A Unifying Framework for Causal Modeling With Infinitely Many Variables

SPENCER PETERS*, Cornell University, USA
JOSEPH Y. HALPERN, Cornell University, USA

Structural-equations models (SEMs) are perhaps the most commonly used framework for modeling causality, but they do not capture all domains of interest. For example, dynamical systems that evolve in continuous time are an important class of domains that are not (naturally) captured by SEMs. A wide variety of approaches have been proposed to fill the gap, including *dynamical structural causal models* (Bongers, Blom and Mooij 2018), *causal constraints models* (Blom, Bongers and Mooij 2019), and *counterfactual resimulation* (Laurent, Yang, and Fontana 2018). These models complement common-sense causal interpretations of specific dynamical systems, such as systems of ODEs. All these approaches look quite different from each other and from SEMs. They are hard to compare, and concepts developed for one approach may not make sense for another. But they are capturing the same notion of causality as SEMs do, in the sense that interventions map to outcomes. We propose a class of models that are, in a certain natural sense, the most expressive generalization of SEMs. Our generalized SEMs (GSEMs) can be viewed as a unifying framework that recovers structural dynamical causal models, causal constraints models, counterfactual resimulation, and common-sense causal interpretations of systems of ODEs and hybrid automata (Alur et al. 1992) as special cases. The input-output behavior, or "interface", of GSEMs is exactly that of SEMs, which means that definitions of concepts like actual cause, responsibility, blame, and explanation, can be immediately lifted from SEMs to GSEMs. The generality of GSEMs also makes them ideally suited to studying causality in the abstract; for example, they have been used to establish independence relationships among Halpern's axioms for SEMs (Peters and Halpern 2022).

JAIR Associate Editor: Sriraam Natarajan

JAIR Reference Format:
Spencer Peters and Joseph Y. Halpern. 2025. A Unifying Framework for Causal Modeling With Infinitely Many Variables. *Journal of Artificial Intelligence Research* 83, Article 36 (August 2025), 21 pages. DOI: 10.1613/jair.1.15612

1 Introduction

Structural equations models (SEMs) are a popular paradigm for modeling causality. A SEM describes causal relationships between variables using *structural equations*, one for each variable. Actions (interventions) taken by a scientist or policymaker are modeled as modifications to the structural equations; for example, if X represents whether a patient is taking a certain drug, assigning the patient to the treatment group amounts to replacing the equation for X with the constant equation $X = 1$, producing a new SEM.

While SEMs have proved their utility in many application domains, their structure makes them difficult to apply to *dynamical systems*, where the state changes continuously over time. To fill this gap, a variety of approaches have been developed to capture the causal semantics of dynamical systems. For example, Laurent, Yang, and Fontana developed a causal semantics for *rule-based models*, a popular class of models in molecular biology, and Bongers, Blom, and Mooij developed models for systems of ordinary differential equations [4] and their

*Corresponding Author.

Authors' Contact Information: Spencer Peters, ORCID: 0000-0002-9248-107X, sp2473@cornell.edu, Cornell University, Ithaca, New York, USA; Joseph Y. Halpern, ORCID: 0000-0002-9229-1663, halpern@cs.cornell.edu, Cornell University, Ithaca, New York, USA.

DOI: 10.1613/jair.1.15612

equilibrium behavior [3]. In addition, systems of ordinary differential equations have a natural causal semantics that is often used implicitly. Intuitively, intervening on the state of the system at time t sets a new initial condition, but does not affect the underlying rules (equations) governing the evolution of the system from that initial condition. Interestingly, these semantics are not fully captured by the structural dynamical causal models of Bongers, Blom, and Mooij 2018, who consider interventions on the rules, not on the state of the system.

Although the existence of approaches like structural dynamical causal models is compelling evidence that SEMs fail to handle dynamical systems in general, it's worth considering the reason why. SEMs are perfectly capable of handling systems that evolve in *discrete* time, using X_1 to represent the value of X at time 1, X_2 to represent the value of X at time 2, and so on. The structural equation for X_2 determines X_2 in terms of, say, X_1 and Y_1. However, in continuous time, the situation is more complicated. For example, the solution to a system of differential equations is described by a set of conditions on the variables on a neighborhood of each point in time; there is no natural way to break these conditions into one equation for each timestamped variable X_t. The definition of SEMs in terms of structural equations is simply too constrained to accommodate the rules by which dynamical systems evolve. For a concrete example of the limitations of SEMs, notice that the discrete-time SEMs above (indeed all so-called *acyclic* SEMs) have unique solutions. But dynamical systems that evolve in continuous time may have multiple solutions. For example, the differential equation $dx/dt = x^{2/3}$ with initial condition $x(0) = 0$ has the solution $x(t)$ defined by $x(t) = 0$ for $t \leq 0$ and $x(t) = t^3/27$ for $t > 0$. But this is not the only solution; for any $c > 0$, $x_c(t) := f(t - c)$ is also a solution. Hence no acyclic SEM (in particular, no SEM whose equation for X_t depends only on X_s for $s < t$) can capture the behavior of this differential equation. While this technically does not rule out the possibility that a cyclic SEM could, we believe that any such SEM would be considerably less natural (and more difficult to work with) than the timestamped SEMs described above.

Thus, to capture the natural causal semantics of dynamical systems, we propose a more flexible class of models that we call *generalized structural-equations models* (GSEMs). It is easy to show that GSEMs indeed generalize SEMs (see Theorem 3.1). Given a SEM and an intervention, we can produce a new SEM that represents the result of the intervention by modifying the relevant equations. GSEMs represent the same input-output relationships as SEMs—a GSEM and an allowed intervention maps to a new GSEM, while a GSEM and a context together determine a set of possible outcomes (assignments to the endogenous variables)—without committing to a specific mechanism for producing the outcomes. Indeed, any mapping from interventions and contexts to outcomes (all defined with respect to some set of variables) is a GSEM. In this sense, GSEMs are the most general causal model that has the same "interface" or input-output behavior as SEMs. In fact, even if we restrict to settings with only finitely many variables, GSEMs are more expressive than SEMs; see Example 3.6.

We show that GSEMs can capture a wide variety of approaches, including rule-based models [14], structural dynamical causal models [4], causal constraints models [3], common-sense causal interpretations of systems of ODEs, and hybrid automata [1]. Thus, GSEMs serve as a unifying framework for many disparate models of causality. Moreover, because GSEMs have the same interface as SEMs, any definition depending only on inputs and outputs (interventions and their outcomes) can be immediately applied to GSEMs. In particular, the notion of actual causality (whether event X caused event Y in some concrete situation) given by Halpern and Pearl 2005 and later modified by Halpern 2015, 2016 can be applied to GSEMs almost without modification (see Section 4). This means that environmental scientists studying predator-prey dynamical systems and molecular biologists studying chemical reaction pathways can use the same language of definitions to describe actual causality and related notions. GSEMs have also proven useful for studying causality in the abstract. In a follow-up paper [15], we give sound and complete axiomatizations for GSEMs and several subclasses of GSEMs, building from Halpern's axiomatization for SEMs [9]. These axiomatizations clarify the properties originally captured by Halpern's axioms.

GSEMs, by virtue of their generality, do not have the features that more specific models like SEMs have. SEMs have a very useful graphical representation; in general, GSEMs do not. Finite acyclic SEMs come with an efficient algorithm to compute their unique outcome; like recursive SEMs, GSEMs do not. However, when GSEMs are used as a unifying framework, these drawbacks disappear. Using a GSEM to capture another model, we still retain all the specific advantages of that model, and in addition gain access to generic notions like actual cause that might not have been defined for that model.[1]

Causal constraints models (CCMs) [3], which were introduced to get around some of the restrictions of SEMs when modeling equilibrium solutions of dynamical systems, can be shown to be as expressive as GSEMs (Theorem 6.1), and thus, in principle, can also capture these other models of causality.[2] However, the definition of GSEMs is much simpler than the definition of CCMs. Moreover, since GSEMs are a more straightforward generalization of SEMs than CCMs, they are arguably easier to use for those familiar with SEMs (as evidenced by the ease with which we carry over the definition of actual causality). More generally, the fact that GSEMs have the same input-output behavior as SEMs makes it straightforward to lift definitions like that of actual causality [10, 8], explanation [11, 8], and responsibility and blame [5] from SEMs to GSEMs.

The rest of the paper is organized as follows. In Section 2, we review the standard SEM framework. In Section 3, we formally define GSEMs, relate them to SEMs, and describe some of their advantages in more detail. In Section 4 we review the definition of actual cause in SEMs and explain that it can be used without change in GSEMs. In Section 5, we show how to use GSEMs to model three different dynamical-systems formalisms: systems of ordinary differential equations, rule-based models, and hybrid automata. In Section 6 we briefly compare GSEMs to several alternative modeling approaches.

2 SEMs: a Review

Formally, a *structural-equations model* M is a pair $(\mathcal{S}, \mathcal{F})$, where $\mathcal{S}$ is a *signature*, which explicitly lists the endogenous and exogenous variables and characterizes their possible values, and $\mathcal{F}$ defines a set of *modifiable structural equations*, relating the values of the variables. We extend the signature to include a set of *allowed interventions*, as was done in earlier work [2, 16]. Intuitively, allowed interventions are the ones that are feasible or meaningful.

A signature $\mathcal{S}$ is a tuple $(\mathcal{U}, \mathcal{V}, \mathcal{R}, \mathcal{I})$. $\mathcal{U}$ is a set of exogenous variables, $\mathcal{V}$ is a set of endogenous variables, and $\mathcal{R}$ associates with every variable $Y \in \mathcal{U} \cup \mathcal{V}$ a nonempty, finite set $\mathcal{R}(Y)$ of possible values for Y (i.e., the set of values over which Y *ranges*). Since we will want to compare to SEMs as they are typically defined in the literature, we will make the standard assumption that $\mathcal{U}$ and $\mathcal{V}$ are finite sets. However, as explained in Section 1, relaxing this restriction does not solve the issues that GSEMs are designed to solve.

We also adopt the convention that for $\vec{Y} \subseteq \mathcal{U} \cup \mathcal{V}$, $\mathcal{R}(\vec{Y})$ denotes the product of the ranges of the variables appearing in $\vec{Y}$; that is, $\mathcal{R}(\vec{Y}) \coloneqq \times_{Y \in \vec{Y}} \mathcal{R}(Y)$. Finally, an intervention $I \in \mathcal{I}$ is a set of pairs (X, x), where $X \in \mathcal{V}$ and $x \in \mathcal{R}(X)$. For each $X \in \mathcal{V}$, there is at most one $x \in \mathcal{R}(X)$ with $(X, x) \in I$.

We abbreviate an intervention I by $\vec{X} \leftarrow \vec{x}$, where $\vec{X} \subseteq \mathcal{V}$ and, unless $\vec{X}$ is empty, $\vec{x} \in \mathcal{R}(\vec{X})$. Although this notation makes most sense if $\vec{X}$ is nonempty, we allow $\vec{X}$ to be empty (which amounts to not intervening at all). If I consists of exactly one pair (Y, y), we abbreviate I as $Y \leftarrow y$.

$\mathcal{F}$ associates with each endogenous variable $X \in \mathcal{V}$ a function denoted F_X such that $F_X : \mathcal{R}(\mathcal{U} \cup \mathcal{V} - \{X\}) \to \mathcal{R}(X)$. This mathematical notation just makes precise the fact that F_X determines the value of X, given the values of all the other variables in $\mathcal{U} \cup \mathcal{V}$. If there is one exogenous variable U and three endogenous variables, X,

[1]See Section 5.2 for a concrete example.

[2]However, this was certainly not the intention of Blom, Bongers, and Mooij. They wanted to investigate systems with constraints, such as the equilibria reached by dynamical systems. Concretely, the models of these equilibria in [3] only involve finitely many real-valued variables, in contrast to our GSEM models of systems of ODEs, which involve uncountably many.

Y, and Z, then F_X defines the values of X in terms of the values of Y, Z, and U. For example, we might have $F_X(u, y, z) = u + y$, which is usually written as $X = U + Y$. Thus, if $Y = 3$ and $U = 2$, then $X = 5$, regardless of how Z is set.

The structural equations define what happens in the presence of external interventions. Setting the value of some variable X to x in a SEM $M = (\mathcal{S}, \mathcal{F})$ results in a new SEM, denoted $M_{X \leftarrow x}$, which is identical to M, except that the equation for X in $\mathcal{F}$ is replaced by $X = x$. Interventions on subsets $\vec{X}$ of $\mathcal{V}$ are defined similarly. Notice that $M_{\vec{X} \leftarrow \vec{x}}$ is always well defined, even if $(\vec{X} \leftarrow \vec{x}) \notin \mathcal{I}$. In earlier work, the reason that the model included allowed interventions was that, for example, relationships between two models were required to hold only for allowed interventions (i.e., the interventions that were meaningful). Here the set of allowed interventions also plays a role in how GSEMs relate to SEMs (see Section 3.2).

Given a context $\mathbf{u} \in \mathcal{R}(\mathcal{U})$, the *outcomes* of a SEM M under intervention $\vec{X} \leftarrow \vec{x}$ are all assignments of values $\mathbf{v} \in \mathcal{R}(\mathcal{V})$ such that the assignments $\mathbf{u}$ and $\mathbf{v}$ together satisfy the structural equations of $M_{\vec{X} \leftarrow \vec{x}}$. This set of outcomes is denoted $M(\mathbf{u}, \vec{X} \leftarrow \vec{x})$. Given an outcome $\mathbf{v}$, we denote by $\mathbf{v}[X]$ and $\mathbf{v}[\vec{X}]$ the value that $\mathbf{v}$ assigns to X and the restriction of $\mathbf{v}$ to $\mathcal{R}(\vec{X})$ respectively. We also use this notation for interventions; for example, $\vec{y}[X]$ is the value that intervention $\vec{Y} \leftarrow \vec{y}$ assigns to variable $X \in \vec{Y}$.

As discussed in the introduction, an important special case of SEMs are acyclic (or recursive) SEMs. Formally, an acyclic SEM is one for which, for every context $\mathbf{u} \in \mathcal{R}(\mathcal{U})$, there is some total ordering $\prec_{\mathbf{u}}$ of the endogenous variables (the ones in $\mathcal{V}$) such that if $X \prec_{\mathbf{u}} Y$, then X is *independent* of Y, that is, $F_X(\mathbf{u}, \dots, y, \dots) = F_X(\mathbf{u}, \dots, y', \dots)$ for all $y, y' \in \mathcal{R}(Y)$. Intuitively, if a model is acyclic, there is no feedback. Acyclic models always have unique outcomes; this is a consequence of assuming that $\mathcal{V}$ is finite.

In order to talk about SEMs and the information they represent more precisely, we use the formal language $\mathcal{L}(\mathcal{S})$ for SEMs having signature $\mathcal{S}$, introduced by Halpern 2000; see also [6]. An informal description of this language follows; for more details, see [9]. We restrict the language used by Halpern 2000 to formulas that mention only allowed interventions. Fix a signature $\mathcal{S} = (\mathcal{U}, \mathcal{V}, \mathcal{R}, \mathcal{I})$. Given an assignment $\mathbf{v} \in \mathcal{R}(\mathcal{V})$, the *primitive event* $X = x$ is true of $\mathbf{v}$, written $\mathbf{v} \models (X = x)$, if $\mathbf{v}[X] = x$; otherwise it is false. We extend this definition to *events* φ, which are Boolean combinations of primitive events, in the obvious way. Given a SEM M with signature $\mathcal{S}$ and an allowed intervention $\vec{Y} \leftarrow \vec{y} \in \mathcal{I}$, the *atomic causal formula* $[\vec{Y} \leftarrow \vec{y}]\varphi$ is true in context $\mathbf{u}$, written $(M, \mathbf{u}) \models [\vec{Y} \leftarrow \vec{y}]\varphi$ if, for all outcomes $\mathbf{v} \in M(\mathbf{u}, \vec{Y} \leftarrow \vec{y})$, we have $\mathbf{v} \models \varphi$. Again, we extend this definition to *causal formulas*, which are Boolean combinations of atomic formulas, in the obvious way. The language $\mathcal{L}(\mathcal{S})$ consists of all causal formulas (over $\mathcal{S}$). Using these formulas, we can also talk about properties that only *some* of the outcomes $\mathbf{v} \in M(\mathbf{u}, \vec{Y} \leftarrow \vec{y})$ have. For an event φ, define $\langle \vec{Y} \leftarrow \vec{y} \rangle \varphi$ as $\neg[\vec{Y} \leftarrow \vec{y}](\neg\varphi)$. This formula is true exactly when φ is true of at least one outcome $\mathbf{v} \in M(\mathbf{u}, \vec{Y} \leftarrow \vec{y})$.

The language of causal formulas completely characterizes the outcomes of a causal model with finite outcome sets, in the following precise sense. (For the purposes of this paper, a *causal model* is either a SEM, a GSEM, or a CCM.)

Theorem 2.1. *If M and M' are causal models over the same signature $\mathcal{S}$ that, given a context and intervention, return a finite set of outcomes, then M and M' have the same outcomes (that is, for all $\mathbf{u} \in \mathcal{R}(\mathcal{U})$ and $I \in \mathcal{I}$, $M(\mathbf{u}, I) = M'(\mathbf{u}, I)$) if and only if they satisfy the same set of causal formulas (that is, for all $\mathbf{u} \in \mathcal{R}(\mathcal{U})$ and $\psi \in \mathcal{L}(\mathcal{S})$, $M, \mathbf{u} \models \psi \Leftrightarrow M', \mathbf{u} \models \psi$).*

We note that it suffices to state and prove the result for GSEMs, since SEMs and CCMs are special cases of GSEMs. The proof of this and all other results not in the main text can be found in the appendix.

A short note on notation; as in the current section, throughout the paper, we use capital letters X, Y to denote variables, lowercase letters x, y to denote the corresponding values; letters with arrows $\vec{X}, \vec{Y}, \vec{x}, \vec{y}$ to denote vectors of variables and their corresponding vectors of values. We use boldface $\mathbf{u}$ and $\mathbf{v}$ for contexts and outcomes, respectively. We use M, M' to denote causal models, script letters $\mathcal{S}, \mathcal{U}, \mathcal{V}, \mathcal{R}, \mathcal{I}$ to denote a model's signature and its components, script $\mathcal{F}$ for a SEM's structural equations, and boldface $\mathbf{F}$ for a GSEM's outcomes mapping (see below).

3 Generalized Structural-Equations Models

The main purpose of causal modeling is to reason about a system's behavior under intervention. A SEM can be viewed as a function that takes a context $\mathbf{u}$ and an intervention $\vec{Y} \leftarrow \vec{y}$ and returns a set of outcomes, namely, the set of all solutions to the structural equations after replacing the equations for the variables in $\vec{Y}$ with $\vec{Y} = \vec{y}$. Viewed in this way, generalized structural-equations models (GSEMs) are a natural generalization of SEMs. In a GSEM, there is a function $\mathbf{F}$ that takes a context $\mathbf{u}$ and an intervention $\vec{Y} \leftarrow \vec{y}$ and returns a set of outcomes. However, the outcomes need not be determined by solving a set of suitably modified equations as they are for SEMs. This relaxation gives GSEMs the ability to concisely represent dynamical systems and other systems with infinitely many variables, and the flexibility to handle situations involving finitely many variables that cannot be modeled by SEMs.

3.1 GSEMs and SEMs

Formally, a *generalized structural-equations model (GSEM) M* is a pair $(\mathcal{S}, \mathbf{F})$, where $\mathcal{S}$ is a signature, and $\mathbf{F}$ is a mapping from contexts and interventions to sets of outcomes. More precisely, a signature $\mathcal{S}$ is a quadruple $(\mathcal{U}, \mathcal{V}, \mathcal{R}, \mathcal{I})$ where, as before, $\mathcal{U}$ is a set of exogenous variables, $\mathcal{V}$ is a set of endogenous variables, and $\mathcal{R}$ associates with every variable Y in $\mathcal{U} \cup \mathcal{V}$ a nonempty set $\mathcal{R}(Y)$ of possible values for Y; we extend $\mathcal{R}$ to subsets of $\mathcal{V}$ in the same way as before. However, we no longer require that $\mathcal{U}, \mathcal{V}$ or the sets $\mathcal{R}(Y)$ for $Y \in \mathcal{U} \cup \mathcal{V}$ be finite. The mapping $\mathbf{F}$ is a function $\mathbf{F} : \mathcal{I} \times \mathcal{R}(\mathcal{U}) \rightarrow \mathcal{P}(\mathcal{R}(\mathcal{V}))$, where $\mathcal{P}$ denotes the powerset operation. That is, it maps a context $\mathbf{u} \in \mathcal{R}(\mathcal{U})$ and an allowed intervention $I \in \mathcal{I}$ to a set of *outcomes* $\mathbf{F}(\mathbf{u}, I) \in \mathcal{P}(\mathcal{R}(\mathcal{V}))$. As with SEMs, we denote these outcomes by $M(\mathbf{u}, I)$. To rule out unintuitive possibilities, we require that outcomes $\mathbf{v} \in \mathbf{F}(\mathbf{u}, \vec{X} \leftarrow \vec{x})$ satisfy $\mathbf{v}[\vec{X}] = \vec{x}$. In the special case where all interventions are allowed, we take $\mathcal{I} = \mathcal{I}_{univ}$, the set of all interventions.

We note that, while these semantics are deterministic, we can bring probability back into the picture just as we do for SEMs: by putting a probability distribution on contexts. It is true that, since the set of contexts may be infinite, defining a distribution on it requires extra work (e.g, measurability assumptions). But this extra work is not specific to causal modeling; it is needed simply to get a probabilistic model of the scenario at hand. And, if that has already been accomplished, the probabilistic model can be reused for causal modeling. To keep things simple, we consider only deterministic examples in the rest of the paper.

We now make precise the sense in which GSEMs generalize SEMs. Two causal models M and M' are *equivalent*, denoted $M \equiv M'$, if they have the same signature and they have the same outcomes, that is, if for all sets of variables $\vec{X} \subseteq \mathcal{V}$, all values $\vec{x} \in \mathcal{R}(\vec{X})$ such that $\vec{X} \leftarrow \vec{x} \in \mathcal{I}$, and all contexts $\mathbf{u} \in \mathcal{R}(\mathcal{U})$, we have $M(\mathbf{u}, \vec{X} \leftarrow \vec{x}) = M'(\mathbf{u}, \vec{X} \leftarrow \vec{x})$.

Proposition 3.1. *For all SEMs M, there is a GSEM M' such that $M \equiv M'$.*

Just as for SEMs, the intervention $I = \vec{Y} \leftarrow \vec{y}$ on a GSEM M induces another GSEM M_I. To define M_I precisely, we must first define the composition of interventions.

Given interventions $\vec{X} \leftarrow \vec{x}$ and $\vec{Y} \leftarrow \vec{y}$, let their composition $I = \vec{X} \leftarrow \vec{x}; \vec{Y} \leftarrow \vec{y}$ be the intervention that results by letting the intervention performed second ($\vec{Y} \leftarrow \vec{y}$) override the first on variables that both interventions affect; that is, $I = \vec{X} \cup \vec{Y} \leftarrow \vec{z}$, where for $Z \in \vec{X} \cup \vec{Y}$,

$$\vec{z}[Z] = \begin{cases} \vec{y}[Z] & \text{if } Z \in \vec{Y}, \\ \vec{x}[Z] & \text{if } Z \in \vec{X} - \vec{Y}. \end{cases}$$

Given a GSEM $M = ((\mathcal{U}, \mathcal{V}, \mathcal{R}, \mathcal{I}), \mathbf{F})$ and an intervention $I \in \mathcal{I}$, define the intervened model M_I to be $((\mathcal{U}, \mathcal{V}, \mathcal{R}, \mathcal{J}), \mathbf{F}')$, where $\mathcal{J} = \{J \in \mathcal{I}_{univ} : I; J \in \mathcal{I}\}$ and, for $J \in \mathcal{J}$, $\mathbf{F}'(\mathbf{u}, J) = \mathbf{F}(\mathbf{u}, I; J)$. (The same relationship holds between the signatures $\mathcal{I}$ of M and $\mathcal{J}$ of M_I when M is a SEM.) Notice that if the set $\mathcal{I}$ is closed under composition, that is, if for all $I, J \in \mathcal{I}$ we have $I; J \in \mathcal{I}$, then $\mathcal{J} = \{J \in \mathcal{I}_{univ} : I; J \in \mathcal{I}\} \supseteq \mathcal{I}$, so that with M_I we have all the interventions that we had with M, and perhaps more.

The skeptical reader may wonder if the mechanism of equation modification in SEMs really is doing the same thing as the mechanism of intervention composition in GSEMs. This is indeed the case. There are two equivalent ways to see this. The first is to show that equation modification and intervention composition are the same for SEMs.

Proposition 3.2. *For all SEMs M and interventions $I, J \in \mathcal{I}$ such that $I; J \in \mathcal{I}$, we have that $M_I(\mathbf{u}, J) = M(\mathbf{u}, I; J)$.*

The second is to show that interventions respect equivalences that hold between SEMs and GSEMs.

Proposition 3.3. *If M is a SEM, M' is a GSEM, and $M \equiv M'$, then for all $I \in \mathcal{I}$, we have that $M_I \equiv M'_I$.*

3.2 Finite GSEMs

GSEMs clearly differ from SEMs in that the sets of endogenous and exogenous variables and the range of each individual variable can be infinite. Consider the class of GSEMs where these restrictions are retained, which we call *finite GSEMs*. How do finite GSEMs relate to SEMs? Halpern 2000 showed that all SEMs satisfy an axiom system called AX^+(see Appendix B for more details). For example, one axiom (effectiveness) states that after setting $X \leftarrow x$, all outcomes have $X = x$: $[\vec{W} \leftarrow \vec{w}; X \leftarrow x](X = x)$. While we imposed this constraint explicitly on GSEMs (and hence this axiom is *valid* in GSEMs—it is true in all contexts of all GSEMs), in SEMs there is no need to impose it; it is a property of the way outcomes are calculated. However, there are additional axioms, for example, one that requires unique outcomes if we intervene on all but one endogenous variable, that finite GSEMs do not satisfy. If we impose these axioms on finite GSEMs, we recover SEMs.

Theorem 3.4. *For all finite GSEMs over a signature $\mathcal{S}$ such that $\mathcal{I} = \mathcal{I}_{univ}$ in which all the axioms of AX^+ are valid, there is an equivalent SEM, and vice versa.*

Acyclic SEMs satisfy an axiom system called AX^+_{rec} (also described in Appendix B), which consists of the axioms in AX^+ along with two additional conditions. Imposing these axioms on finite GSEMs when all interventions are allowed gives us exactly the class of acyclic SEMs.

Theorem 3.5. *For all finite GSEMs over a signature $\mathcal{S}$ such that $\mathcal{I} = \mathcal{I}_{univ}$ and all the axioms AX^+_{rec} are valid, there is an equivalent acyclic SEM, and vice versa.*

We remark that the axiom system AX^+ can be generalized so as to deal with arbitrary (not necessarily finite) GSEMs, and soundness and completeness results for GSEMs can be proved. We defer these results to a companion paper [15].

Theorems 3.4 and 3.5 show that finite GSEMs satisfying AX^+ and AX^+_{rec}, respectively, are equivalent to SEMs and acyclic SEMs, respectively, *if all interventions are allowed*. This equivalence breaks down once we restrict the set of interventions; GSEMs are then strictly more expressive than SEMs, as shown by the following example.

Example 3.6. Suppose that Suzy is playing a shell game with two shells. One of the shells conceals a dollar; the other shell is empty. Suzy can choose to flip over a shell. If she does, the house flips over the other shell. If

Suzy picks shell 1, which hides the dollar, she wins the dollar; otherwise she wins nothing. This example can be modeled by a GSEM M_{shell} with two binary endogenous variables S_1, S_2 describing whether shell 1 is flipped over and shell 2 is flipped over, respectively, and a binary endogenous variable Z describing the change in Suzy's winnings. (The GSEM also has a trivial exogenous variable whose range has size 1, so that there is only one context $\mathbf{u}$.) That defines $\mathcal{U}, \mathcal{V}$, and $\mathcal{R}$; we set $\mathcal{I} = \{S_1 \leftarrow 1, \mathcal{S}_2 \leftarrow 1\}$; and $\mathbf{F}$ is defined as follows:

$$\begin{aligned} \mathbf{F}(\mathbf{u}, S_1 \leftarrow 1) &= M_{shell}(\mathbf{u}, S_1 \leftarrow 1) \\ &= \{(S_1 = 1, S_2 = 1, Z = 1)\} \\ \mathbf{F}(\mathbf{u}, S_2 \leftarrow 1) &= M_{shell}(\mathbf{u}, S_2 \leftarrow 1) \\ &= \{(S_1 = 1, S_2 = 1, Z = 0)\}. \end{aligned}$$

M_{shell} is clearly a valid GSEM. Furthermore, checking that M_{shell} satisfies all the axioms in AX^+ is straightforward; see Appendix C (Theorem C.2) for details. However, no SEM M' with the same signature can have the outcomes $M'(\mathbf{u}, S_1 \leftarrow 1) = \{(S_1 = 1, S_2 = 1, Z = 1)\}$ and $M'(\mathbf{u}, S_2 \leftarrow 1) = \{(S_1 = 1, S_2 = 1, Z = 0)\}$. This is because in a SEM, the value of Z would be specified by a structural equation $Z = \mathcal{F}_Z(\mathcal{U}, S_1, S_2)$. This cannot be the case here, since there are two outcomes having $S_1 = S_2 = 1$, but with different values of Z. ∎

This example shows that finite GSEMs (even restricted to those satisfying the axioms of AX^+) are more expressive than SEMs when not all interventions are allowed. The fundamental issue here is that Z is determined by the intervention (which shell Suzy picks), not the state of the shells. In SEMs, the system's behavior cannot depend explicitly on the intervention, only on the variables altered by the intervention. We note that Suzy's situation can be modeled by a SEM with an additional variable describing Suzy's action. (More precisely, one where the only allowed interventions set this variable's value to match Suzy's action.) However, this variable is redundant in the sense that Suzy's action is already described by the intervention. Thus, arguably, the GSEM is the more natural model for this situation.

4 Actual Causes

As we said, the fact that GSEMs have the same input-ouput semantics as SEMs makes it easy to lift the definitions of notions like actual causality, explanation, responsibility, and blame from SEMs to GSEMs, and that doing so gives reasonable results. We illustrate this point with actual causality.

One important application of causal modeling is to deducing the *actual cause(s)* of $X = x$, that is (roughly speaking) the specific reasons that X takes value x in a given context $\mathbf{u}$ and outcome $\mathbf{v}$. Many definitions of actual cause in SEMs have been proposed; for definiteness, we use that of Halpern and Pearl 2005, as later modified by Halpern 2016, except we require the intervention $\vec{W} \leftarrow \mathbf{v}[\vec{W}]; \vec{X} \leftarrow \vec{x}$ appearing in AC2 below to be allowed (previous definitions did not consider allowed interventions). We can use this definition without change in GSEMs.

Definition 4.1 (Actual cause). Given a causal model M, $\mathbf{u} \in \mathcal{R}(\mathcal{U})$, and $\mathbf{v} \in \mathcal{R}(\mathcal{V})$, $\vec{X} = \vec{x}$ is an *actual cause* of the event φ in $(M, \mathbf{u}, \mathbf{v})$ if the following three conditions hold:

AC1. $\mathbf{v} \models \vec{X} = \vec{x}$ and $\mathbf{v} \models \varphi$.
AC2. There is some $\vec{W} \subseteq \mathcal{V}$ and a setting $\vec{x}'$ of $\vec{X}$ such that
$\vec{W} \leftarrow \mathbf{v}[\vec{W}]; \vec{X} \leftarrow \vec{x}' \in \mathcal{I}$ and
$(M, \mathbf{u}) \models \langle \vec{W} \leftarrow \mathbf{v}[\vec{W}]; \vec{X} \leftarrow \vec{x}' \rangle \neg\varphi$.
AC3. No proper subset of $\vec{X}$ satisfies conditions AC1 and AC2.

Intuitively, this definition captures the fact that when reasoning about counterfactuals in a concrete scenario, we often want to fix some details $\vec{W}$ to the values $\mathbf{v}[\vec{W}]$ that they actually had in that scenario; see [7]. In the next section, we use GSEMs to apply this definition in the dynamical-systems setting.

5 Using GSEMs to Capture Dynamical-Systems Formalisms

In this section, we show how GSEMs can be used to capture various formalisms for modeling dynamical systems.

5.1 Ordinary Differential Equations

Suppose that we have a system of ordinary differential equations (ODEs) of the form

$$\begin{aligned}\dot{\mathcal{X}}_1 &= F_1(\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n)\\ \dot{\mathcal{X}}_2 &= F_2(\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n)\\ &\ldots\\ \dot{\mathcal{X}}_n &= F_n(\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n),\end{aligned}$$

where the $\mathcal{X}_i$ are real-valued functions of time, called *dynamical variables*, and $\dot{\mathcal{X}}_i$ denotes the derivative of $\mathcal{X}_i$ with respect to time.[3] Together with the initial values $\mathcal{X}_1(0), \mathcal{X}_2(0), \ldots, \mathcal{X}_n(0)$, the equations determine a set of solutions over the interval $[0, T]$ for $T > 0$ or the interval $[0, \infty)$. This system of ODEs has a natural causal semantics; namely, (1) intervening on a dynamical variable at a given instant in time does not change the equations that govern the system's evolution, and (2), intervening on a variable's trajectory on an interval does not change the equations governing the other variables. In this section, we implement this semantics with a GSEM M_{ODE}.

A GSEM Model for Systems of ODEs. The GSEM M_{ODE} consists of variable sets $\mathcal{U}, \mathcal{V}$, their ranges $\mathcal{R}$, the allowed interventions $\mathcal{I}$, and the mapping $\mathbf{F}$. The exogenous variables $\mathcal{U} := \{X_1^0, \ldots, X_n^0\}$ correspond to initial conditions (here, the superscript represents time). The endogenous variables $\mathcal{V} := \{X_i^s : 1 \leq i \leq n, s \in (0, T]\}$ are indexed by time and describe the evolution of the system. All variables range over the reals; that is, $\mathcal{R}(V) = \mathbb{R}$ for all $V \in \mathcal{U} \cup \mathcal{V}$. The natural causal semantics described above deals with point interventions $X_i^t \leftarrow k$, and trajectory interventions $\{X_i^t \leftarrow f(t) \mid t \in (a, b)\}$, where f is a smooth function (n times differentiable). We will abbreviate the latter as $X_i(a, b) \leftarrow f$, and if f is the constant function that outputs $k \in \mathbb{R}$, we will also write $X_i(a, b) \leftarrow k$. The class of interventions $\mathcal{I} = \mathcal{I}_{intervals}$ we will handle consists of all finite compositions $I_1; I_2; \ldots; I_m$, where each I_j is either a point intervention $X_i^t \leftarrow k$ or a trajectory intervention $X_i(a, b) \leftarrow f$. Notice that this class also contains intervals on half-closed or closed intervals; we will denote such interventions in the natural way, for example, $X_i[a, b) \leftarrow f$.

It remains to describe the mapping $\mathbf{F}$ of M_{ODE}. Before we do, we need some preliminary definitions. A dynamical variable $\mathcal{X}_i$ is *intervention-free with respect to* $I = \vec{X} \leftarrow \vec{x}$ *on an interval* (a, b) if for all $t \in (a, b)$, we have $X_i^t \notin \vec{X}$; the interval (a, b) is *intervention-free* if all dynamical variables are intervention-free on (a, b). Given a smooth function $f(t)$, a variable is *set to* f during an open interval (a, b) if for all $t \in (a, b)$, $\vec{x}[X_i^t] = f(t)$. An open interval (a, b) is *intervention-constant* if for all $1 \leq i \leq m$, $\mathcal{X}_i$ is either intervention-free on (a, b) or set to f (for some smooth function f) during (a, b). Given a context $\mathbf{u}$ and an intervention $I = \vec{X} \leftarrow \vec{x}$, let $\mathbf{F}(\mathbf{u}, I)$ consist of all outcomes $\mathbf{v}$ that satisfy the following conditions (where $\mathcal{X}_i(t) = \mathbf{v}[X_i^t]$ for $1 \leq i \leq n$).

ODE1. The outcome $\mathbf{v}$ agrees with I, that is, $\mathbf{v}[\vec{X}] = \vec{x}$.

[3] Nearly all systems of ODEs occurring in practice can be put into this form by adding auxiliary variables [17]. For example, $d^2\mathcal{X}/dt^2 = -\mathcal{X}$ becomes the pair of equations $d\mathcal{X}/dt = \mathcal{Y}; d\mathcal{Y}/dt = -\mathcal{X}$.

ODE2. For all i, $\mathcal{X}_i$ is left-continuous except when intervened on; that is, $\mathcal{X}_i$ is left-continuous at all points t such that $X_i^t \notin \vec{X}$.

ODE3. For every intervention-constant open interval (a, b), if $\mathcal{X}_i$ is intervention-free on (a, b), then $\mathcal{X}_i$ is right-continuous at a, differentiable on (a, b), and its derivative $\dot{\mathcal{X}}_i$ satisfies $\dot{\mathcal{X}}_i(t) = F_i(\mathcal{X}_1, ..., \mathcal{X}_m)(t)$ for all $t \in (a, b)$.

These conditions require, rather straightforwardly, that the outcomes, when not intervened on, are continuous and obey the governing differential equations. In Appendix A, we give an algorithm that computes the outcomes of M_{ODE} by solving the underlying differential equations, further reinforcing the view that the outcomes of M_{ODE} correspond to the natural causal semantics of the underlying equations.

Example: an LC Circuit Model. We conclude this section by showing how a textbook dynamical system—an LC circuit—can be modeled as a GSEM. An LC circuit consists of a voltage source, a capacitor, and an inductor. The dynamical variable of interest is the charge $Q(t)$ on the capacitor; the voltage V, capacitance C, and inductance L are fixed parameters (although we encode them as dynamical variables with zero derivatives). The differential equations governing this circuit's behavior are

$$\dot{\mathcal{Q}}(t) = \mathcal{K}(t)$$
$$\dot{\mathcal{K}}(t) = \frac{\mathcal{V}}{\mathcal{L}} - \frac{1}{\mathcal{L}\mathcal{C}}\mathcal{Q}(t)$$
$$\dot{\mathcal{V}}(t) = \dot{\mathcal{C}}(t) = \dot{\mathcal{L}}(t) = 0,$$

where $\mathcal{K}$ is the current. The solutions of these differential equations (for $\mathcal{Q}(t)$) take the form

$$\mathcal{Q}(t) = \mathcal{V}\mathcal{C} + A\cos(\omega t + B),$$

where $\omega = 1/\sqrt{\mathcal{L}\mathcal{C}}$, and A and B are determined by the initial conditions on $\mathcal{Q}$ and $\mathcal{K}$: $B = \arctan(\frac{\mathcal{K}(0)}{\omega(\mathcal{V}\mathcal{C} - \mathcal{Q}(0))})$ and $A = \frac{\mathcal{Q}(0) - \mathcal{V}\mathcal{C}}{\cos(B)}$. Note that these expressions make sense for all initial conditions except when $\mathcal{V}\mathcal{C} - \mathcal{Q}(0) = 0$; in this case the solution is instead $\mathcal{Q}(t) = \mathcal{V}\mathcal{C}$ for all t.

It follows from the explicit forms of the solutions given above that all initial-value problems involving these differential equations have unique solutions. It is similarly easy to see that if any of the differential equations (for variable $\mathcal{X}$) is replaced with $\dot{\mathcal{X}} = 0$, all initial-value problems involving the resulting modified system of differential equations also have unique solutions. Using this, it is not too hard to see that for all contexts $\mathbf{u}$ and $I \in \mathcal{I}_{intervals}$, there is a unique outcome $\mathbf{v} \in M(\mathbf{u}, I)$. Suppose that the initial conditions of the circuit are given by the context $\mathbf{u} = \{Q^0 = 0, K^0 = 0, V^0 = 2, C^0 = 2, L^0 = 2\}$. The unique outcome $\mathbf{v} \in M(\mathbf{u}, \emptyset)$ under the empty intervention is shown in Figure 1 (blue curve).

Using the LC Circuit Model to Determine Actual Causes. Suppose that the capacitor breaks down if $\mathcal{Q}$ exceeds 6. Figure 1 shows that in the absence of intervention, $\mathcal{Q}$ exceeds 6 just after $t = 4$. To prevent this, the operators ground the capacitor at time 4 (i.e., make the intervention $Q^4 \leftarrow 0$). However, this does not help. As shown in Figure 1 (orange curve), this initially reduces $\mathcal{Q}$, but eventually it exceeds 6. Next, the operators try opening the circuit at time 3 and closing it again at time 4; that is, performing the intervention $\{K^t \mid t \in (3, 4)\} \leftarrow 0$. This results in the circuit entering an operating regime where $\mathcal{Q}$ is nearly constant (the green curve in Figure 1), and never exceeds 6.

We can show that, in agreement with intuition, the first intervention is not an actual cause of the capacitor not breaking down, and neither is the second (because it is not the minimal intervention required to bring about the outcome). But a sub-intervention $K^4 \leftarrow 0$ of the second intervention is an actual cause (corresponding to the red curve in Figure 1). Indeed, recall that the capacitor breaks down if $\mathcal{Q}$ exceeds 6, so the capacitor not

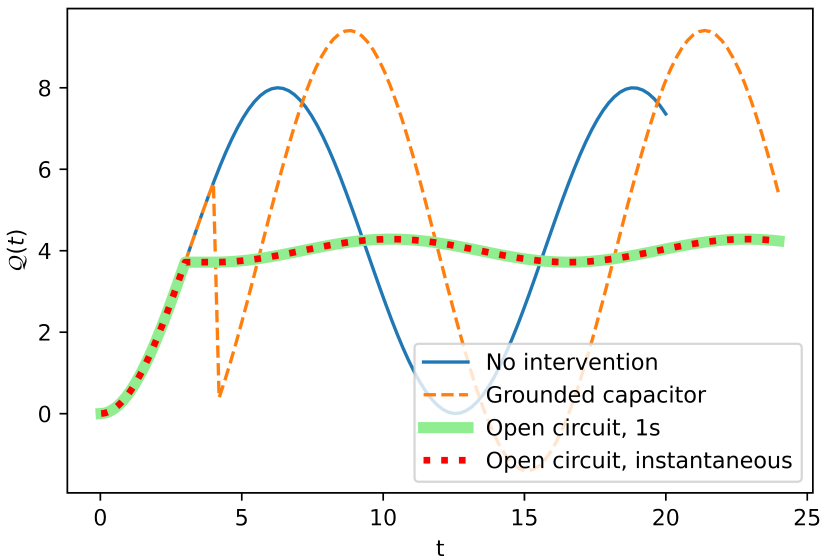

Fig. 1. LC circuit outcomes under different interventions.

breaking down corresponds to the statement $\varphi = \forall t > 4,\ Q(t) \leq 6$.[4] It is obvious that the first intervention is not an actual cause of the capacitor breaking down, because under the first intervention, intervention, Q^t does exceed 6 (at, say, $t = 10$). Thus $\mathbf{v} \models \neg\varphi$, where $\mathbf{v}$ is the outcome under $Q^t \leftarrow 0$, which violates AC1. To argue the claim that $K^4 \leftarrow 0$ *is* an actual cause of the capacitor breaking down, note that on intervention-free intervals, the solutions are periodic in time, with period $2\pi/\omega = 2\pi\sqrt{L^0C^0} = 4\pi \approx 12.6$. Since Q does not exceed 6 on the interval $(4, 4 + 12.6)$, it will never exceed 6. Thus $\mathbf{v} \models \varphi$, where $\mathbf{v}$ is the outcome under $K^4 \leftarrow 0$, satisfying AC1. AC2 is satisfied, because if we instead set $K^t \leftarrow a_t$ for $t \in (3, 4)$ where a_t is the value K^t had under the empty intervention, then Q^t does exceed 6 (at, say, $t = 6$). (Here we are taking $\vec{W} = \emptyset$.) Finally, AC3 is satisfied, because the intervention $K^4 \leftarrow 0$ is on a single variable, and therefore minimal.

The fact that we can apply the definition of causality without change in this way emphasizes the benefits of using GSEMs. We can similarly apply the definitions of explanation [11, 8] and responsibility and blame [5] without change. Essentially, any definition that is couched in terms of interventions and outcomes can be lifted to GSEMs.

5.2 Rule-Based Models

A rule-based model is a dynamical system that transitions probabilistically between states, with the transition defined by rewrite rules. In this section, we construct a GSEM corresponding to the generic rule-based model given by Laurent et al. 2018.

[4] This statement φ, since it is universally quantified over all $t > 4$, cannot be expressed in the language of causal formulas $\mathcal{L}(\mathcal{S})$ we defined in Section 2. However, we do not view this as a serious problem. The definition of actual causality (Definition 4.1) allows arbitrary φ, and it still makes sense if we take φ to be a formula in a richer language containing the first-order quantification we need here.

An Example of a Rule-Based Model. Laurent et al. 2018 show how a rule-based model can describe a reaction between a set S of *substrates* and a set K of *kinases*. The state of the mixture at time t is a binary relation $\mathsf{Bound} \subseteq S \times K$ that specifies which substrates are bound to which kinases, and a unary relation $\mathsf{Phos} \subseteq S \cup K$ that defines which substrates and kinases have a phosphate group attached. Chemical interactions between groups of molecules are intended to take place spontaneously, in an analogous fashion to radioactive decay. For example, if at time t there is a substrate $s \in S$ and a kinase $k \in K$ such that $(s,k) \in \mathsf{Bound}$ and $s \notin \mathsf{Phos}$, then at time $t + \Delta t$, where Δt is drawn from an exponential distribution with a time constant that depends on the rule being applied, s will gain a phosphate group (unless in the meantime some other rule has changed the state of the mixture so that the precondition for s and k no longer holds). These updates are called *events*. Interventions correspond to blocking some interactions from taking place at specific times, for example, "between $t = 1$ and $t = 2$, even if s and k satisfy the above conditions, s cannot gain a phosphate group."

Laurent et al. explain how to simulate these dynamics using the following algorithm. For every possible target Targ of a rule r—in the example above, this would be every substrate-kinase pair (s,k)—sample from a Poisson process with parameter $\tau/\ln(2)$ to obtain a schedule of times when the rule applies to this target. Then, starting with the initial mixture and moving through time, whenever a rule applies according to the schedule, check if the rule's condition—in our example $(s,k) \in \mathsf{Bound} \land s \notin \mathsf{Phos}$—is satisfied for the target, and that the rule is not currently blocked by an intervention. If so, update the mixture using the rule's mapping (e.g. $\mathsf{Phos} \mapsto \mathsf{Phos} \cup \{s\}$); otherwise, do nothing.

A Rule-Based GSEM. . This algorithm can immediately be described in a GSEM model. In the example above, for each time $t \in [0,\infty)$ we would have binary endogenous variables $\mathsf{Bound}^t_{(s,k)}$ for each $s \in S, k \in K$, along with variables Phos^t_x for each $x \in S \cup K$. The exogenous variables correspond to the firing schedule; we have timestamped variables $T^t_{r,\mathsf{Targ}}$, one for each rule r, each target Targ compatible with that rule. (There are also exogenous variables describing the initial state of the mixture.) In order to match the intervention model of [14], we add additional binary variables $B^t_{r,\mathsf{Targ}}$. Intuitively, $B^t_{r,\mathsf{Targ}} = 1$ means that the firing of rule r applied to target Targ at time t is blocked. Finally, for bookkeeping, we have binary endogenous variables of the form $X^t_{r,\mathsf{Targ}}$ that model whether rule r actually fired on target Targ at time t. The unique outcome is specified in the obvious way: $X^t_{r,\mathsf{Targ}}$ is true exactly if, at time t, Targ satisfies the condition of r, $T^t_{r,\mathsf{Targ}}$ is true, and $B^t_{r,\mathsf{Targ}}$ is false. If $X^t_{r,\mathsf{Targ}}$ is true, then at time t the state (i.e., the relations Bound and Phos) gets updated using the rule's mapping. Interventions such as the one above can simply be described by setting some of the $B^t_{r,\mathsf{Targ}}$ to false; we take the set of allowed interventions $\mathcal{I}$ to be all interventions of this form. The *trace* $T(\mathbf{u}, I)$ is simply the (countable) sequence of variables $X^t_{r,\mathsf{Targ}}$ (in ascending order of t) for which $X^t_{r,\mathsf{Targ}} = 1$.

The Analysis of Laurent et al. Laurent et al. 2018 defined notions of *enablement* and *prevention*. Enablement and prevention happen at the level of *events*, or updates to the mixture. Every event e corresponds to a variable $X^t_{r,\mathsf{Targ}}$; it *occurs* if $X^t_{r,\mathsf{Targ}} = 1$. We can think of the relations Bound and Phos as binary vectors; each entry in these vectors is called a *site*. For any given event to occur, certain sites must have certain values. Hence, intuitively, given two events e, e', e enables e' if e is the last event before e' that modifies some site to the value that is needed for e to occur. Likewise, e prevents e' (roughly) if e is the last event before e' to set a site s, and it sets s to a value such that e cannot occur. Given a context $\mathbf{u}$ and an intervention $I \in \mathcal{I}$, they considered the difference between the trace $T(\mathbf{u})$ and the trace $T(\mathbf{u}, I)$. They showed that for every element of the first sequence absent from the intervened sequence, a chain of enablements and preventions could be traced back from that element to an element that was directly blocked by I. That is, enablements and preventions were sufficient to explain why each element of $T(\mathbf{u})$ no longer in $T(\mathbf{u}, I)$ was missing.

A Complementary Analysis Utilizing the Rule-Based GSEM. . Thinking in terms of actual cause (see Section 4) complements this analysis. For example, if one rule firing is the actual cause of another rule firing, then a chain of enablements and preventions can be traced back from the trace entry for the second rule to the trace entry for the first. More precisely, if $B^t_{r,\mathrm{Targ}} = 0$ is an actual cause of $X^{t'}_{r',\mathrm{Targ}'} = 1$ in context $\mathbf{u}$, then a chain of enablements and preventions can be traced back from $X^{t'}_{r',\mathrm{Targ}'}$ to $X^t_{r,\mathrm{Targ}}$ in the pair of traces $T(\mathbf{u})$, $T(\mathbf{u}, B^t_{r,\mathrm{Targ}} \leftarrow 1)$. Without going into the formalism of [14], a sketch of the proof of this claim is as follows. The actual cause statement implies that $X^{t'}_{r',\mathrm{Targ}'}$ is in $T(\mathbf{u})$ but not in $T(\mathbf{u}, B^t_{r,\mathrm{Targ}} \leftarrow 1)$, because the intervention $B^t_{r,\mathrm{Targ}} \leftarrow 1$ is the only one that can satisfy AC2 and AC3. (The other blocking variables take value zero in both outcomes $M(\mathbf{u}, \emptyset)$ and $M(\mathbf{u}, B^t_{r,\mathrm{Targ}} \leftarrow 1)$, so setting them to 0 is redundant and violates AC3.) The only element blocked by $B^t_{r,\mathrm{Targ}} \leftarrow 1$ is $X^t_{r,\mathrm{Targ}}$. Hence, a chain of enablements and preventions can be traced back from $X^{t'}_{r',\mathrm{Targ}'}$ to $X^t_{r,\mathrm{Targ}'}$.

The GSEM machinery can also be used to answer questions not addressed by the analysis of Laurent et al. We can ask counterfactual questions like "What would the state at $t = 7$ be if every kinase gained a phosphate group at time $t = 5$?" (potentially corresponding to the addition of a test tube's worth of phosphate solution) or "Is the fact that substrate s was bound to kinase k at $t = 1$ the actual cause of kinase k gaining a phosphate group at $t = 2$?" For this reason, we believe that GSEMs are a useful addition to the rule-based causal modeling toolkit developed by Laurent et al.

5.3 Hybrid Automata

Hybrid automata are a well-developed class of models for systems that have both continuous and discrete components [1]; for example, a thermostat controlling a heater to keep the temperature within a certain tolerance of a set point. The state variables for this system are both continuously varying in time (the temperature) and discretely changing in time (whether the heater is on). In this section, we show how to construct a GSEM model corresponding to an arbitrary hybrid automaton. We demonstrate our construction on a simple example from [12] and show how the resulting GSEM can be used to answer causal questions.

Definition. Mathematically, a hybrid automaton is a finite directed multigraph $G = (V, E)$, a set $\boldsymbol{\mathcal{X}} = \{\mathcal{X}_1, \ldots, \mathcal{X}_n\}$ of real-valued *dynamical variables*,[5] and some predicates (discussed below).[6] The states $v \in V$ are called *control modes*, and they represent the state of the discrete component of the system. The edges $e = (u, v, n) \in E$, where u and v are control modes and n indexes the edge among all edges from u to v, are called *control switches*, and they describe transitions between control modes. Recall that a multigraph is a graph which may have multiple edges between any given pair of nodes; different control switches between the same pair of control modes represent different modes of transition between them. For example, the heater may have multiple triggers that change its state from OFF to ON.

The semantics of a hybrid automaton is defined by predicates init, flow, jump, and inv. Possible starting configurations are given by $\mathsf{init}(v, \boldsymbol{\mathcal{X}})$. Possible continuous dynamics within a control mode are given by $\mathsf{flow}(v, \boldsymbol{\mathcal{X}}, \dot{\boldsymbol{\mathcal{X}}})$, where $\dot{\boldsymbol{\mathcal{X}}}$ represents the vector of first derivatives. Possibly discontinuous changes are given by $\mathsf{jump}(e, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{X}}')$, where $\boldsymbol{\mathcal{X}}'$ represents values at the conclusion of the change. Hard constraints are represented by $\mathsf{inv}(v, \boldsymbol{\mathcal{X}})$, which may be thought of as describing invariants of the different control modes. Hybrid automata are nondeterministic in general; all dynamics compatible with the automaton's predicates are possible.

A GSEM Hybrid Automaton. Let $A = (V, E, \boldsymbol{\mathcal{X}}, \mathsf{init}, \mathsf{flow}, \mathsf{jump}, \mathsf{inv})$ be a hybrid automaton. We now construct a GSEM M corresponding to A. The endogenous variables are as follows. For each $\mathcal{X}_i \in \boldsymbol{\mathcal{X}}$, M has real-valued variables $\{X^t_i \mid t \geq 0\}$ corresponding to the value of $\mathcal{X}_i$ at time t. M also has V-valued variables $\{S_t \mid t \geq 0\}$

[5] In the literature, these are usually just called variables; we call them dynamical variables to avoid confusing them with GSEM variables.
[6] There is also a set of events Σ used to disambiguate discrete transitions, but this is not relevant for us.

corresponding to the control mode of the system at time t. M has a single exogenous variable with a single value, so there is only one (trivial) context.

The outcomes for intervention $\vec{Y} \leftarrow \vec{y}$ are, similar to ODEs, all assignments to the variables that (1), agree with $\vec{Y} \leftarrow \vec{y}$, and (2), are otherwise compatible with the predicates of the automaton. More precisely, $\mathbf{v} \in M(\mathbf{u}, \vec{Y} \leftarrow \vec{y})$ iff all the following conditions hold. For convenience, we define functions $\mathcal{X}_i(t) = \mathbf{v}[X_i^t]$ for $i = 1, \ldots, n$, $\boldsymbol{\mathcal{X}}(t) = (\mathcal{X}_1(t), \ldots, \mathcal{X}_n(t))$, and $\mathcal{D}(t) = \mathbf{v}[S_t]$.

HA1. $\mathbf{v}[\vec{Y}] = \vec{y}$.
HA2. $\mathsf{init}(\mathbf{v}[S_t], \boldsymbol{\mathcal{X}}(0))$ holds.
HA3. For all t, $\mathsf{inv}(\mathbf{v}[S_0], \boldsymbol{\mathcal{X}}(t))$ holds.
HA4. For all $t \geq 0$, if $\boldsymbol{\mathcal{X}}(t)$ is not continuous, at least one of (a) or (c) below holds; and if $\mathcal{D}(t)$ is not continuous, at least one of (b) or (c) below holds.
(a) $X_i^t \in \vec{Y}$ for some $i = 1, \ldots, n$.
(b) $S_t \in \vec{Y}$.
(c) There is an edge $e = (u, v, n) \in E$ such that $\mathsf{jump}(e, \lim_{s \to t^-} \boldsymbol{\mathcal{X}}(t), \boldsymbol{\mathcal{X}}(t))$ holds.
HA5. Defining *intervention-free* intervals in the same way as in Section 5, the following holds. For all intervention-free intervals (a, b) such that $\boldsymbol{\mathcal{X}}$ is continuous on (a, b), we have that $\boldsymbol{\mathcal{X}}$ is right-continuous at a, differentiable on (a, b), and its derivative $\dot{\boldsymbol{\mathcal{X}}}$ satisfies the flow condition $\mathsf{flow}(\mathbf{v}[S_t], \boldsymbol{\mathcal{X}}(t), \dot{\boldsymbol{\mathcal{X}}}(t))$ for all $t \in (a, b)$.

Note that HA5 is analogous to the condition ODE3 in Section 5.[7] The modeler is free to select a set of allowed interventions that fits the task at hand. In the example below, we choose $\mathcal{I}$ to be the set of finite compositions of point and interval interventions on the dynamical variables and control mode variables, but the outcomes of M are well-defined for arbitrary interventions. This concludes the construction of M.

A Thermostat Example. A simple thermostat and heater automaton given by [12] (with the flow conditions slightly simplified) is as follows. There is a single continuous variable $\mathcal{T} \in \boldsymbol{\mathcal{X}}$ and two control modes $V = \{OFF, ON\}$. The system starts in OFF with $\mathcal{T} = 20$, which is the desired set point. This defines $\mathsf{init}(v, \boldsymbol{\mathcal{X}})$. In OFF, the temperature drifts slowly downward; we have $\dot{\mathcal{T}} = -0.1$. An invariant of OFF is that $\mathcal{T} \geq 18$; if at any point $\mathcal{T} < 18$, the system transitions to ON. Likewise, in ON, the temperature increases rapidly as $\dot{\mathcal{T}} = 0.5$, and an invariant of ON is $\mathcal{T} \leq 22$. These conditions define $\mathsf{flow}(v, \boldsymbol{\mathcal{X}}, \dot{\boldsymbol{\mathcal{X}}})$ and $\mathsf{inv}(v, \boldsymbol{\mathcal{X}})$. There are two control switches, e_{on} from OFF to ON, and e_{off} from ON to OFF. Finally, the heater can transition from OFF to ON when $\mathcal{T} < 19$, and from ON to OFF when $\mathcal{T} > 21$. Neither of these transitions affects the instantaneous value of $\mathcal{T}$; that is, $\mathcal{T}' = \mathcal{T}$. These conditions define $\mathsf{jump}(e, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{X}}')$.

Let us analyze the GSEM M corresponding to this automaton. In our case, since neither of the two possible discrete jumps change the value of $\mathcal{T}$, HA4 implies that $\mathcal{T}$ must be continuous everywhere it is not intervened on. Interventions on $\mathcal{T}$ are finite compositions of point and interval interventions. Furthermore, when $\mathcal{T}$ is not intervened on, it cannot change very quickly; by HA5, it must obey the flow condition (either $\dot{\mathcal{T}} = -0.1$ or $\dot{\mathcal{T}} = 0.5$). Hence, given a time horizon τ, $\mathcal{T}$ can cross the vertical lines $\mathcal{T} = 19$ and $\mathcal{T} = 21$ only finitely many times prior to τ. The state of the heater is specified by the control mode S^t. By HA4, the heater switches ON only at times t when the state of the heater (i.e., S^t) is intervened on, or when $\mathcal{T} < 19$; likewise, the heater switches OFF only when S^t is intervened on, or when $\mathcal{T} > 21$. Again, interventions on the state of the heater are finite compositions of point and interval interventions. It follows that the state of the heater changes only a finite

[7] It is possible (and probably desirable) to strengthen HA5 analogously to the way we strengthened ODE3 to ODE3′, so that interval interventions on one dynamical variable do not interfere with the flow conditions on another dynamical variable. We do not do this here, because it is not necessary for our simple example (which has only a single variable), and because doing this requires some knowledge of the structure of the predicate $\mathsf{flow}(v, \boldsymbol{\mathcal{X}}, \dot{\boldsymbol{\mathcal{X}}})$.

number of times before any given time horizon τ. Hence, it is meaningful to talk about the heater discretely changing state—before any given time τ, the heater turns on at t_1, t_3 and so on, and turns off at t_2, t_4 and so on.

Answering Questions of Actual Cause with the Hybrid Automaton GSEM. . Now that we have some intuition for the behavior of M, we examine how M can be used to answer questions of actual cause. By HA2, the heater is initially OFF, and the temperature is initially $\mathcal{T} = 20$. In the absence of intervention or discrete jumps, the heater will stay OFF and the temperature will drop at the rate of 0.1 per unit time.

Consider an outcome $\mathbf{v}$ where the heater does not turn on until, at $t = \frac{18-20}{-0.1} = 20$, it is required to do so by HA3; specifically, by the invariant of OFF that $\mathcal{T} \geq 18$. If the heater had been on over any open subinterval (a, b) of $[0, 20)$, the temperature would have been higher than 18 by $t = 20$ by at least $0.5(b - a)$. Hence, intuitively, the heater being off over any such subinterval should be considered a cause of $T^{20} = 18$. However, if we fix any subinterval $(a, b) \subset [0, 20)$ and ask the formal question of whether $S(a, b) = \mathit{OFF}$ is an actual cause of $T^{20} = 18$ in $(M, \mathbf{u}, \mathbf{v})$ (where $S(a, b) = \{S^t \mid t \in (a, b)\}$), we run into problems.[8]

AC1 and AC2 both hold, but AC3 does not. AC1 holds, because $\mathbf{v}[S^t] = \mathit{OFF}$ for all $t \in [0, 20)$, and $\mathbf{v}[T^{20}] = 18$. AC2 holds, since if we choose $\vec{W} = \emptyset$ and $\vec{x}' = \mathit{ON}$ (recall that $\vec{X} = S(a, b)$ we find that the outcome $\mathbf{v}'$ of M under intervention $\vec{X} \leftarrow \vec{x}'$ where the heater is on only during (a, b) has $\mathbf{v}'[T^{20}] = 18 + 0.5(b - a) \neq 18$. However, AC3 does not hold, because the open subinterval (a, b) contains other open subintervals for which AC1 and AC2 also hold, by the same arguments. This implies that there is no open interval on which the heater being off is an actual cause of $T^{20} = 18$.

This creates a dilemma. Since turning the heater on results in $T^{20} > 18$, a good definition of actual cause should provide for some cause. In this case, the resolution is that the equality $S^t = \mathit{OFF}$ for any *point* $t \in (0, 20)$ is in fact an actual cause of $T^{20} = 18$. This is because one of the solutions to $S^t \leftarrow 1$ has $S^r = \mathit{ON}$ for all r in some nonempty interval starting at t, which as before implies $T^{20} > 18$. So AC2 holds. AC1 clearly holds, and AC3 holds because $S^t \leftarrow 1$ is a point intervention, therefore minimal. We believe this resolution to the dilemma is always possible in hybrid automata (if point interventions are allowed), since we see no way of defining a hybrid automaton such that when the control mode is intervened on at a point in time, the control mode does not remain at the intervened value for some nonzero amount of time in some solution (although we have not attempted to prove this).

However, we see no reason for this resolution to work in general. Other models of dynamical systems may not respond in the same way to intervention. This resolution even fails for our example hybrid automaton, if it is modified so that point interventions are not allowed. In these cases, the definition of actual causality presented in Section 4 fails to provide for any cause of $T^{20} > 18$ involving only the heater state. The issue is that AC3 requires a minimal cause; but minimal causes do not exist in general when causes can involve infinitely many variables. It is an open problem to find a new definition of actual cause that handles infinitely many variables well in general. One potential solution is to broaden the set of things that count as causes in infinitary settings. In the definition presented in Section 4, only conjunctions of equalities $X = x$ can be actual causes. One could consider expanding possible causes to include infinite disjunctions over these equalities, for example, the statement that there is *some* nonempty interval on which the heater is off:

$$\exists (a, b) \forall t \in (a, b) S^t = \mathit{OFF}.$$

However, we do not pursue this approach further in this paper.

Note that if we ask instead whether $S^t(a, b) = \mathit{OFF}$ is an actual cause of $T^{20} \leq 19$ instead, there are no problems. The answer is yes, iff $b - a = 2$. This agrees with the natural intuition that the heater being off for a sufficiently

[8]Similar to Footnote 4, there is a technical issue here, because the event $S(a, b) = \mathit{OFF}$ (which is an infinite conjunction of equalities) is not in the language $\mathcal{L}(\mathcal{S})$, even though the intervention $S(a, b) \leftarrow \mathit{OFF}$ is. However, again, we do not view this as a problem, since the definition of actual causality makes perfect sense for this formula.

short time is not enough to cause the temperature to be low. (There is nothing special about 19; the solution for any value $x > 18$ is similar.)

6 Comparison with Alternative Modeling Approaches

In this section we briefly compare GSEMs to several modeling approaches that have related objectives.

6.1 Monotone Simulation Models

Ibeling and Icard 2019 consider SEMs with infinitely many variables and a second class of models that they call *monotone simulation models*. They show that monotone simulation models are equivalent to *computable* SEMs (ones where, roughly speaking, the structural equations are computable functions). Our examples showing that GSEMs are more general than SEMs apply immediately to the models considered by Icard and Ibeling.

6.2 Causal Constraints Models

At a coarse-grained level of modeling, such as when describing equilibrium solutions of dynamical systems, constraints between variables are natural objects of study. In order to describe equilibrium solutions and functional laws (roughly, dependencies that cannot be violated via intervention), [3] introduced the notion of a causal constraints model (CCM). These models are composed of a set of constraints, each of which are active only under selected intervention targets; their outcomes under a given intervention are the solutions of the constraints active under the intervention's target.

A CCM M can be viewed as a pair $(\mathcal{S}, \mathcal{C})$, where $\mathcal{S} = (\mathcal{U}, \mathcal{V}, \mathcal{R}, \mathcal{I})$ is a signature and $\mathcal{C}$ is a set of causal constraints.[9] Each constraint $C \in \mathcal{C}$ is a pair (f_C, a_C), where $f_C : \mathcal{R}(\mathcal{U}) \times \mathcal{R}(\mathcal{V}) \to \{0, 1\}$ and $a_C \subseteq \mathcal{P}(\mathcal{V})$. Given a context $\mathbf{u}$ and an intervention $\vec{X} \leftarrow \vec{x} \in \mathcal{I}$, the outcomes $M(\mathbf{u}, \vec{X} \leftarrow \vec{x})$ are all assignments $\mathbf{v}$ to the variables of $\mathcal{V}$ such that, for all $C \in \mathcal{C}$ with $\vec{X} \in a_C$, C is satisfied (e.g., we have $f_C(\mathbf{u}, \mathbf{v}) = 1$) and $\mathbf{v}[\vec{X}] = \vec{x}$. [10] Causal constraints models are equivalent to GSEMs.

Theorem 6.1. *For all GSEMs, there is an equivalent CCM, and vice versa.*

Of course, if we restrict to CCMs that satisfy the axioms in AX^+ (resp. AX^+_{rec}), we can prove an analogue of Theorem 6.1 for GSEMs satisfying AX^+ (resp. AX^+_{rec}).

CCMs were designed for characterizing equilibrium solutions of dynamical systems. Their constraint structure, where the outcomes are the solutions to a subsystem of constraints depending on the intervention, is useful there, but seems less well suited for a general-purpose causal modeling framework. Thus, even though they are equivalent in expressive power, we believe that CCMs and GSEMs will find complementary applications.

6.3 Structural Dynamical Causal Models

Structural dynamical causal models (SDCMs) [4] are designed with a different type of intervention in mind. Interventions act on entire *stochastic processes* (dynamical variables), and not, as in our previous examples, on state variables at a particular instant in time. Any SDCM can be captured quite directly by a GSEM whose variables range over functions of time. Without going into too much detail, the endogenous variables of the GSEM correspond to endogenous stochastic processes of the SDCM; the exogenous variables correspond to exogenous

[9] The definition of CCMs in [3] does not include a set of allowed interventions. We include one here (in effect, generalizing CCMs slightly) to allow for a fair comparison with GSEMs.

[10] The semantics given here is equivalent to the semantics in [3], but we simplified the exposition slightly. In particular, in [3], f_C is allowed to map to an arbitrary measurable space, each constraint has an additional constant c_C, and C is satisfied if $f_C(\mathbf{u}, \mathbf{v}) = c_C$. In addition, rather than just requiring that $\mathbf{v}[\vec{X}] = \vec{x}$, Blom et al. add constraints to enforce this condition. The changes that we have made do not affect the expressive power of CCMs. We do not give the semantics of the intervened CCMs M_I, since they can be derived from the semantics of outcomes in exactly the same way that we derived the semantics of intervened GSEMs in Section 3 (using composition of interventions).

stochastic processes of the SDCM, and the mapping $\mathbf{F}$ of the GSEM implements the dynamic structural equations of the SDCM. Although the GSEM variables represent different objects (functions, not values), the construction is analogous to the construction of M_{ODE} in Section 5.1.

References

[1] R. Alur, C. Courcoubetis, T. A. Henzinger, and P.-H. Ho. 1992. Hybrid automata: An algorithmic approach to the specification and verification of hybrid systems. In *Hybrid Systems*. Springer, 209–229.

[2] S. Beckers and J. Y. Halpern. 2019. Abstracting causal models. In *Proc. Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*. The full version appears at arxiv.org/abs/1812.03789.

[3] T. Blom, S. Bongers, and J. M. Mooij. 2019. Beyond structural causal models: causal constraints models. In *Proc. 35th Conference on Uncertainty in Artificial Intelligence (UAI 2019)*.

[4] S. Bongers, T. Blom, and J. M. Mooij. Causal modeling of dynamical systems. Available at https://arxiv.org/abs/1803.08784, (2018).

[5] H. Chockler and J. Y. Halpern. 2004. Responsibility and blame: a structural-model approach. *Journal of A.I. Research*, 20, 93–115.

[6] D. Galles and J. Pearl. 1998. An axiomatic characterization of causal counterfactuals. *Foundation of Science*, 3, 1, 151–182.

[7] J. Y. Halpern. 2015. A modification of the Halpern-Pearl definition of causality. In *Proc. 24th International Joint Conference on Artificial Intelligence (IJCAI 2015)*, 3022–3033.

[8] J. Y. Halpern. 2016. *Actual Causality*. MIT Press, Cambridge, MA.

[9] J. Y. Halpern. 2000. Axiomatizing causal reasoning. *Journal of A.I. Research*, 12, 317–337.

[10] J. Y. Halpern and J. Pearl. 2005. Causes and explanations: a structural-model approach. Part I: Causes. *British Journal for Philosophy of Science*, 56, 4, 843–887.

[11] J. Y. Halpern and J. Pearl. 2005. Causes and explanations: a structural-model approach. Part II: Explanations. *British Journal for Philosophy of Science*, 56, 4, 889–911.

[12] T. A. Henzinger. 2000. The theory of hybrid automata. In *Verification of Digital and Hybrid Systems*. Springer, 265–292.

[13] D. Ibeling and T. Icard. 2019. On open-universe causal reasoning. In *Proc. 35th Conference on Uncertainty in Artificial Intelligence (UAI 2019)*.

[14] J. Laurent, J. Yang, and W. Fontana. 2018. Counterfactual resimulation for causal analysis of rule-based models. In *Proc. Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI '18)*, 1882–1890.

[15] S. Peters and J. Y. Halpern. 2022. Reasoning about causal models with infinitely many variables. In *Proc. Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-20)*, 5668–5675.

[16] P. K. Rubenstein, S. Weichwald, S. Bongers, J. M. Mooij, D. Janzing, M. Grosse-Wentrup, and B. Schölkopf. 2017. Causal consistency of structural equation models. In *Proc. 33rd Conference on Uncertainty in Artificial Intelligence (UAI 2017)*.

[17] T. Young and M. J. Mohlenkamp. 2017. *Introduction to Numerical Methods and MATLAB Programming for Engineers*. Available at https://openlibra.com/en/book/introduction-to-numerical-methods-and-matlab-programming-for-engineers. Ohio University.

A An Algorithm for Finding the Outcomes of M_{ODE}

Recall that $\mathcal{I}_{intervals}$ consists of all finite compositions $I_1; I_2; \ldots; I_m$, where each I_j is either a point intervention $X_i^t \leftarrow k$ or an interval intervention $\{X_i^t \mid t \in (a,b)\} \leftarrow k$, which we abbreviate as $X_i(a,b) \leftarrow k$ for readability. (We similarly use the abbreviations $X_i[a,b)$, $X_i(a,b]$, and $X_i[a,b]$. Note that intervening on each of these sets can be achieved by composing two or three point or interval interventions.)

Fix an intervention $\vec{X} \leftarrow \vec{x} = I = I_1; \ldots; I_m \in I_{intervals}$. Let $t_1 < t_2 < \cdots < t_l$ be the endpoints of the intervals $I_1, \ldots, I_m$. (The endpoint of $X_i^t \leftarrow k$ is t and the endpoints of $X_i(a,b) \leftarrow f$ are a and b.) For convenience, define $t_0 = 0$. It is easy to see that each interval (t_i, t_j) is intervention-constant. Thus, we can find outcomes of the model step by step. The following algorithm finds an outcome of M_{ODE} under the intervention I with initial conditions $\mathbf{u} = (X_1^0, \ldots, X_n^0)$.

Algorithm A

(1) For $1 \leq i \leq n$, define $\mathcal{X}_i(0) = X_i^0$.

(2) For $i = 1, \ldots, l$:

(a) For $j = 1, \ldots, n$, if $\mathcal{X}_j$ is set to f on (t_{i-1}, t_i), define $\mathcal{X}_j(t) = f(t)$ for all $t \in (t_{i-1}, t_i)$.

(b) Define the remaining (intervention-free) dynamical variables on (t_{i-1}, t_i) so that

for $j = 1, \dots, n$, if $\mathcal{X}_j$ is intervention-free, then $\mathcal{X}_j$ is right-continuous at t_{i-1}, differentiable on (t_{i-1}, t_i), and its derivative $\dot{\mathcal{X}}_j$ satisfies $\dot{\mathcal{X}}_j(t) = F_j(\mathcal{X}_1(t), \dots, \mathcal{X}_m(t))$ for all $t \in (t_{i-1}, t_i)$.
If there is no way to do this, output "No solution".
(c) For $j = 1, \dots, n$, define $\mathcal{X}_j(t_i)$ as follows.
(i) If $X_j^{t_i} \in \vec{X}$, define $\mathcal{X}_j(t_i) = \vec{x}[X_j^{t_i}]$.
(ii) If $X_j^{t_i} \notin \vec{X}$,
define $\mathcal{X}_j(t_i) = \lim_{t \to t_i^-} \mathcal{X}_j(t)$.
(3) Define the functions $\mathcal{X}_i$, $1 \le i \le n$, on (t_l, ∞) so that $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n$ solve the initial-value problem on (t_l, ∞),
as defined in ODE3′. Again, if there is no way to do this, output "No solution".
(4) Output the outcome $\mathbf{v}$ defined by $\mathbf{v}[X_i^t] = \mathcal{X}_i(t)$ for all $1 \le i \le n$, $t > 0$.

As we mentioned before, this algorithm is underspecified. The modeler can choose how to find and represent the solutions of the initial-value problems appearing in steps 2(b) and 3. Depending on the context, the solution could be computed exactly and stored as an algebraic expression, or computed numerically and stored as a sequence of real values. Moreover, these solutions are not unique; there will in general be uncountably many solutions to each initial value problem. For each initial-value problem that arises, any solution of that problem can be used. Each different choice of solutions to the initial-value problems leads to a different valid execution of the algorithm, and each valid execution outputs a different output outcome.

The correspondence between the solutions of the model and outputs of the algorithm is formalized in the following theorem.

THEOREM A.1. *The set of outcomes output by valid executions of Algorithm A on input* $\mathbf{u} \in \mathcal{U}, I \in \mathcal{I}_{intervals}$ *are exactly the outcomes* $M_{ODE}(\mathbf{u}, I)$.

An alternative approach to modeling this situation would allow "partial outcomes": that is, whenever "No solution" would be output in iteration i, instead define $\mathbf{v}[X_j^t] = error$ for all $1 \le j \le n$ and all $t > t_{i-1}$ (where $error$ is a special value indicating that the differential equations to go have no solution). The alternative model M'_{ODE} with the outcomes corresponding to this modified algorithm satisfies an analogue of acyclicity, which we explore in a companion paper. (A very similar reformulation yields a version of the hybrid automaton model in Section 5 that also satisfies this acyclicity condition.)

B An Axiom System for Causal Reasoning

We now review the axiom systems considered by Halpern 2000 for reasoning about causality. Note that there are two slight differences between our presentation and that of Halpern. First, as we mentioned earlier, we have weakened the language of causal formulas so that primitive causal formulas are no longer parameterized by contexts. Thus, our language has formulas such as $[\vec{Y} \leftarrow \vec{y}](X = x)$ rather than $[\vec{Y} \leftarrow \vec{y}](X(\mathbf{u}) = x)$. Second, the list of axioms given below does not include two of Halpern's axioms, which he called D10 and D11. D11 is a technical axiom that was needed only to reason about formulas with contexts (to reduce to formulas that mentioned ony one context); D10 says that there are unique outcomes, and is redundant in acyclic systems. A minor modification of Halpern's proof shows that the axiom systems AX^+ and AX^+_{rec} defined below (which are identical to the system Halpern called AX^+ and AX^+_{rec}, respectively, except that they omit the axioms D10 and D11) are sound and complete for SEMs and acyclic SEMs, respectively, with respect to the language that we are considering (just as Halpern's versions of AX^+ and AX^+_{rec} were sound and complete for his language); the proof is essentially identical to Halpern's, so we omit it here. To axiomatize acyclic SEMs, following Halpern, we define

$Y \rightsquigarrow Z$, read "Y affects Z", as an abbreviation for the formula

$$\bigvee_{\vec{X} \subseteq \mathcal{V}, \vec{x} \in \mathcal{R}(\vec{X}), y \in \mathcal{R}(y), z \neq z' \in \mathcal{R}(Z)} ([\vec{X} \leftarrow \vec{x}](Z = z) \wedge [\vec{X} \leftarrow \vec{x}, Y \leftarrow y](Z = z'));$$

that is, Y affects Z if there is some setting of some endogenous variables $\vec{X}$ for which changing the value of Y changes the value of Z. This definition is used in axiom D6 below, which characterizes acyclicity.

Definition B.1. AX^+ consists of axiom schema D0-D5 and D7-D9, and inference rule MP. AX^+_{rec} results from adding D6 to AX^+.

- D0. All instances of propositional tautologies.
- D1. $[\vec{Y} \leftarrow \vec{y}](X = x \Rightarrow X \neq x')$ if $x, x' \in \mathcal{R}(X)$, $x \neq x'$ (functionality)
- D2. $[\vec{Y} \leftarrow \vec{y}](x \in \mathcal{R}(X))$ (definiteness)
- D3. $\langle \vec{X} \leftarrow \vec{x} \rangle (W = w \wedge \vec{Y} = \vec{y}) \Rightarrow \langle \vec{X} \leftarrow \vec{x}; W \leftarrow w \rangle (\vec{Y} = \vec{y})$ (composition)
- D4. $[\vec{W} \leftarrow \mathbf{w}; X \leftarrow x](X = x)$ (effectiveness)
- D5. $(\langle \vec{X} \leftarrow \vec{x}; Y \leftarrow y \rangle (W = w \wedge \vec{Z} = \vec{z}) \wedge \langle \vec{X} \leftarrow \vec{x}; W \leftarrow w \rangle (Y = y \wedge \vec{Z} = \vec{z}))$ $\Rightarrow \langle \vec{X} \leftarrow \vec{x} \rangle (W = w \wedge Y = y \wedge \vec{Z} = \vec{z})$, where $\vec{Z} = \mathcal{V} - (\vec{X} \cup \{W, Y\})$ (reversibility)
- D6. $(X_0 \rightsquigarrow X_1 \wedge \ldots \wedge X_{k-1} \rightsquigarrow X_k) \Rightarrow \neg(X_k \rightsquigarrow X_0)$ (recursiveness)
- D7. $([\vec{X} \leftarrow \vec{x}]\varphi \wedge [\vec{X} \leftarrow \vec{x}](\varphi \Rightarrow \psi)) \Rightarrow [\vec{X} \leftarrow \vec{x}]\psi$ (distribution)
- D8. $[\vec{X} \leftarrow \vec{x}]\varphi$ if φ is a propositional tautology (generalization)
- D9. $\langle \vec{Y} \leftarrow \vec{y} \rangle true \wedge \big(\langle \vec{Y} \leftarrow \vec{y} \rangle (X = x) \Rightarrow \langle Y \leftarrow y \rangle (X \neq x')\big)$, if $x \neq x'$ and $Y = \mathcal{V} - \{X\}$. (unique outcomes for $\mathcal{V} - \{X\}$)
- MP. From φ and $\varphi \Rightarrow \psi$, infer ψ (modus ponens)

C Proofs

Theorem 2.1: *If M and M' are causal models over the same signature $\mathcal{S}$ that, given a context and intervention, return a finite set of outcomes, then M and M' have the same outcomes (that is, for all $\mathbf{u} \in \mathcal{R}(\mathcal{U})$ and $I \in \mathcal{I}$, $M(\mathbf{u}, I) = M'(\mathbf{u}, I)$) if and only if they satisfy the same set of causal formulas (that is, for all $\mathbf{u} \in \mathcal{R}(\mathcal{U})$ and $\psi \in \mathcal{L}(\mathcal{S})$, $M, \mathbf{u} \models \psi \Leftrightarrow M', \mathbf{u} \models \psi$).*

Proof. Let M and M' be causal models with the same set of solutions. It suffices to consider the primitive causal formulas $[\vec{Y} \leftarrow \vec{y}](X = x)$, since the truth of other formulas in $\mathcal{L}(\mathcal{S})$ are derived from these. Recall that $M, \mathbf{u} \models [\vec{Y} \leftarrow \vec{y}](X = x)$ iff for all outcomes $\mathbf{v} \in M(\mathbf{u}, \vec{Y} \leftarrow \vec{y})$, $\mathbf{v}[X] = x$. But $M(\mathbf{u}, \vec{Y} \leftarrow \vec{y}) = M'(\mathbf{u}, \vec{Y} \leftarrow \vec{y})$, so $M, \mathbf{u} \models [\vec{Y} \leftarrow \vec{y}](X = x)$ if and only if $M', \mathbf{u} \models [\vec{Y} \leftarrow \vec{y}](X = x)$. Conversely, suppose that M and M' satisfy the same set of causal formulas. Suppose for contradiction that there exists some $\mathbf{u}$ and I with $M(\mathbf{u}, I) \neq M'(\mathbf{u}, I)$. Then without loss of generality, there is an outcome $\mathbf{v}$ in $M(\mathbf{u}, I)$ that is not in $M'(\mathbf{u}, I)$. This outcome must differ from each of the finitely many outcomes $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} = M'(\mathbf{u}, I)$ in at least one variable; that is, there must be variables $X_1, X_2, \ldots, X_n \in \mathcal{V}$ with $\mathbf{v}[X_i] \neq \mathbf{v}_i[X_i]$ for $i \in 1, 2, \ldots, n$. Consider the causal formula $\varphi = \langle I \rangle (\bigwedge_{1 \leq i \leq n} X_i = \mathbf{v}[X_i])$. We have that $M, \mathbf{u} \models \varphi$, since $\mathbf{v} \in M(\mathbf{u}, I)$. However, it is not true that $M', \mathbf{u} \models \varphi$, because no outcome $\mathbf{v}_i$ of M' satisfies $\bigwedge_{1 \leq i \leq n} X_i = \mathbf{v}[X_i]$. This contradicts the assumption s that M and M' satisfy the same set of causal formulas; hence $M(\mathbf{u}, I) = M'(\mathbf{u}, I)$ for all $\mathbf{u} \in \mathcal{R}(\mathcal{U})$, $I \in \mathcal{I}$. □

Proposition 3.1: *For all SEMs M, there is a GSEM M' such that $M \equiv M'$.*

Proof. Given a SEM $M = ((\mathcal{U}, \mathcal{V}, \mathcal{R}, \mathcal{I}), \mathcal{F})$, define $M' = ((\mathcal{U}, \mathcal{V}, \mathcal{R}, \mathcal{I}), \mathcal{F}')$, where for all $\vec{X} \leftarrow \vec{x} \in \mathcal{I}$ and $\mathbf{u} \in \mathcal{U}$, $\mathcal{F}'(\mathbf{u}, \vec{X} \leftarrow \vec{x}) = M(\mathbf{u}, \vec{X} \leftarrow \vec{x})$. Since $M'(\mathbf{u}, \vec{X} \leftarrow \vec{x}) = \mathcal{F}'(\mathbf{u}, \vec{X} \leftarrow \vec{x})$, M' is equivalent to M by definition. □

Proposition 3.2: *For all SEMs M and interventions $I, J \in \mathcal{I}$ such that $I;J \in \mathcal{I}$, we have that $M_I(\mathbf{u}, J) = M(\mathbf{u}, I; J)$.*

Proof. We prove the equivalent statement $(M_I)_J(\mathbf{u}) = M_{I;J}(\mathbf{u})$. Since the outcomes of SEMs are determined by the structural equations, it suffices to show that the structural equations of $(M_I)_J$ are the same as those of $M_{I;J}$. Let $X \in \mathcal{V}$ be arbitrary and consider the structural equation $\mathcal{F}_X$. Without loss of generality, let $I = \vec{Y} \leftarrow \vec{y}$ and $J = \vec{Z} \leftarrow \vec{z}$. There are three cases to consider: $X \notin \vec{Y} \cup \vec{Z}$, $X \in \vec{Z}$, and $X \in \vec{Y} \setminus \vec{Z}$. The first case is trivial; $\mathcal{F}_X$ is unmodified in both models. In the second case, letting $\mathbf{s} \in \mathcal{R}(\mathcal{V} \setminus \{X\})$ denote an arbitrary input to $\mathcal{F}_X$, in $(M_I)_J$, we have that $\mathcal{F}_X(\mathbf{s}) = \vec{Z}[X]$. But by the definition of $I;J$, we also have $\mathcal{F}_X(\mathbf{s}) = \vec{Z}[X]$ in $M_{I;J}$. In the third case, in $(M_I)_J$, $\mathcal{F}_X(\mathbf{s}) = \vec{Y}[X]$, since in M_I, $\mathcal{F}_X(\mathbf{s}) = \vec{Y}[X]$, and applying the intervention J does not affect $\mathcal{F}_X$ since $X \notin \vec{Z}$. □

Proposition 3.3: *Suppose that M is a SEM, M' is a GSEM, and $M \equiv M'$. Then for all $I \in \mathcal{I}$, we have that $M_I \equiv M'_I$.*

Proof. Clearly M_I and M'_I have the same signatures. It remains to show that for all contexts $\mathbf{u}$ and all intervention J allowed in M_I, we have that $M_I(\mathbf{u}, J) = M'_I(\mathbf{u}, J)$. Applying the definition and the fact that $M \equiv M'$, we have that $M'_I(\mathbf{u}, J) = M'(\mathbf{u}, I; J) = M(\mathbf{u}, I; J)$. Therefore, it suffices to show $M(\mathbf{u}, I; J) = M_I(\mathbf{u}, J)$, which is exactly Theorem 3.2. □

The following theorem is needed to prove Theorem 3.4.

Theorem C.1. *If M and M' are causal models (either SEMs or GSEMs) with a common signature $\mathcal{S} = (\mathcal{U}, \mathcal{V}, \mathcal{R}, \mathcal{I}_{univ})$, where $\mathcal{V}$ is finite and $\mathcal{R}(X)$ is finite for all $X \in \mathcal{V}$, that both satisfy the axioms in AX^+ and have the same outcomes under complete interventions—that is, for all $\mathbf{u} \in \mathcal{R}(\mathcal{U})$ and $X \in \mathcal{V}$, if $\vec{Y} = \mathcal{V} \setminus X$, then for all $\vec{y} \in \mathcal{R}(\vec{Y})$, $M(\mathbf{u}, \vec{Y} \leftarrow \vec{y}) = M'(\mathbf{u}, \vec{Y} \leftarrow \vec{y})$—*

then M and M' agree on all causal formulas.

Proof. Fix an arbitrary context $\mathbf{u}$. M satisfies axiom D9, so for every variable $X \in \mathcal{V}$, and for every assignment $\vec{y} \in \mathcal{R}(\vec{Y})$ to the variables $\vec{Y} = \mathcal{V} \setminus \{X\}$, there is a unique $x \in \mathcal{R}(X)$ such that $M, \mathbf{u} \models [\vec{Y} \leftarrow \vec{y}](X = x)$. Using this fact, we can define a SEM M'' with signature $\mathcal{S}$ as follows. Define $\mathcal{F}''_X(\mathbf{u}, \vec{y})$ to be the unique x such that $M, \mathbf{u} \models [\mathcal{V} \setminus \{X\} \leftarrow \vec{y}](X = x)$. Let C be the set of all formulas $\varphi = [\mathcal{V} \setminus \{X\} \leftarrow \vec{y}](X = x)$ such that $M, \mathbf{u} \models \varphi$. By assumption, C is also the set of all such formulas φ for which $M', \mathbf{u} \models \varphi$. Let χ be the conjunction of all the formulas in C. Since there are finitely many variables, and all ranges are finite, this set of formulas is finite, and so taking the conjunction makes sense. We know that M and M' satisfy all axioms of AX^+, and both models satisfy χ. This means that if $\chi \Rightarrow \psi$ is provable in AX^+, then M and M' both satisfy ψ. We now show that, for all formulas ψ, either $\chi \Rightarrow \psi$ or $\chi \Rightarrow \neg\psi$ is provable in AX^+. This means that either both $M, \mathbf{u} \models \psi$ and $M', \mathbf{u} \models \psi$ (if $\chi \Rightarrow \psi$ is provable in AX^+), or both $M, \mathbf{u} \not\models \psi$ and $M, \mathbf{u} \not\models \psi$ (if $\chi \Rightarrow \neg\psi$ is provable in AX^+). That is, M and M' agree on all causal formulas. Note that χ is false in all SEMs over $\mathcal{S}$ other than models that agree with the M'' that we defined using χ in context $\mathbf{u}$. Thus, if $(M'', \mathbf{u}) \models \psi$, then $\chi \Rightarrow \psi$ is valid; and if $(M'', \mathbf{u}) \models \neg\psi$, then $\chi \Rightarrow \neg\psi$ is valid. Since AX^+ is a sound and complete axiomatization, it follows that either $\chi \Rightarrow \psi$ or $\chi \Rightarrow \neg\psi$ is provable, as desired. □

Theorem 3.4: *For all finite GSEMs over a signature $\mathcal{S}$ such that $\mathcal{I} = \mathcal{I}_{univ}$ and all the axioms of AX^+ are valid, there is an equivalent SEM, and vice versa.*

Proof. Given a SEM M, define a GSEM M' with the same signature by taking $\mathbf{F}'(\mathbf{u}, I) = M(\mathbf{u}, I)$, as in Theorem 3.1. This GSEM is clearly equivalent to M. Furthermore, all the axioms in AX^+ are valid in M. This follows from the facts that (1) equivalent causal models have the same outcomes (by definition), (2) finite causal models with the same outcomes satisfy the same causal formulas (Theorem 2.1), and (3) M is a SEM, so all the axioms in AX^+ are valid in M. Conversely, given a finite GSEM M' in which all the axioms of AX^+ are valid, the GSEM must have unique solutions for $\mathcal{V} \setminus X$ (D9). That is, for each context $\mathbf{u} \in \mathcal{R}(\mathcal{U})$ and each variable $X \in \mathcal{V}$, if we define $\vec{Y} = \mathcal{V} \setminus X$, for every $\vec{y} \in \mathcal{R}(Y)$, there is a unique $x \in \mathcal{R}(X)$ such that $M', \mathbf{u} \models [\vec{Y} \leftarrow \vec{y}](X = x)$. Here we use the fact that $\mathcal{I} = \mathcal{I}_{univ}$ to ensure that the relevant instances of D9 are in the language. We can use this property to define the structural equations of the SEM M. That is, define a SEM M with the same signature by defining $\mathcal{F}_X(\mathbf{u}, \vec{y}) = x$, where x is the value guaranteed above. We must show that M has the same outcomes as M'. But this is just Theorem C.1. □

Theorem 3.5: *For all finite GSEMs over a signature $\mathcal{S}$ such that $\mathcal{I} = \mathcal{I}_{univ}$ and all the axioms AX^+_{rec} are valid, there is an equivalent acyclic SEM, and vice versa.*

Proof. Given a finite GSEM M' satisfying AX^+_{rec}, Theorem 3.4 guarantees the existence of an equivalent SEM M. Since M is equivalent to M', M satisfies AX^+_{rec}. This implies that M is acyclic. To prove this, suppose not. Then there is $k > 1$ and endogenous variables $V_1, \ldots, V_k$ having cyclic dependencies; that is, V_{i+1} is not independent of V_i for $i = 1, \ldots, k-1$, and V_1 is not independent of V_k. But it is easy to see that if Y is not independent of X, then X affects Y, i.e., $X \rightsquigarrow Y$. Thus, $V_k \rightsquigarrow V_1 \wedge V_1 \rightsquigarrow V_2 \wedge \cdots \wedge V_{k-1} \rightsquigarrow V_k$. This is the negation of an instance of D6. Hence not all the axioms of AX^+_{rec} are valid in M, a contradiction. For the converse, given an acyclic SEM M, Theorem 3.1 guarantees the existence of an equivalent GSEM M'. This equivalent GSEM satisfies the same formulas as M, so it satisfies AX^+_{rec}. □

Theorem A.1: *The set of outcomes output by valid executions of Algorithm A on input $u \in \mathcal{U}, I \in \mathcal{I}_{intervals}$ are exactly the outcomes $M_{ODE}(\mathbf{u}, I)$.*

Proof. We walk through the algorithm's execution and show that whenever it defines a dynamical variable (and thus a model variable, via the translation in step 4), it can make all the choices compatible with ODE1, ODE2, and ODE3′, and cannot make any other choices:

- In step 1,
  $\mathcal{X}_i(0) = X_i^0$ is the only choice consistent with the right-continuity requirement of ODE3′.
- In step 2(a), $\mathcal{X}_j(t) = f(t)$ is the only choice consistent with ODE1. It is compatible with ODE2 and ODE3, since ODE2 and ODE3 require nothing of intervened points.
- In step 2(b) and step 3, the possible settings for the intervention-free variables are exactly the settings allowed by ODE3′.
  They are compatible with ODE1, since the intervention-free variables are not intervened on in (a, b), and compatible with ODE2, since solutions to initial value problems are always continuous.
- In step 2(c)(i), $\mathcal{X}_j(t_i) = \vec{x}[X_j^{t_i}]$ is the only choice consistent with ODE1. It is compatible with ODE2 and ODE3 since, again, ODE2 and ODE3 require nothing of intervened points.
- In step 2(c)(ii), $\mathcal{X}_j(t_i) = \lim_{t \to t_i^-} \mathcal{X}_j(t)$ is the only choice that maintains left-continuity (is consistent with ODE2). It is compatible with ODE1, since $\mathcal{X}_j$ is not intervened on at time t_i, and compatible with ODE3, since by construction, there is no intervention-constant open interval containing t_i.
  Finally, the limit always exists, because the values of $\mathcal{X}_j$ on (t_{i-1}, t_i) were set in step 2(b), so $\mathcal{X}_j$ is continuous on the open interval (t_{i-1}, t_i).

□

Theorem 6.1: *For all GSEMs, there is an equivalent CCM, and vice versa.*

Proof. Given a CCM $M' = (\mathcal{S}, \mathcal{C}, I)$, define a GSEM $M = (\mathcal{S}, \mathbf{F}, \mathcal{I})$ by taking $\mathbf{F}(\mathbf{u}, I) = M'(\mathbf{u}, I)$; it is immediate that M and M' have the same outcomes. For the converse, given a GSEM $M = (\mathcal{S}, \mathbf{F}, \mathcal{I})$, define a CCM $M' = (\mathcal{S}, \mathcal{C}, \mathcal{I})$ as follows. For every intervention $\vec{X} \leftarrow \vec{x} \in \mathcal{I}$, $\mathcal{C}$ contains a constraint $C_{\vec{X} \leftarrow \vec{x}}$ such that $a_{C_{\vec{X} \leftarrow \vec{x}}} = \{\vec{X}\}$, and for every context $\mathbf{u}$, $f_{C_{\vec{X} \leftarrow \vec{x}}}(\mathbf{u}, \mathbf{v}) = 1$ iff either $\mathbf{v}[\vec{X}] \neq \vec{x}$ or $\mathbf{v} \in \mathbf{F}(\mathbf{u}, \vec{X} \leftarrow \vec{x})$. We claim that the outcomes

$M'(\mathbf{u}, \vec{X} \leftarrow \vec{x})$ are exactly the GSEM outcomes $\mathbf{F}(\mathbf{u}, \vec{X} \leftarrow \vec{x})$. Indeed, suppose $\mathbf{v} \in M'(\mathbf{u}, \vec{X} \leftarrow \vec{x})$. Then $\mathbf{v}[\vec{X}] = \vec{x}$, and it follows from the constraint $C_{\vec{X} \leftarrow \vec{x}}$ that $\mathbf{v} \in \mathbf{F}(\mathbf{u}, \vec{X} \leftarrow \vec{x})$. For the opposite implication, suppose that $\mathbf{v} \in \mathbf{F}(\mathbf{u}, \vec{X} \leftarrow \vec{x})$. Then $\mathbf{v}[\vec{X}] = \vec{x}$ (since GSEMs satisfy effectiveness). Moreover, $\mathbf{v}$ satisfies all active constraints. It satisfies the constraint corresponding to $\vec{X} \leftarrow \vec{x}$, since $\mathbf{v} \in \mathcal{F}(\mathbf{u}, I)$. And it satisfies the constraints $C_{\vec{X} \leftarrow \vec{x}'}$ for $\vec{x}' \neq \vec{x}$, since $\mathbf{v}[\vec{X}] = \vec{x} \neq \vec{x}'$. □

Theorem C.2. *M_{shell} satisfies all the axioms in AX^+.*

Proof. D0, D1, D2, D7 and D8 are trivial. No joint interventions are allowed, so the only way to instantiate D3 is to have $\vec{X} = W = S_1$ (or symmetrically, $\vec{X} = W = S_2$). But if $\vec{X} = W$, then $\vec{X} \leftarrow \vec{x}; W \leftarrow w = W \leftarrow w$ and D3 follows trivially by eliminating the conjunction. D4 (effectiveness) holds by inspection. D5 holds for the same reason as D3. Finally, D9 cannot be instantiated because no complete interventions are allowed. □

Received 29 October 2023; revised 02 June 2024; accepted 04 June 2024