulas.

3. If α has the $\sim$ -restriction, then $\models \alpha \rightarrow \chi$ iff $\models \alpha \supset \chi$.

Proof. 1) Let T be a causal team of domain $\{X\}$, which features at least two values 0,1 for X . Then, $T \models X = 0 \rightarrow X = 1$ because the antecedent is false, but $T^{X=0} \not\models X = 1$, so $T \not\models X = 0 \supset X = 1$.

2) Suppose $T \models \alpha \supset \chi$ and $T \models \alpha$. The latter (by the flatness of α) gives $T^\alpha = T$, and so the former gives $T \models \chi$. Thus, $T \models \alpha \rightarrow \chi$.

3) The right-to-left direction immediately follows from point 2.

For the left-to-right direction, assume that $\models \alpha \rightarrow \chi$, and let T be a causal team. Since α is flat, $T^\alpha \models \alpha$. But then $T^\alpha \models \chi$. So, $T \models \alpha \supset \chi$. Thus, $\models \alpha \supset \chi$. $\square$

In particular, $\psi \supset \chi$ as a subformula cannot be replaced with $\psi \rightarrow \chi$, or vice versa (we will see that these kinds of substitutions are allowed

within the scope of a modal operator). Point 3 can be taken as a pair of inference rules, which may be applied only to valid formulas.

To see that the $\sim$ -restriction is needed in 2., and in the right-to-left direction of 3., consider a signature with a Boolean variable X ; any causal team of such signature that features both values for X satisfies $\sim X = 0 \supset X = 1$; but such a causal team does not satisfy $\sim X = 0 \rightarrow X = 1$ (the consequent is false). For point 3, a counterexample to the left-to-right implication is given by the formulas $(X = 0 \sqcup X = 1) \rightarrow (X = 0 \sqcup X = 1)$ and $(X = 0 \sqcup X = 1) \supset (X = 0 \sqcup X = 1)$.

The following are a number of formal properties of $\supset$, which might serve as a starting point toward axiomatization.

Proposition 10.10. If $\alpha, \beta, \psi, \chi$ are $\mathcal{HO}$ formulas, and α, β satisfy the $\sim$ -restriction, then:

1. $\alpha \supset (\psi \& \chi) \equiv (\alpha \supset \psi) \& (\alpha \supset \chi)$.
2. $\alpha \supset \sim \chi \equiv \sim (\alpha \supset \chi)$.
3. $\alpha \supset (\beta \supset \psi) \equiv (\alpha \& \beta) \supset \psi$.
4. $\langle \mathbf{X} = \mathbf{x} \rangle \alpha \rightarrow ([\mathbf{X} = \mathbf{x}] (\alpha \supset \beta) \equiv ([\mathbf{X} = \mathbf{x}] \alpha) \supset ([\mathbf{X} = \mathbf{x}] \beta))$.
5. $(\Diamond \top \& \langle \mathbf{X} = \mathbf{x} \rangle \alpha) \rightarrow ([\mathbf{X} = \mathbf{x}] (\alpha \supset \chi) \equiv ([\mathbf{X} = \mathbf{x}] \alpha) \supset ([\mathbf{X} = \mathbf{x}] \chi))$.
6. $\alpha \supset (\psi \sqcup \chi) \equiv (\alpha \supset \psi) \sqcup (\alpha \supset \chi)$.
7. $\alpha \supset (\psi \rightarrow \chi) \equiv (\alpha \supset \psi) \rightarrow (\alpha \supset \chi)$.
8. $\models (\alpha \subset \supset \beta) \rightarrow [(\alpha \supset \chi) \rightarrow (\beta \supset \chi)]$.
9. If $\models \alpha \leftrightarrow \beta$, then $\models (\alpha \supset \chi) \rightarrow (\beta \supset \chi)$.
10. If $\models \psi \rightarrow \chi$, then $\models (\alpha \supset \psi) \rightarrow (\alpha \supset \chi)$.
11. $\langle \mathbf{X} = \mathbf{x} \rangle \varphi \leftrightarrow \langle \mathbf{X} = \mathbf{x} \rangle \varphi [\alpha \rightarrow \chi / \alpha \supset \chi]$.
12. $[\mathbf{X} = \mathbf{x}] \varphi \leftrightarrow [\mathbf{X} = \mathbf{x}] \varphi [\alpha \rightarrow \chi / \alpha \supset \chi]$.
13. $\langle \mathbf{X} = \mathbf{x} \rangle \varphi \leftrightarrow \langle \mathbf{X} = \mathbf{x} \rangle \varphi [\sim (\alpha \& \sim \chi) / \alpha \supset \chi]$.
14. $[\mathbf{X} = \mathbf{x}] \varphi \leftrightarrow [\mathbf{X} = \mathbf{x}] \varphi [\sim (\alpha \& \sim \chi) / \alpha \supset \chi]$.

Proof. 1) Straightforward.

2) $T \models \alpha \supset \sim \chi$ iff $T^\alpha \models \sim \chi$ iff $T^\alpha \not\models \chi$ iff $T \not\models \alpha \supset \chi$ iff $T \models \sim (\alpha \supset \chi)$.

3) $T \models \alpha \supset (\beta \supset \psi)$ iff $(T^\alpha)^\beta \models \psi$. But $(T^\alpha)^\beta = T^{\alpha \& \beta}$, so the former holds iff $T^{\alpha \& \beta} \models \psi$, iff $T \models (\alpha \& \beta) \supset \psi$.

4) Let $T = (T^-, \mathcal{F})$. If T is empty, note that, since α and β have the $\sim$ -restriction, the two formulas in the equivalence also have it. Thus, by Lemma 10.8, 1. they both have the empty team property; so, both formulas are satisfied by T . Let us then assume that T is nonempty. Note that $T \models [\mathbf{X} = \mathbf{x}] (\alpha \supset \beta)$ iff $(T_{\mathbf{X}=\mathbf{x}})^\alpha \models \beta$, while $T \models ([\mathbf{X} = \mathbf{x}] \alpha) \supset ([\mathbf{X} = \mathbf{x}] \beta)$ iff $(T^{[\mathbf{X}=\mathbf{x}] \alpha})_{\mathbf{X}=\mathbf{x}} \models \beta$. Then, it suffices to prove that $((T_{\mathbf{X}=\mathbf{x}})^\alpha)^- = ((T^{[\mathbf{X}=\mathbf{x}] \alpha})_{\mathbf{X}=\mathbf{x}})^-$.

If $s \in ((T_{\mathbf{X}=\mathbf{x}})^\alpha)^-$, then $(\{s\}, \mathcal{F}_{\mathbf{X}=\mathbf{x}}) \models \alpha$ and (since T is nonempty) there is a $t \in T^-$ such that $s \in t_{\mathbf{X}=\mathbf{x}}^\mathcal{F}$. Thus, $t \in (T^{[\mathbf{X}=\mathbf{x}] \alpha})^-$. But then $s \in ((T^{[\mathbf{X}=\mathbf{x}] \alpha})_{\mathbf{X}=\mathbf{x}})^-$.

Vice versa, assume $s \in ((T^{[\mathbf{X}=\mathbf{x}] \alpha})_{\mathbf{X}=\mathbf{x}})^-$. We want to prove, first, that $s \in (T_{\mathbf{X}=\mathbf{x}})^\alpha$. Since $T \models \langle \mathbf{X} = \mathbf{x} \rangle \alpha$, $T^{[\mathbf{X}=\mathbf{x}] \alpha}$ is nonempty; but then, $s \in (T^{[\mathbf{X}=\mathbf{x}] \alpha})_{\mathbf{X}=\mathbf{x}}^- = \bigcup_{t \in (T^{[\mathbf{X}=\mathbf{x}] \alpha})^-} t_{\mathbf{X}=\mathbf{x}}^\mathcal{F} \subseteq \bigcup_{t \in T^-} t_{\mathbf{X}=\mathbf{x}}^\mathcal{F} = T_{\mathbf{X}=\mathbf{x}}^-$.⁴⁴

Since $s \in (T_{\mathbf{X}=\mathbf{x}})^\alpha$ and T is nonempty, there is a $t \in T^-$ such that $s \in t_{\mathbf{X}=\mathbf{x}}^\mathcal{F}$. Since furthermore, $s \in (T^{[\mathbf{X}=\mathbf{x}] \alpha})_{\mathbf{X}=\mathbf{x}}^-$ and $T^{[\mathbf{X}=\mathbf{x}] \alpha}$ is nonempty, t can be chosen so that $(\{t\}, \mathcal{F}) \models [\mathbf{X} = \mathbf{x}] \alpha$. Thus, by the clause for the modal operator, $(\{s\}, \mathcal{F}) \models \alpha$. Thus, $s \in ((T_{\mathbf{X}=\mathbf{x}})^\alpha)^-$.

5) The proof proceeds as for 4., using the fact that $T \models \Diamond \top$ to exclude the case that T is empty.

6) Straightforward.

7) We observe that

$$\begin{aligned} T \models \alpha \supset (\psi \rightarrow \chi) &\iff T^\alpha \models \psi \rightarrow \chi \\ &\iff T^\alpha \not\models \psi \text{ or } T^\alpha \models \chi \\ &\iff T \not\models \alpha \supset \psi \text{ or } T \models \alpha \supset \chi \\ &\iff T \models (\alpha \supset \psi) \rightarrow (\alpha \supset \chi). \end{aligned}$$

⁴⁴ This step could also be proved, by a longer argument, without the assumption of nonemptiness of $T^{[\mathbf{X}=\mathbf{x}] \alpha}$.

8) Suppose $T \models \alpha \sqsubset \beta$ and $T \models \alpha \supset \chi$. From the latter, $T^\alpha \models \chi$. From the former, we obtain $T^\alpha = T^\beta$. Thus, $T^\beta \models \chi$, i.e., $T \models \beta \supset \chi$.

9) Analogous to the proof of 7, using the flatness of α , β .

10) Straightforward.

11) Fix a $T = (T^-, \mathcal{F})$. For any given assignment $s \in T^-$, we will use the notation $\eta \equiv_s \theta$ to say that $(s, \mathcal{F}) \models \eta$ iff $(s, \mathcal{F}) \models \theta$. We also write $\eta \equiv_0 \theta$ in case $(\emptyset, \mathcal{F}) \models \eta$ iff $(\emptyset, \mathcal{F}) \models \theta$.

The statement is equivalent to saying that there is an assignment $s \in T_{\mathbf{x}=\mathbf{x}}^-$ such that $\varphi \equiv_s \varphi[\alpha \rightarrow \chi / \alpha \supset \chi]$. We prove this simultaneously with the statement that $\varphi \equiv_0 \varphi[\alpha \rightarrow \chi / \alpha \supset \chi]$, by induction on φ .

If φ has no occurrence of $\alpha \supset \chi$, then the statement trivially holds; this covers the atomic case. The inductive steps are easy to prove, except for the case in which the most external operator of φ is $\supset$. We have two subcases, depending on whether φ is the occurrence $\alpha \supset \chi$ to be substituted, or another formula, namely $\beta \supset \psi$.

In the former case, showing $\varphi \equiv_s \varphi[\alpha \rightarrow \chi / \alpha \supset \chi]$ is straightforward. We show the statement for $\equiv_0$. Note that, if $(\emptyset, \mathcal{F}) \models \alpha \supset \chi$, then $(\emptyset, \mathcal{F})^\alpha \models \chi$. But $(\emptyset, \mathcal{F})^\alpha = (\emptyset, \mathcal{F})$. Thus, $(\emptyset, \mathcal{F}) \models \chi$. Thus, $(\emptyset, \mathcal{F}) \models \alpha \rightarrow \chi$. In the other direction, if $(\emptyset, \mathcal{F}) \models \alpha \rightarrow \chi$ we have either $(\emptyset, \mathcal{F}) \not\models \alpha$ or $(\emptyset, \mathcal{F}) \models \chi$. The former is excluded by the $\sim$ -restriction on α (since that makes α flat, and thus satisfy the empty team property). So, $(\emptyset, \mathcal{F}) = (\emptyset, \mathcal{F})^\alpha \models \chi$, i.e., $(\emptyset, \mathcal{F}) \models \alpha \supset \chi$.

If φ is of the form $\beta \supset \psi$, first note that $(\emptyset, \mathcal{F}) \models \beta \supset \psi$ iff $(\emptyset, \mathcal{F})^\beta = (\emptyset, \mathcal{F}) \models \psi$ iff (by the i.h.) $(\emptyset, \mathcal{F}) = (\emptyset, \mathcal{F})^{\beta[\alpha \rightarrow \chi / \alpha \supset \chi]} \models \psi[\alpha \rightarrow \chi / \alpha \supset \chi]$ iff $(\emptyset, \mathcal{F}) \models \beta[\alpha \rightarrow \chi / \alpha \supset \chi] \supset \psi[\alpha \rightarrow \chi / \alpha \supset \chi]$. On the other hand, $(\{s\}, \mathcal{F}) \models \beta \supset \psi$ iff $(\{s\}, \mathcal{F})^\beta \models \psi$. We can show that the latter is equivalent to $(\{s\}, \mathcal{F})^\beta \models \psi[\alpha \rightarrow \chi / \alpha \supset \chi]$. Indeed, if $(\{s\}, \mathcal{F})^\beta$ is nonempty, then the equivalence is given by the i.h. $\psi \equiv_s \psi[\alpha \rightarrow \chi / \alpha \supset \chi]$; if it is empty, the equivalence is given by the i.h. $\psi \equiv_0 \psi[\alpha \rightarrow \chi / \alpha \supset \chi]$. Finally, $(\{s\}, \mathcal{F})^\beta \models \psi[\alpha \rightarrow \chi / \alpha \supset \chi]$ is equivalent to $(\{s\}, \mathcal{F})^{\beta[\alpha \rightarrow \chi / \alpha \supset \chi]} \models \psi[\alpha \rightarrow \chi / \alpha \supset \chi]$ by the inductive hypothesis on β . Thus, it is equivalent to $(\{s\}, \mathcal{F}) \models \beta[\alpha \rightarrow \chi / \alpha \supset \chi] \supset \psi[\alpha \rightarrow \chi / \alpha \supset \chi]$.

The proofs of 11., 12., and 13. are analogous. $\square$

11 Comparisons with related work

As mentioned in the introduction, the idea of extending the Galles-Pearl framework of causal models with *one-to-many* causal laws is already suggested as a final remark in Halpern's original axiomatization work^[12]; on the other hand, we are not aware of any further development in this direction in the following two decades.

In 2021, Peters & Halpern^[17,25] developed a vast generalization of causal models (the *generalized structural equation models*, GSEM). Their generalizations go into two directions: on one hand, they remove the restriction to finite variable domains and finite value ranges, seemingly with the aim of modeling physical phenomena; and on the other hand, they allow for highly arbitrary interventions, which do not require causal laws for their computation. Such interventions associate a *set* of possible outcomes for the endogenous variables to each setting for the exogenous variables, and the only constraint that the interventions are enforced to abide to is that, say, an intervention $do(\mathbf{X} = \mathbf{x})$ should only produce sets of assignments that set the variables $\mathbf{X}$ to $\mathbf{x}$. It is clear that a causal model with indeterministic causal laws can be seen as a rather special case of a GSEM, although this is not remarked upon in these works; no talk of laws at all seems to be made in these papers. Given the large increase in generality—well beyond the simple addition of indeterministic laws—the GSEM framework seems quite unsuitable for obtaining a clear comparison between the usual causal models and their indeterministic generalization. And indeed, even when restricting attention to finite signatures, some of the

usual axioms like *Weak composition* (our I3) and *Weak reversibility* (our I6) fail.

In 2023, Wysocki^[3] focused explicitly on indeterministic causal laws (there called "underdeterministic") as an addition to the usual recursive causal models (although, not having to deal with the problem of axiomatization, he does not bother restricting value ranges to the finite case). The paper is mainly a philosophical defense of qualitative indeterministic laws as opposed to probabilistic laws, and provides, e.g., several examples where indeterministic laws may be available (either as a representation of real indeterminism or of lack of exact knowledge) while a probabilistic law is unavailable due to pragmatic or mathematical constraints. Its most interesting technical contributions are, arguably, 1) the idea of redefining direct causation in terms of variation in the range of possible values of Y when intervening on X having fixed all other variables (which we tried to formalize in section 7.1); and 2) the observation that the d -separation criterion is correct also for qualitative independence statements. For the rest, the presentation and some details of the framework—e.g., the definition of intervention, the semantics, and the syntax—are unusual and a bit vague, which makes it rather difficult to say to what extent Wysocki's framework agrees or differs with that which is considered in the present paper. It would look like formulas of the form $\Diamond\psi$ or $\Box\psi$ are evaluated over teams that are *maximal*, in the sense we mentioned in section 9.2, but allowing, rather than a single tuple of values for the exogenous variables, a cartesian product of value ranges for each exogenous variable (so as to ensure a qualitative form of independence among the exogenous variables). We have not considered this class of causal teams in the present paper.

In 2024, the proceedings version of the present paper^[38] appeared in print. Our work on the subject of indeterministic causal models began as early as 2020, and was influenced only in the later stages by the publication of Wysocki^[3] (from which we essentially only borrowed the generalized idea of direct causation that is discussed here in section 7.1). The paper introduced a formal definition of relational causal team (with *simplified* signatures) and sketched a description of the formalization obstacles that are presented more sharply here (sections 2 and 8). The paper also made some claims about axiomatization; regretfully, it was later discovered that the system therein introduced (as a purported axiomatization of the general class of models) is unsound. Indeed, the system in question includes axiom I6 (*Weak reversibility*), which, as we have seen, can be falsified by teams with more than one assignment. Secondly, its version of axiom I8 (which, differently from ours, does not rely on having the external/internal distinction in the signature) relies on the definability, in $\mathcal{H}$, of what we call here *visible direct cause* in order to enforce syntactically the assumption that this axiom only applies to external variables; unfortunately, doing so amounts to evaluating externality by looking at the incorrect graph (the visible causal graph instead of the parenthood graph). Lastly, there is an axiom $\Diamond\top \leftrightarrow \langle \mathbf{W} = \mathbf{w} \rangle \top$. Its right-to-left direction is in general false, since, as we have seen, the semantics allows for the possibility that intervening on an empty model may produce a nonempty team (this is actually *always* the case when intervening on the set $\mathbf{W}$ of all variables). Our system A replaces this incorrect principle with axioms I9 and I10. We also remark that, while I6 is at least sound on singleton models, the other two axioms we discussed here are not.

In Beckers^[26], Beckers contributed his own theory of indeterministic counterfactuals, which was then used in Beckers^[27] to formulate a new definition of actual causality. The indeterministic causal models used by Beckers (at least, in what he calls *swc semantics*) are essentially identical to our total recursive singleton models. The paper is

directly inspired by our work^[38], but Beckers proposes a different definition of counterfactual. Roughly, the truth value of his counterfactuals is assessed after a phase of observation (of values actually taken by the variables), followed by an intervention in the sense of Halpern. The observation phase is used (to the best of our understanding) to exclude interventions that would modify the actual value of a variable Y even though they do not change the values of the parents of Y (thus, in practice, enforcing the validity of the *Strong composition* axiom we discussed in section 8.2). So, for example, in a coin-tossing scenario where the coin has been tossed and it came heads, the counterfactual "had the coin been tossed, it would have come heads" is taken to be true (while it is false in our framework). Our impression is that Beckers' counterfactuals might be closer to Lewisian counterfactuals; their consequents are evaluated in (some of the) worlds that differ from the actual one only for making the antecedent true. The counterfactuals considered in our paper seem to be closer to the spirit of *interventionist* counterfactuals; their consequents are evaluated in worlds where not only is the antecedent true but it is also enforced by external intervention. Note also that it is only our paper's version of the counterfactual that preserves the *predictive* power of causal statements. Suppose, for example, that the Federal Council of Switzerland decides to declare war on Luxembourg upon the outcome of a coin toss (war shall be declared if a *tails* is tossed). *Heads* is tossed, thus preventing a suicidal war declaration. If we made the mistake of using Beckers' notion of counterfactual for policy-making, we would reach the obviously mistaken conclusion that it will be a good policy to use coin-tossing as a device to decide on the future diplomatic relations between Switzerland and Luxembourg. Using the counterfactuals from our paper, one will instead conclude that further coin-tossing *might* lead to a counter-productive war.^[45] It would seem, then, that we have a distinction between a philosophical and a practical notion of counterfactual; the differences between the two approaches are certainly worth future investigation. As for what concerns axiomatization, Beckers provides deduction systems for two rather special classes of models. First, he axiomatizes total recursive singleton models (what he calls *swc semantics*); his system differs from our A + Tot + R (see section 8) in that it features both *Strong composition* (which is unsound in our context) and *Weak composition* (our I3), it lacks *Nonemptiness* (I10) and has axiom $\Diamond\varphi \rightarrow \Box\varphi$ in place of our I12b. Furthermore, axioms I7 and I8 are absent due to differences in the syntax.^[46] The recursivity axiom is formulated in terms of an indeterministic version of "causally affecting" instead of our choice of direct cause. Second, Beckers provides an axiomatization for what he calls *scc semantics*, i.e., essentially for the (total, recursive) *maximal* causal teams that we briefly mentioned in section 9.2.

There is a different line of inquiry connecting counterfactuals and (qualitative) indeterminism, and that is the study of (using Bohm's terminology) *many-to-one* causation, i.e., cases where "many different kinds of causes can produce essentially the same effect"^[2]. We are aware of only three papers attempting to develop the technicalities behind this idea. The first one, from 2022, is Barbero & Yang^[19], in whose conclusion section, an idea is described of *indeterminate interventions*, i.e., interventions that are in some ways underspecified (e.g., "pressing *any* button on the remote", "picking one direction", "switching to the third or fifth gear"). Such interventions admit a natural

definition in the *generalized causal teams* studied in that paper.^[47] These interventions are introduced with the purpose of giving an intervention-based semantics for interventionist counterfactuals with disjunctive antecedents. While several semantic frameworks for such counterfactuals had been proposed elsewhere, e.g., studies^[14,39,40], all these earlier proposals seemed rather *ad hoc* and ungrounded in interventionist ideas. In 2023, Wysocki^[41] makes a similar proposal under the curious name of *semaphore interventions*.^[48] Since he does not cite our paper^[19], we can only assume that the idea was developed independently. This paper makes the interesting observation that for a simple enough language (say, $\mathcal{H}$ extended with the possibility of having disjunctive counterfactual antecedents), the semantics returns the same judgments of truth and falsity as Briggs's theory. Thus, the indeterministic interventions provide a clear interventionist motivation for Briggs's seemingly *ad hoc* truth-maker semantics.

12 Conclusions

We have shown that extending causal modeling to the case of indeterministic causal laws, as suggested in Halpern^[12], is doable but not as straightforward as Halpern suggested. We have seen that the notions of direct cause and causal parenthood are more complex than in the deterministic case, and even more complicated than suggested in Wysocki^[3], since it turns out that even the dummy arguments of causal laws can be direct causes. These insights lead us to the definition of appropriate models (relational causal teams).

We then produced two types of axiomatizations: one, exemplified by system A and its extensions, axiomatizes the Halpern-style language $\mathcal{H}$ but is only applicable to singleton models (since the *Weak reversibility* axiom I6 is unsound over multiple-assignment causal teams). This semantics is the closest to the original causal models, and we used it to compare the deterministic and indeterministic cases. The second type of proof system, given by system B and its extensions, axiomatizes the language $\mathcal{H}^+$ with *right-nested* counterfactuals and applies to the class of *all* models (i.e., all relational causal teams). The axiomatization is somewhat unsatisfactory, as one of the axioms (J17) is rather *ad hoc* and heavily relies on the semantics; what remains unsettled is finding an elegant family of "rules" to transform $\mathcal{H}^+$ formulas into unnested, $\mathcal{H}$ formulas. We believe that the result has in any case some value, in that it is (to the best of our knowledge) the first known axiomatization of right-nested counterfactuals outside of the recursive deterministic case.^[49] Its proof also provides a detailed exploration (perhaps the first) of the connections between interventionist counterfactuals and (dynamic) modal logic. On the other hand, while system B does provide a recursive enumeration of the validities of language $\mathcal{H}$, the direct axiomatization of $\mathcal{H}$ over the class of all models remains an open question.

We also saw how to specialize systems A and B (by adding new axioms) to a number of significant subclasses of models: the strictly/visibly recursive, deterministic, total and total recursive subclasses (the corresponding additional axioms can also be combined modularly). The logic of the recursive class differs from its deterministic counterpart in a few respects, among which we may remark the

^[45] One might retort that it is only a lucky accident that interventionist counterfactuals turn out to have predictive power in the deterministic case; after all, counterfactuals concern past, not future events. We thank S. Beckers for this observation.

^[46] Becker's syntax only uses endogenous variables—thus, our I8 cannot even be formulated, and identifies modal-free formulas φ with $\Box\varphi$, making axiom I7 useless.

^[47] While a causal team encodes uncertainty about the state of the variables under a fixed hypothesis about the causal laws, a generalized causal team encodes uncertainty about both the state of the variables and the causal laws.

^[48] The idea is also described, less explicitly, in Wysocki^[3].

^[49] The recursive deterministic case is easily dealt with, as in that case the diamond modalities are interchangeable with box modalities, and the sequence of modalities $[\mathbf{X} = \mathbf{x}][\mathbf{Y} = \mathbf{y}]$ can always be contracted into a single modality. See Briggs^[14] and Barbero & Sandu^[28].

failure of the (strong) *Composition* law. We also observed that the *Reversibility* law holds in the strictly recursive case, showing that its traditional connection to the property of uniqueness of solutions breaks in the recursive indeterministic context.

Our axiomatization results rely to some extent on using signatures that fix the distinction between internal and external variables. We have seen that, abandoning this constraint, some important notions, such as determinism and the external/internal distinction, become undefinable in the usual language $\mathcal{H}$. We have seen that the severity of these problems decreases when restricting attention to singleton models (determinism becomes again definable), so these technical obstacles might be seen as a counterindication to using teams rather than singleton models. However, they may instead be seen as evidence that the usual causal language $\mathcal{H}$ is insufficiently expressive in the indeterministic context; we have seen that determinism is definable when using the ordinary signatures, and becomes definable also with simplified signatures as soon as one allows for (right-)nested counterfactuals. We have also seen that something similar happens if $\mathcal{H}$ is enriched with the conditional $\supset$ expressing the results of learning and observations, which was proposed in Barbero & Sandu^[28]. We conjecture that these two extended languages, $\mathcal{H}^+$ and $\mathcal{HO}$, might be expressively complete (further work will be needed just to make this statement precise). For the moment, an important open question is whether using such languages, or by restricting attention to singleton models, one might be able to define the important notion of direct cause; we have seen that $\mathcal{H}$ is insufficient for this purpose, with both types of signatures. Other language extensions might be worth considering in the future; e.g., in Barbero & Yang^[19], we suggested a strategy for allowing complex, disjunctive antecedents, by assigning them a semantics by means of *indeterminate interventions*. This extension requires a further generalization of the models and is not explored in the present paper, although plausibly the generalization is straightforward.

A number of natural directions of investigation open ahead, among which is the axiomatization problem for classes of models with simplified signatures; the axiomatization of $\mathcal{HO}$; a systematical comparison with the logic of Stalnaker-Lewis counterfactuals^[42] in the spirit of studies^[15,22,43]; and the analysis of the computational and descriptive complexity of the formalism (cf., for the deterministic case, paper^[44] and Halpern's *Actual causality* book^[13], Chapter 5. A rather different direction to explore is the development of languages that can express data dependencies along with interventionist counterfactuals. This was done in the deterministic, recursive case in Barbero & Sandu^[28] and Barbero & Yang^[19]. We note that the definition of counterfactual given in the present paper is clearly inadequate for this task: the consequence of a counterfactual is evaluated assignment by assignment, while data dependencies are global properties of a whole team. The natural move is then to use the following definition of interventionist counterfactual from Barbero & Sandu^[28]:

- $\mathbf{X} = \mathbf{x} \square \rightarrow \psi$ iff $T_{\mathbf{X}=\mathbf{x}} \models \psi$.

One might also need a setup more similar to that of the common logics with team semantics, featuring connectives such as the *dual negation* and the *tensor disjunction* instead of those used here. In this case, the *might* counterfactual would need to be introduced as a distinct operator, as follows:

- $\mathbf{X} = \mathbf{x} \diamond \rightarrow \psi$ iff there is a nonempty $S \leq T_{\mathbf{X}=\mathbf{x}} \models \psi$ such that $S \models \psi$.

We may note that $\diamond \rightarrow$ is an operator that preserves some interesting closure properties of formulas—such as union closure and convexity—but does not preserve downward closure. Thus, the purely counterfactual part of such a language—without the addition of data dependencies—is already quite more complex than in the deterministic

recursive case, where it happened to be just a flat, "classical" language. The properties and proof theory of such languages are still largely unexplored, and plausibly depart considerably from what may be familiar to the experts of causal reasoning; for these reasons, in this paper, we chose to stick to languages closer to the Galles-Pearl-Halpern tradition.

Author contributions

The author confirmed sole responsibility for all aspects of this study and approved the final version of the manuscript.

Data availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Acknowledgments

The author wishes to thank Christopher Hitchcock, Sanders Beckers, and the late Joseph Halpern for intense discussions on the topic and contents of the paper. The author's research was supported by the Research Council of Finland's grant no. 349803.

Conflict of interest

The author declares that there is no conflict of interest.

Supplementary information accompanies this paper online at: <https://doi.org/10.48130/ker-0026-0005>.

Dates

Received 14 December 2024; Revised 29 October 2025; Accepted 10 March 2026; Published online 5 August 2026

References

- [1] Paul LA, Hall EJ. 2013. *Causation: a user's guide*. UK: Oxford University Press. doi: [10.1093/acprof:oso/9780199673445.001.0001](https://doi.org/10.1093/acprof:oso/9780199673445.001.0001)
- [2] Bohm D. 1957. *Causality and chance in modern physics*. Princeton, NJ: Van Nostrand Company, Inc.
- [3] Wysocki T. 2026. The underdeterministic framework. *The British Journal for the Philosophy of Science* 77:245–270
- [4] Spirtes P, Glymour C, Scheines RN. 1993. *Causation, prediction, and search*. Vol. 81. New York: Springer. doi: [10.1007/978-1-4612-2748-9](https://doi.org/10.1007/978-1-4612-2748-9)
- [5] Pearl J. 2000. *Causality: models, reasoning, and inference*. New York, NY, USA: Cambridge University Press.
- [6] Peters J, Janzing D, Schölkopf B. 2017. *Elements of Causal Inference: Foundations and Learning Algorithms*. US: The MIT Press.
- [7] Schölkopf B. 2022. Causality for machine learning. In *Probabilistic and causal inference: the works of Judea Pearl*. New York, USA: Association for Computing Machinery. pp. 765–804 doi: [10.1145/3501714.3501755](https://doi.org/10.1145/3501714.3501755)
- [8] Hernan MA, Robins J. 2020. *Causal inference: what if*. Boca Raton: Chapman & Hill/CRC.
- [9] Heckman JJ, Vytlacil EJ. 2007. Econometric evaluation of social programs, part I: causal models, structural models and econometric policy evaluation. *Handbook of Econometrics* 6:4779–4874
- [10] Morgan SL, Winship C. 2015. *Counterfactuals and causal inference*. New York, UK: Cambridge University Press. doi: [10.1017/CBO9781107587991](https://doi.org/10.1017/CBO9781107587991)
- [11] Galles D, Pearl J. 1998. An axiomatic characterization of causal counterfactuals. *Foundations of Science* 3(1):151–182
- [12] Halpern JY. 2000. Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research* 12(1):317–337

- [13] Halpern JY. 2016. *Actual Causality*. Cambridge, Massachusetts: The MIT Press.
- [14] Briggs R. 2012. Interventionist counterfactuals. *Philosophical Studies* 160(1):139–166
- [15] Zhang J. 2013. A Lewisian logic of causal counterfactuals. *Minds and Machines* 23(1):77–93
- [16] Ibeling D, Icard T. 2020. Probabilistic reasoning across the causal hierarchy. *Proceedings of the AAAI Conference on Artificial Intelligence* 34:10170–10177
- [17] Halpern JY, Peters S. 2022. Reasoning about causal models with infinitely many variables. *Proceedings of the AAAI Conference on Artificial Intelligence* 36:5668–5675
- [18] Barbero F, Schulz K, Velazquez-Quesada FR, Xie K. 2022. Observing interventions: a logic for thinking about experiments. *Journal of Logic and Computation* 33:1152–1185
- [19] Barbero F, Yang F. 2022. Characterizing counterfactuals and dependencies over (generalized) causal teams. *Notre Dame Journal of Formal Logic* 63(3):301–341
- [20] Beckers S, Halpern J, Hitchcock C. 2023. Causal models with constraints. *Proceedings of the Second Conference on Causal Learning and Reasoning, Tubingen, Germany, 2023*. PMLR. pp. 866–879. <https://proceedings.mlr.press/v213/beckers23a/beckers23a.pdf> (accessed on 30 April 2026)
- [21] Barbero F, Virtema J. 2023. Strongly complete axiomatization for a logic with probabilistic interventionist counterfactuals. *Logics in artificial intelligence*, eds. Gaggl S, Martinez MV, Ortiz M. Cham: Springer. pp. 649–664 doi: [10.1007/978-3-031-43619-2_44](https://doi.org/10.1007/978-3-031-43619-2_44)
- [22] Fang J, Zhang J. 2023. A characterization of Lewisian causal models. *International Workshop on Logic, Rationality and Interaction. LORI 2023. Lecture Notes in Computer Science*. Cham: Springer. pp. 94–108 doi: [10.1007/978-3-031-45558-2_8](https://doi.org/10.1007/978-3-031-45558-2_8)
- [23] Kawamoto Y, Sato T, Suenaga K. 2023. Formalizing statistical causality via modal logic. *European Conference on Logics in Artificial Intelligence, JELIA 2023. Lecture Notes in Computer Science*. Cham: Springer. pp. 681–696 doi: [10.1007/978-3-031-43619-2_46](https://doi.org/10.1007/978-3-031-43619-2_46)
- [24] Ding Y, Manoorkar K, Tzimoulis A, Wang R, Wang X. 2023. Causal kripke models. *Electronic Proceedings in Theoretical Computer Science* 379:185–200
- [25] Peters S, Halpern JY. 2021. Causal modeling with infinitely many variables. *arXiv Preprint*
- [26] Beckers S. 2025. Nondeterministic causal models. *Proceedings of the Fourth Conference on Causal Learning and Reasoning. Proceedings of Machine Learning Research, Lausanne, Switzerland, 2025*, eds. Huang B, Drton M. Vol. 275. PMLR. pp. 1532–1554 <https://proceedings.mlr.press/v275/beckers25b.html>
- [27] Beckers S. 2025. Actual causation and nondeterministic causal models. *arXiv Preprint*
- [28] Barbero F, Sandu G. 2021. Team semantics for interventionist counterfactuals: observations vs. interventions. *Journal of Philosophical Logic* 50:471–521
- [29] Strasser C, Antonelli GA. 2024. Non-monotonic Logic. In *The Stanford Encyclopedia of Philosophy (Winter 2024 Edition)*, eds. Zalta EN, Nodelman U. <https://plato.stanford.edu/entries/logic-nonmonotonic/index.html>
- [30] Barbero F, Galliani P. 2022. Embedding causal team languages into predicate logic. *Annals of Pure and Applied Logic* 173:103159
- [31] Pearl J. 2017. Physical and metaphysical counterfactuals: evaluating disjunctive actions. *Journal of Causal Inference* 5(2):20170018
- [32] Blackburn P, De Rijke M, Venema Y. 2001. *Modal Logic*. Vol. 53. Cambridge, England: Cambridge University Press.
- [33] Wang Y, Cao Q. 2013. On axiomatizations of public announcement logic. *Synthese* 190:103–134
- [34] Woodward J. 2003. *Making things happen: a theory of causal explanation*. Vol. 114. Oxford Studies in the Philosophy of Science. Oxford University Press.
- [35] Zhang J, Lam WY, De Clercq R. 2013. A peculiarity in Pearl's logic of interventionist counterfactuals. *Journal of Philosophical Logic* 42:783–794
- [36] Barbero F, Virtema J. 2024. Expressivity landscape for logics with probabilistic interventionist counterfactuals. *32nd EACSL Annual Conference on Computer Science Logic, CSL 2024*. Vol. 288. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. pp. 15:1–15:19 doi: [10.4230/LIPIcs.CSL.2024.15](https://doi.org/10.4230/LIPIcs.CSL.2024.15)
- [37] Barbero F, Sandu G. 2024. Multiteam semantics for interventionist counterfactuals: probabilities and causation. *Journal of Philosophical Logic* 53(6):1537–1577
- [38] Barbero F. 2024. On the logic of interventionist counterfactuals under indeterministic causal laws. *International Symposium on Foundations of Information and Knowledge Systems*. Cham: Springer. pp. 203–221 doi: [10.1007/978-3-031-56940-1_11](https://doi.org/10.1007/978-3-031-56940-1_11)
- [39] Ciardelli I, Zhang L, Champollion L. 2018. Two switches in the theory of counterfactuals: a study of truth conditionality and minimal change. *Linguistics and Philosophy* 41(6):577–621
- [40] Santorio P. 2019. Interventions in premise semantics. *Philosophers' Imprint* 19(1):1–27
- [41] Wysocki T. 2023. An event algebra for causal counterfactuals. *Philosophical Studies* 180(12):3533–3565
- [42] Lewis D. 1973. *Counterfactuals*. Oxford: Blackwell Publishers. <https://perso.uclouvain.be/peter.verdee/counterfactuals/lewis.pdf>
- [43] Halpern JY. 2013. From causal models to counterfactual structures. *The Review of Symbolic Logic* 6(2):305–322
- [44] Mossé M, Ibeling D, Icard T. 2024. Is causal reasoning harder than probabilistic reasoning? *The Review of Symbolic Logic* 17:106–131



Copyright: © 2026 by the author(s). Published by Maximum Academic Press, Fayetteville, GA. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.