Generalized Structural Causal Models

Tineke Blom
Informatics Institute
University of Amsterdam
The Netherlands
t.blom2@uva.nl

Joris M. Mooij
Informatics Institute
University of Amsterdam
The Netherlands
j.m.mooij@uva.nl

Abstract

Structural causal models are a popular tool to describe causal relations in systems in many fields such as economy, the social sciences, and biology. In this work, we show that these models are not flexible enough in general to give a complete causal representation of equilibrium states in dynamical systems that do not have a unique stable equilibrium independent of initial conditions. We prove that our proposed *generalized structural causal models* do capture the essential causal semantics that characterize these systems. We illustrate the power and flexibility of this extension on a dynamical system corresponding to a basic enzymatic reaction. We motivate our approach further by showing that it also efficiently describes the effects of interventions on functional laws such as the ideal gas law.

1   INTRODUCTION

Real world processes are often complex and evolve over time. A popular approach to model such systems is to describe the dynamics by a system of (random) differential equations. This method has many applications in a wide range of fields (e.g. classical mechanics, thermodynamics, biochemistry, population biology, and economy).

An interesting question is whether dynamical models also admit a causal interpretation, see for instance [1, 2]. One approach is to model the changes of the system in a discrete time-setting as in [3]. Another promising idea is to model the causal semantics of the time-independent equilibrium states of a dynamical system as in [4, 5, 6, 2]. In this article we consider the causal semantics of the equilibrium states of a dynamical system, allowing for and modeling the dependence on initial conditions.

A popular causal modeling framework to model static systems is provided by Structural Causal Models [7]. These are well-understood and have recently been extended to also include the cyclic case, see [8] and [9]. In [4, 10, 5, 11, 12, 6, 13] it is shown how *cyclic* structural causal models may arise from studying the stationary behavior of certain dynamical time series models or differential equations. But can the equilibrium states of a dynamical system be represented by a causal model in general?

In this work, we prove that a structural causal model is not flexible enough in general to capture the rich causal semantics of the equilibrium behavior in a dynamical system whose asymptotic behavior depends on their initial conditions. The main contribution of this paper is the introduction of a novel way to represent the causal semantics of complex data-generating processes and dynamical systems: We generalize the SCM framework and prove that our proposed *generalized structural causal model* (GSCM) represents the behavior of equilibrium states in dynamical systems while pertaining all causal semantics.

In an SCM, each endogenous variable is associated with a structural equation that describes its causal dependence on other variables in the system, which induces a set of probability distributions over the space of endogenous variables. We generalize the notion of a structural equation to the concept of a *causal constraint*, which is a functional relation between variables that is invariant under a specified set of interventions. A generalized structural causal model is then a set of causal constraints in combination with a probability distribution on the exogenous variables.

We illustrate the power and flexibility of our approach on a basic enzymatic reaction that can be modeled as a dynamical system. These reactions are typically described by a system of differential equations. We show that, on the one hand, an SCM can never give a *complete* descrip-

tion of the equilibrium behavior in this model, while a GSCM, on the other hand, fully captures the rich causal semantics pertaining to this system. We demonstrate how the causal constraints of this GSCM can be derived naturally from the equations and constants of motion in the dynamical system. Furthermore, we use our framework to identify a marginal model that can be described by an SCM, and which can thus be analyzed using existing techniques.

We further motivate our approach by considering functional laws, which describe a relation between variables that is invariant under all interventions. This is another popular tool, especially for modeling laws of nature. We show that a GSCM can efficiently and naturally describe such functional laws. As an example, we consider the ideal gas law, which relates physical quantities in an ideal gas. We demonstrate that a GSCM representation, contrary to an SCM representation, does not admit solutions that violate the ideal gas law.

1.1 PRELIMINARIES

Throughout this paper we will denote tuples or sets by capital letters and sets of variables by bold capital letters. We will assume that all sets of variables X are indexed by a set $\mathbb{I}_X$ and take values in the product space of standard measurable spaces, $\mathcal{X} = \prod_{i \in \mathbb{I}_X} \mathcal{X}_i$.

There are many types of interventions that can be modeled that correspond to different experimental procedures. A *perfect* ("surgical") intervention on a variable forces that variable to take on a specific value through some external force acting on the system. Henceforth, we will always refer to perfect interventions.

1.1.1 Structural Causal Models

A statistical model over variables X typically is a pair $(\mathcal{X}, \mathbb{P}^{\mathcal{X}})$ where $\mathbb{P}^{\mathcal{X}}$ is a (parametrized) family of probability distributions on $\mathcal{X}$. It cannot be used to predict the effects of interventions, but causal models *can*. A causal model can be thought of as a family of statistical models, one for each intervention,

$$\mathbb{P}^{\mathcal{X}} = \left(\mathbb{P}^{\mathcal{X}}_{\text{do}(I, \boldsymbol{\xi}_I)} : \quad I \in \mathcal{P}(\mathbb{I}_X), \quad \boldsymbol{\xi}_I \in \mathcal{X}_I \right), \quad (1)$$

where $\mathcal{P}(\mathbb{I}_X)$ denotes the power set of $\mathbb{I}_X$, I is an intervention *target* and $\boldsymbol{\xi}_I$ a tuple of values which the target is forced to take on. We let $I = \emptyset$ denote the null-intervention corresponding to the observed system.

Structural causal models are a special type of causal models that are specified by so-called structural equations. Our formal treatment of SCMs follows that of [8] and is slightly different from that of [7], because we do not assume acyclicity (i.e. recursiveness).

Definition 1. *(Structural Causal Model (SCM))* Let $\mathbb{I}_X$ be an index set for endogenous variables X taking value in the product of standard measurable spaces, $\mathcal{X} = \prod_{i \in \mathbb{I}_X} \mathcal{X}_i$. Let E be a tuple of latent exogenous variables taking value in a product of standard measurable spaces $\mathcal{E} = \prod_{i \in \mathbb{I}_E} \mathcal{E}_i$. An SCM is a triple $(\mathcal{X}, F, \mathbb{P}_{\mathcal{E}})$ where,

- $\mathbb{P}_{\mathcal{E}}$ is a factorizing probability measure on $\mathcal{E}$.[1]

- F is a family of functions[2],

$$f_j : \mathcal{X}_{\text{pa}(j)} \times \mathcal{E}_{\text{pa}(j)} \to \mathcal{X}_j \qquad \forall j \in \mathbb{I}_X,$$

Note that a cyclic structural causal model does not need to imply a unique joint distribution $\mathbb{P}^{\mathcal{X}}_{\text{do}(\emptyset)}$ on the space of endogenous variables in the observed system, although acyclic SCMs do [8].

Definition 2. A uniquely solvable SCM is an SCM $\mathcal{M} = (\mathcal{X}, F, \mathbb{P}_{\mathcal{E}})$ such that the *structural equations*

$$x_j = f_j(\boldsymbol{x}_{\text{pa}(j)}, \boldsymbol{e}_{\text{pa}(j)}) \qquad \forall j \in \mathbb{I}_X,$$

with $(\boldsymbol{x}, \boldsymbol{e}) \in \mathcal{X} \times \mathcal{E}$ have a unique solution $\boldsymbol{x}^* \in \mathcal{X}$ $\mathbb{P}_{\mathcal{E}}$-a.s.

We say that a pair of random variables (X^*, E^*), where $\mathbb{P}^{E^*} = \mathbb{P}_{\mathcal{E}}$, that satisfies the structural equations in an SCM $\mathbb{P}_{\mathcal{E}}$-a.s. is a *solution* of the SCM. If the SCM is not uniquely solvable, there may be multiple solutions or there may not exist a solution at all.

An intervention $\text{do}(I, \boldsymbol{\xi}_I)$ with target $I \subseteq \mathcal{P}(\mathbb{I}_X)$ and value $\boldsymbol{\xi}_I \in \mathcal{X}_I$ on an SCM $\mathcal{M} = (\mathcal{X}, F, \mathbb{P}_{\mathcal{E}})$ maps it to the intervened SCM $\mathcal{M}_{\text{do}(I, \boldsymbol{\xi}_I)} = (\mathcal{X}, \widetilde{F}, \mathbb{P}_{\mathcal{E}})$ such that

$$\tilde{f}_j(\boldsymbol{x}_{\text{pa}(j)}, \boldsymbol{e}_{\text{pa}(j)}) = \begin{cases} \xi_j & j \in I \\ f_j(\boldsymbol{x}_{\text{pa}(j)}, \boldsymbol{e}_{\text{pa}(j)}) & j \in \mathbb{I}_X \backslash I. \end{cases}$$

2 GENERALIZED STRUCTURAL CAUSAL MODELS

As we will show in Section 3, SCMs cannot in general capture the causal semantics of equilibrium states in dynamical systems with initial conditions. To that end, we introduce *generalized* structural causal models here, so that a larger class of systems can be described. Later on, we prove that these *do* fully describe the equilibrium behavior of dynamical systems with initial conditions.

---

[1]In the case that the set of exogenous variables is empty, one can take a trivial probability measure over a point for $\mathbb{P}_{\mathcal{E}}$.

[2]$\text{pa}(j) \subseteq \mathbb{I}_X \cup \mathbb{I}_E$ denotes a subset of indexes that are sufficient to determine the values of f_j.

SCMs are specified by structural equations while generalized structural causal models are specified by *causal constraints*, which describe a relation between endogenous and exogenous variables which is active under specific intervention targets.

Definition 3. (*Generalized Structural Causal Model (GSCM)*) Let $\mathbb{I}_{\boldsymbol{X}}$, $\boldsymbol{X}$, $\boldsymbol{\mathcal{X}}$, $\boldsymbol{E}$, and $\boldsymbol{\mathcal{E}}$ be as in Definition 1. A GSCM is a triple $(\boldsymbol{\mathcal{X}}, G, \mathbb{P}_{\boldsymbol{\mathcal{E}}})$ where $G = \{(g_i, A_i) : i \in \mathbb{I}_G\}$ is a set of causal constraints, and a *causal constraint* is a pair (g_i, A_i) where,

- $g_i : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(i)} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(i)} \to \mathbb{R}$ is a measurable function[3].
- $A_i \subseteq \mathcal{P}(\mathbb{I}_{\boldsymbol{X}})$ is a set of intervention targets under which g_i is *active*.

For a model $(\boldsymbol{\mathcal{X}}, G, \mathbb{P}_{\boldsymbol{\mathcal{E}}})$, let $\boldsymbol{g}_\emptyset$ denote all functions that are *active* in the observational system, that is $\boldsymbol{g}_\emptyset = \{g : (g, A) \in G : \emptyset \in A\}$. We refer to

$$\boldsymbol{g}_\emptyset(\boldsymbol{x}, \boldsymbol{e}) = \boldsymbol{0}, \tag{2}$$

with $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ and $\boldsymbol{e} \in \boldsymbol{\mathcal{E}}$ as the *generalized* structural equations of the model.

The solution space of a generalized structural causal model consists of all values that satisfy the generalized structural equations:

$$\boldsymbol{S} = \{(\boldsymbol{x}, \boldsymbol{e}) \in \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}} : \boldsymbol{g}_\emptyset(\boldsymbol{x}, \boldsymbol{e}) = \boldsymbol{0}\}. \tag{3}$$

The main idea behind Definition 3 is that we can use an arbitrary number of flexible equations to restrict the solution space of the causal model, but these restrictions do not need to hold under all interventions. The following example illustrates how a causal constraint for a GSCM can be constructed and interpreted.

Example 1. Consider the price, supply, and demand of a certain product, which will be denoted by P, S, D respectively. Suppose a simple economic model states that supply equals demand, unless the price of the product is intervened upon (e.g. when there is price-fixing)[4]. We can capture this description by the following causal constraint,

$$(g, A) = (S - D, \quad \{\emptyset, \{D\}, \{S\}, \{D, S\}\}). \tag{4}$$

The generalized structural equation $S - D = 0$ constrains the solution space of the model. The set A indicates that the constraint is active in the observational system, when either D or S is targeted by an intervention, or when D and S are both intervened upon. It does

---

[3]We can be more general by taking a standard measurable space $\boldsymbol{\mathcal{Y}}$ as the target space of g. For the sake of simplicity we consider $\mathbb{R}$ here.

[4]Note that this is a toy-example, and is not meant to reflect a realistic economic model.

not constrain the solution space when P is intervened upon. This economic model is captured by the GSCM $(\mathbb{R}^3, G = \{(g, A)\}, \mathbb{P}_\emptyset)$.

2.1 SOLUTIONS OF A GSCM

We define a solution of a GSCM in complete analogy with the definition of a solution to an SCM.

Definition 4. A pair of random variables $(\boldsymbol{X}, \boldsymbol{E})$ taking value in $\boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}}$ is a solution to a GSCM $(\boldsymbol{\mathcal{X}}, G, \mathbb{P}_{\boldsymbol{\mathcal{E}}})$ if

- $\mathbb{P}^{\boldsymbol{E}^*} = \mathbb{P}_{\boldsymbol{\mathcal{E}}}$.
- $\boldsymbol{g}_\emptyset(\boldsymbol{X}, \boldsymbol{E}) = \boldsymbol{0}$ almost surely.

Similar to the solutions of an SCM, Definition 4 implies that a GSCM may have a unique solution, multiple random variables with different distributions as solutions or it may have no solution at all. After we define interventions on GSCMs in the next section, we will show that for *any* SCM there exists a GSCM which has the same solutions under any intervention on endogenous variables.

2.2 INTERVENTIONS

An intervention on an SCM changes the structural equations of variables that are targeted by the intervention. For a GSCM, an intervention changes the set of causal constraints. When an intervention targets a set I and forces it to take on a certain value $\boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$, new causal constraints are introduced to describe this intervention. In addition, some generalized structural equations may no longer be active (i.e. they no longer constrain the solution space of the model), after an intervention. Hence the sets A in (g, A) are altered for causal constraints $(g, A) \in G$.

Definition 5. (*Interventions*) Let $\mathcal{M} = (\boldsymbol{\mathcal{X}}, G, \mathbb{P}_{\boldsymbol{\mathcal{E}}})$ be a GSCM and consider the intervention $\mathrm{do}(I, \boldsymbol{\xi}_I)$, where $I \in \mathcal{P}(\mathbb{I}_{\boldsymbol{X}})$ is the intervention target and $\boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$ the target value. The intervened GSCM is given by $\mathcal{M}_{\mathrm{do}(I, \boldsymbol{\xi}_I)} = (\boldsymbol{\mathcal{X}}, G_{\mathrm{do}(I, \boldsymbol{\xi}_I)}, \mathbb{P}_{\boldsymbol{\mathcal{E}}})$ where,

- For each $i \in I$ we have a causal constraint describing the intervened value of the targets, $(x_i - \xi_i, \mathcal{P}(\mathbb{I}_{\boldsymbol{X}} \backslash \{i\})) \in G_{\mathrm{do}(I, \boldsymbol{\xi}_I)}$.
- For each causal constraint $(g, A) \in G$ we have a causal constraint $(g, A_{\mathrm{do}(I)}) \in G_{\mathrm{do}(I, \boldsymbol{\xi}_I)}$, where

$$A_{\mathrm{do}(I)} = \bigcup_{A_i \in A : A_i \supseteq I} \{A_i \setminus J : J \subseteq I\}.$$

Definition 5 describes how activation sets A of causal constraints (g, A) change under interventions. It says that for any $A_i \in A$, and for any combination of two subsequent interventions such that $I_1 \cup I_2 = A_i$, the constraint will be active. So after I_1 (which needs to be

3

a subset of A_i), any I_2 that adds the remaining elements $A_i \setminus I_1$ (plus possibly any elements that were already in I_1) will activate the constraint.

Example 2. This example illustrates the effect of an intervention on a set $A = \{\emptyset, \{1,2\}, \{2,3\}\}$.

$$
\begin{aligned}
A_{\mathrm{do}(\emptyset)} &= \{\emptyset, \{1,2\}, \{2,3\}\}\,, \\
A_{\mathrm{do}(1)} &= \{\{2\}, \{1,2\}\}\,, \\
A_{\mathrm{do}(2)} &= \{\{1\}, \{1,2\}, \{3\}, \{2,3\}\}\,, \\
A_{\mathrm{do}(\{1,2\})} &= A_{\mathrm{do}(1)\mathrm{do}(2)} = A_{\mathrm{do}(2)\mathrm{do}(1)} \\
&= \{\emptyset, \{1\}, \{2\}, \{1,2\}\}\,, \\
A_{\mathrm{do}(\{1,2,3\})} &= \emptyset\,.
\end{aligned}
$$

Lemma 1 shows that the effect of multiple interventions on a GSCM does not depend on the order in which they are performed, or whether the interventions are performed simultaneously or sequentially.

Lemma 1. *Let $\mathcal{M}$ be a GSCM for variables $\boldsymbol{X}$ and let $I, J \subseteq \mathcal{P}(\mathbb{I}_{\boldsymbol{X}})$ be two disjoint sets of intervention targets with intervention values $\xi_I \in \boldsymbol{\mathcal{X}}_I$ and $\xi_J \in \boldsymbol{\mathcal{X}}_J$ respectively. Then*

$$
\begin{aligned}
(\mathcal{M}_{do(I,\xi_I)})_{do(J,\xi_J)} &= (\mathcal{M}_{do(J,\xi_J)})_{do(I,\xi_I)} \\
&= \mathcal{M}_{do(I \cup J, \xi_{I \cup J})}\,.
\end{aligned}
$$

Proof. The result follows directly from Definition 5. □

In the following example, we extend the model for price, supply, and demand that we considered in Example 1, and demonstrate how an intervention is performed on this extended GSCM.

Example 3. Consider the simple model for price, supply, and demand in Example 1, where supply is equal to demand when price is not intervened upon. Suppose an economic theory further states that the supply S is determined by price P and some exogenous variable E (e.g. cost of production) by some function f_S, unless the supply is targeted by an intervention. The model can be represented by the GSCM $\mathcal{M} = (\mathbb{R}^3, G, \mathbb{P}_{\mathcal{E}})$, where G consists of the causal constraints

$$
\begin{aligned}
S - D &= 0\,, & \{\emptyset, \{D\}, \{S\}, \{D,S\}\}\,, \\
S - f_S(P, E) &= 0 & \{\emptyset, \{D\}, \{P\}, \{D,P\}\}\,.
\end{aligned}
$$

We consider the intervened GSCM $\mathcal{M}_{\mathrm{do}(\mathbb{I}_P, \xi_P)} = (\boldsymbol{\mathcal{X}}, G_{\mathrm{do}(\mathbb{I}_P, \xi_P)}, \mathbb{P}_{\mathcal{E}})$. The causal constraints $(g, A_{\mathrm{do}(\mathbb{I}_P)}) \in G_{\mathrm{do}(\mathbb{I}_P, \xi_P)}$ are

$$
\begin{aligned}
S - D &= 0 & \emptyset\,, \\
S - f_S(P, E) &= 0 & \{\emptyset, \{D\}, \{P\}, \{D,P\}\}\,, \\
P - \xi_P &= 0 & \{\emptyset, \{D\}, \{S\}, \{D,S\}\}\,.
\end{aligned}
$$

Note that after an intervention on P, there is no intervention under which the first generalized structural equation is active. Hence it no longer restricts the solution space of the model under any intervention and the causal constraint can be discarded.

2.3 RELATION TO AN SCM

In this section, we show that GSCMs can represent any model that can be described by an SCM. In the next section we show that the reverse is not true in general. In combination with the result we present here, this will then show that GSCMs are indeed a (strict) generalization of SCMs.

Lemma 2 shows how, for any real-valued SCM, we can construct a GSCM which has the same solutions under all interventions[5].

Lemma 2. *Let $\mathcal{M}^{SCM} = (\boldsymbol{\mathcal{X}}, F, \mathbb{P}^{\mathcal{E}})$ be a real-valued SCM. The GSCM $\mathcal{M}^{GSCM} = (\boldsymbol{\mathcal{X}}, G, \mathbb{P}^{\mathcal{E}})$ with causal constraints G,*

$$
g_j = f_j(\boldsymbol{x}_{pa(j)}, \boldsymbol{e}_{pa(j)}) - x_j, \quad A_j = \mathcal{P}(\mathbb{I}_{\boldsymbol{X}} \setminus j) \quad \forall j \in \mathbb{I}_{\boldsymbol{X}},
$$

has the same solutions as the SCM under any intervention.

Proof. The proof follows from Definitions 4 and 5. □

We give an example of the construction of a GSCM from an SCM for the system in Example 3.

Example 4. Consider the system of price, supply, and demand in Example 3. Suppose we try to capture the same system in an SCM $(\mathbb{R}^3, F, \mathbb{P}_{\mathbb{R}})$ and end up with the following structural equations[6]

$$
\begin{aligned}
P &= P + S - D \\
D &= D \\
S &= f_S(P, E).
\end{aligned}
$$

Applying the construction in Lemma 2 we find the GSCM in Example 3 with the extra causal constraint $(D - D = 0, \{\emptyset, \{S\}, \{P\}, \{S,P\}\})$. The generalized structural equation $D - D = 0$ imposes no extra restrictions on the solutions of the model under any possible intervention. In a GSCM such redundant constraints can be discarded.

---

[5]The more general case, where variables take value in an arbitrary standard measurable space, requires a trivial extension of GSCMs.

[6]Note that the second equation imposes no restrictions on the solutions to this SCM since any random variable D^* satisfies $D^* = D^*$ almost surely. Since an SCM needs a structural equation for each random variable, we need to include this in our model description.

3 EQUILIBRIUM STATES OF DYNAMICAL SYSTEMS

In this section we consider the equilibrium states of dynamical systems with initial conditions. We show that the equilibrium behavior of this class of systems can generally not be fully described using an ordinary structural causal model. We show that the causal semantics of the equilibria are *completely* represented in the more flexible framework of GSCMs.

We start with a concise introduction to the dynamical system that we consider. Before we present our theoretical results, we illustrate the advantages of using a generalized structural causal model to describe equilibrium states on an example from biochemistry.

3.1 DYNAMICAL SYSTEMS

We consider dynamical systems describing n variables $\boldsymbol{X}$ taking value in $\mathbb{R}^n$ by a set of coupled first-order ordinary differential equations (ODE). We assume that the initial conditions $\boldsymbol{X}^0$ and parameters $\boldsymbol{K}$ associated with the ODEs are random variables taking value in $\mathbb{R}^n$ and $\mathbb{R}^k$ respectively that do *not* depend on time.

A dynamical system $\mathcal{D}$ for variables $\boldsymbol{X}$ and exogenous random variables $\boldsymbol{E} = (\boldsymbol{K}, \boldsymbol{X}^0)$ is specified as follows [14]:

$$\frac{dX_i(t)}{dt} = f_i(\boldsymbol{X}_{\mathrm{pa}(i)}(t), \boldsymbol{E}_{\mathrm{pa}(i)}), \qquad \forall i \in \mathbb{I}_{\boldsymbol{X}}$$
$$X_i(0) = X_i^0, \qquad \forall i \in \mathbb{I}_{\boldsymbol{X}},$$

where $\boldsymbol{f}$ is a continuous and measurable function. We let $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$ be a probability distribution over the exogenous variables. For $\boldsymbol{e} \in \mathbb{R}^n$, the solution $\boldsymbol{X}(t, \boldsymbol{e})$ to the initial value problem is given by the integral equation [14]

$$\boldsymbol{X}(t, \boldsymbol{e}) = \boldsymbol{X}(0, \boldsymbol{e}) + \int_0^t \boldsymbol{f}\left(\boldsymbol{X}(s, \boldsymbol{e}), \boldsymbol{e}\right) ds.$$

We say that a dynamical system converges to a random variable $\boldsymbol{X}^*$ if for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e} \in \mathcal{E}$,

$$\lim_{t \to \infty} \boldsymbol{X}(t, \boldsymbol{e}) \to \boldsymbol{X}^*(\boldsymbol{e}).$$

This implies a unique distribution $\mathbb{P}_{\mathrm{do}(\emptyset)}^{\boldsymbol{\mathcal{X}}}$.[7]

There are various ways to define an intervention that targets $I \subseteq \boldsymbol{X}$. We could, for instance, fix the entire sample path of target variables I from $t = 0$ to $t = \infty$ [6].

---

[7]Note that ODEs may have multiple equilibrium states, but since we assume that the initial conditions are specified, the equilibrium states of the corresponding initial-value problems are unique under mild assumptions.
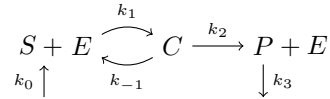
Another option is the push-and-let-go intervention [1], where the targeted variables are forced to take on a specific value at $t = 0$ only. Here we follow [6] and consider interventions that keep the value of the targeted variables fixed to a constant value until the system converges to an equilibrium.

An intervention $\mathrm{do}(I, \boldsymbol{\xi}_I)$ where $I \in \mathcal{P}(\mathbb{I}_{\boldsymbol{X}})$ and $\boldsymbol{\xi}_I \in \mathbb{R}_I^n$ then results in the *intervened* dynamical system $\mathcal{D}_{\mathrm{do}(I, \boldsymbol{\xi}_I)}$ specified by

$$\frac{dX_i(t)}{dt} = 0, \qquad X_i(0) = \xi_i, \ \forall i \in I$$
$$\frac{dX_i(t)}{dt} = f_i(\boldsymbol{X}_{\mathrm{pa}(i)}(t), \boldsymbol{E}_{\mathrm{pa}(i)}), \quad X_i(0) = X_i^0, \ \forall i \in \mathbb{I}_{\boldsymbol{X}} \backslash I$$

3.2 THE BASIC ENZYME REACTION

Enzymes play an important role in biochemical reactions that regulate biological processes in living organisms. A well-known example in biochemistry is the basic enzyme reaction [15]. In this system a substrate S reacts with an enzyme E to form a complex C which is then converted into a product P and the enzyme. For the *open* enzyme reaction a constant influx of substrate and an efflux of product are added [16]. The process can then be presented by the following reaction graph,

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C \xrightarrow{k_2} P + E$$

with $k_0 \uparrow$ on the left and $\downarrow k_3$ on the right,

where $\boldsymbol{k} = [k_0, k_{-1}, k_1, k_2, k_3]$ are the *rate* parameters of the system [15, 16].

The concentrations are $\boldsymbol{X} = (S, C, E, P) \in \mathbb{R}_+$ in a system[8] are strictly positive and may change over time. The law of mass-action states that the rate of a reaction is proportional to the product of the concentration of the reactants [15]. Applying the law of mass-action to the basic enzyme reaction, we obtain the following dynamical system,

$$\dot{S}(t) = k_0 - k_1 S(t) E(t) + k_{-1} C(t), \qquad (5)$$
$$\dot{E}(t) = -k_1 S(t) E(t) + (k_{-1} + k_2) C(t), \qquad (6)$$
$$\dot{C}(t) = k_1 S(t) E(t) - (k_{-1} + k_2) C(t), \qquad (7)$$
$$\dot{P}(t) = k_2 C(t) - k_3 P(t), \qquad (8)$$
$$S(0) = s_0, \quad E(0) = e_0, \quad C(0) = c_0, \quad P(0) = p_0, \qquad (9)$$

where $\boldsymbol{x}_0 = (s_0, e_0, c_0, p_0)$ are the *initial conditions* of the system.

---

[8]For the remainder of this paper we consider the concentrations of the molecules and these will be denoted by the same symbol as the molecules.

To get an idea of the stationary behavior of this system, we simulated the system in (5) to (9) with random initial conditions. The time dependence of the concentrations over time is shown in Figure 1a for 10 different initial conditions. This plot shows that a) the concentration of each reactant converges to an equilibrium b) the equilibrium state of S and E depends on the initial conditions and c) the equilibrium state of C and P does not depend on the initial conditions.

By explicit calculation one can verify that given strictly positive initial conditions $x_0 \in \mathbb{R}_+^4$ and rate parameters $\boldsymbol{k} \in \mathbb{R}_+^5$, the (intervened) dynamical system converges to an equilibrium $\boldsymbol{X}^* \in \mathbb{R}_+^4$ if it exists (see [16] and supplementary material for details). For simplicity we assume that the system is deterministic (i.e. the initial conditions and rate parameters are non-random). The stochastic setting, where initial conditions and rate parameters are sampled from a distribution, can be obtained as a trivial extension.

3.2.1 Equilibrium states

In this section we look at the equilibrium states of the basic enzyme reaction, and how they can be derived from the equations of motion and the constants of motion of the corresponding dynamical system. In the next section we will demonstrate how the causal constraints in a GSCM representation of the equilibrium states can be derived naturally from these equations and constants of motion.

There are many ways to find the equilibrium states of a dynamical system. One approach is to observe that, at equilibrium, the system is at rest and all time derivatives (i.e. equations of motion) are equal to zero. Subsequently one can look for *constants of motion*, i.e. functions of the variables that are time-invariant. A solution to the resulting system of equations is an equilibrium state.

The observed dynamical system in (5) to (9) has four equations of motion, which we set to zero. For example, (5) yields the equilibrium equation

$$k_0 - k_1 S(t)E(t) + k_{-1}C(t) = 0.$$

The system also admits a constant of motion. Since $\dot{C}(t) + \dot{E}(t) = 0$ for all t, we have that

$$C(t) + E(t) = c_0 + e_0 \qquad \forall\, t. \qquad (10)$$

Solving the resulting system of five equations one can find a unique equilibrium solution for the observational system, see also Table 1. Note that when $e_0 + c_0 - \frac{k_0}{k_2} \leq 0$ the system has no solution, because the concentrations need to be positive.

We also consider the constants of motion in the intervened systems:

Table 1: Solutions to the intervened dynamical system of the basic enzyme reaction in (5) to (9) under various interventions, where $y = \frac{1}{2}\sqrt{(e_0 - s_0)^2 + 4\frac{k_0(k_{-1}+k_2)}{k_1 k_2}}$.[9]

I	S^*	C^*	E^*
$\emptyset$	$\frac{k_0 + k_{-1}\frac{k_0}{k_2}}{k_1(e_0 + c_0 - \frac{k_0}{k_2})}$	$\frac{k_0}{k_2}$	$e_0 + c_0 - \frac{k_0}{k_2}$
$S = \xi_s$	ξ_s	$\frac{k_1\xi_s(e_0+c_0)}{k_{-1}+k_2+k_1\xi_s}$	$\frac{(k_{-1}+k_2)(e_0+c_0)}{k_{-1}+k_2+k_1\xi_s}$
$C = \frac{k_0}{k_2}$	$\frac{(e_0 - s_0)}{2} + y$	$\frac{k_0}{k_2}$	$\frac{-(e_0 - s_0)}{2} + y$
$E = \xi_e$	$\frac{k_0 + k_{-1}\frac{k_0}{k_2}}{k_1\xi_e}$	$\frac{k_0}{k_2}$	ξ_e

1 An intervention $\mathrm{do}(I, \boldsymbol{\xi}_I)$ on variables $I \in \mathbb{I}_{\boldsymbol{X}}$ results in $\dot{\boldsymbol{X}}_I(t) = 0$ and thus in a new constant of motion $\boldsymbol{X}_I(t) = \boldsymbol{\xi}_I$.
2 Since an intervention on E results in a differential equation $\dot{E}(t) = 0$, and an intervention on C results in a differential equation $\dot{C}(t) = 0$, the constant of motion in (10) no longer holds if either E, C or both are targeted by an intervention.
3 The system under the intervention $\mathrm{do}(C = \frac{k_0}{k_2})$ implies the following constant of motion,

$$S(t) - E(t) = s_0 - e_0 \qquad \forall\, t, \qquad (11)$$

since $\dot{S}(t) - \dot{E}(t) = 0$ in that case. Similar to the constant of motion in (10) this no longer holds when either S, E, or both are targeted by an intervention.

We obtain the equilibrium solutions in Table 1, by solving the system of equations for each intervention.

Note that the intervention $\mathrm{do}(C = \xi_C)$ and the equations that we obtain by setting (5) and (6) to zero imply that

$$k_1 S(t)E(t) = k_0 + k_{-1}\xi_C = (k_{-1} + k_2)\xi_C.$$

This equation can only hold when $\xi_C = \frac{k_0}{k_2}$, otherwise the intervened dynamical system does not have an equilibrium solution.

From Table 1 we can also observe the rich causal semantics of the equilibrium states of the basic enzyme reaction. For example, an intervention on S makes C^* dependent on the initial conditions and an intervention on E makes S^* independent of the initial conditions. We simulated the intervened dynamical system under intervention on S and E. The change in dependencies on the initial conditions can be seen clearly from the simulation results in Figures 1b and 1c respectively.

3.2.2 GSCM representation

In this section we present a GSCM representation of the basic enzyme reaction that captures *all* causal semantics of its stationary behavior, given the initial conditions
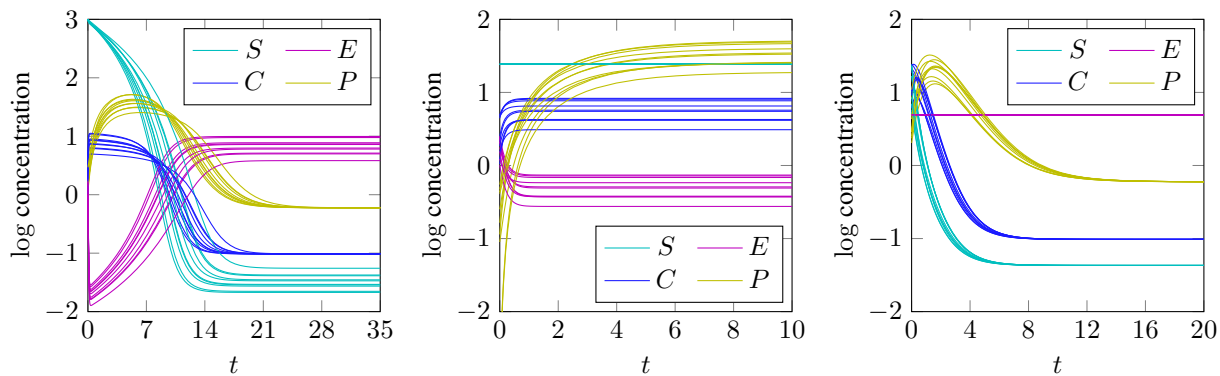
(a) $S, C, E,$ and P converge to an equilibrium in the observational system, where E^* and S^* depend on initial conditions.

(b) $C, E,$ and P converge to an equilibrium that depends on the initial conditions after an intervention on S.

(c) $S, C,$ and P converge to an equilibrium that is independent of the initial conditions after an intervention E.

Figure 1: Temporal dependence of the concentrations in the basic enzyme reaction in (5) to (9) with random initial conditions and $\boldsymbol{k} = [0.4, 0.3, 1.0, 1.1, 0.5]$. Many other choices for the rate parameters give qualitatively similar results.

of the system. We show that the problem of finding an appropriate GSCM representation reduces to finding the constants of motion that the system admits under all interventions such that the equilibrium states are implied.

At equilibrium, the concentration of $\boldsymbol{X}$ is constant. This means that the equations of motion in (5) to (9) for $S, C, E,$ and P must be equal to zero, unless the variable itself is intervened upon. This leads to the causal constraints in (12) to (15) below.

The constant of motion in (10) restricts the system unless either E or C is intervened upon, which corresponds to the causal constraint in (16) below. The constant of motion in (11) only holds when C is intervened upon and S, E are not, which is captured by the causal constraint in (17) below. The causal constraints (12) and (14) ensure that the model has no solution under the intervention $\mathrm{do}(C, \xi_C)$ when $\xi_C \neq \frac{k_0}{k_2}$.

$$k_0 + k_{-1}C - k_1 SE = 0 \quad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus S}), \quad (12)$$
$$k_1 SE - (k_{-1} + k_2)C = 0 \quad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus C}), \quad (13)$$
$$-k_1 SE + (k_{-1} + k_2)C = 0 \quad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus E}), \quad (14)$$
$$k_2 C - k_3 P = 0 \quad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus P}), \quad (15)$$
$$C + E - (c_0 + e_0) = 0 \quad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus\{C,E\}}), \quad (16)$$
$$S - E - (s_0 - e_0) = 0 \quad \{\mathbb{I}_C, (\mathbb{I}_C \cup \mathbb{I}_P)\}. \quad (17)$$

The GSCM for the basic enzyme reaction is then $\mathcal{M} = (\mathbb{R}_+^4, G, \mathbb{P}_\emptyset)$ where G consists of the causal constraints in equations (12) to (17). Recall that given strictly positive rate parameters and initial conditions the system converges to an equilibrium state if it exists, under any intervention. The solutions of the GSCM that we constructed

coincide with the equilibrium states of the dynamical system (see Table 1) under any intervention, since the equations constraining its solution space are identical to the ones that characterize the equilibrium states. Hence the GSCM fully captures the causal semantics pertaining to the equilibrium behavior of the basic enzyme reaction.

Remark 1. The example of the basic enzyme reaction also illustrates the crucial difference between interventions that fix concentrations from $t = 0$ to $t = \infty$, and interventions that only fix it at $t = 0$. In the latter case, there is no effect on the dynamics of the system and the causal constraints in (12) to (15) hold under all such interventions, and the only effect of interventions is that they change the initial state $\boldsymbol{x}_0$. Hence, it would only uncover the (causal) dependence of the equilibrium state on the initial conditions of the system.

3.2.3 What about SCMs?

In this section we show that contrary to a GSCM, an SCM cannot give a *complete* description of the equilibrium states of the basic enzyme reaction.

The construction of the GSCM in the previous section gives a complete representation of the equilibria in the basic enzyme reaction. To construct an SCM from the ODEs we could follow [6] and derive it by setting the equations of motion in the dynamical system to zero, yielding an SCM with structural equations (12)-(15). This would not incorporate the dependence of the equilibrium state on the initial conditions that we observed in the simulations of the observational and intervened dynamical systems. Therefore such an SCM would be *un-*

derspecified.

Theorem 1 below shows that there does not exist an SCM representation of the basic enzyme reaction with initial conditions. This shows that, if we wish to model the equilibrium states of dynamical systems and pertain all causal semantics, we cannot generally use an SCM for this task.

Theorem 1. *The equilibrium states of the basic enzyme reaction cannot be represented by an SCM.*

Proof. From Table 1 it can be seen that the solution $(S_\emptyset^*, E_\emptyset^*)$ in the observational system is different from the solution $(S_{\text{do}(C=\xi_C)}^*, E_{\text{do}(C=\xi_C)}^*$ in the system after an intervention that targets C and sets it equal to $\frac{k_0}{k_2}$ (i.e., the same value as the equilibrium value for C in the observational setting). Clearly, this behavior cannot be captured by an SCM. $\square$

Interestingly, if we would treat the equilibrium state of C as a latent endogenous variable that cannot be intervened upon, there exists a complete SCM representation of the causal semantics of the system with initial conditions. Using Lemma 2 to represent it as a GSCM, it takes the form

$$\frac{k_0 + k_{-1}\frac{k_0}{k_2}}{k_1 E} - S = 0, \qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X'}\setminus S}),$$

$$\frac{(k_{-1}+k_2)(c_0+e_0)}{k_{-1}+k_2+k_1 S} - E = 0, \qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X'}\setminus E}),$$

$$\frac{k_2}{k_3}\frac{k_1 SE}{k_{-1}+k_2} - P = 0, \qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X'}\setminus P}).$$

This SCM can be analyzed using existing methods, see [8, 9].

3.3 EQUILIBRIUM CAUSAL MODELS

We have shown that the equilibrium solutions of dynamical systems change under intervention, and that for a realistic example an SCM is not sufficiently flexible to capture these changes. A GSCM, on the other hand, can represent the equilibrium distributions of dynamical systems under interventions, as is shown in Theorem 2 below.

Theorem 2. *Let $\mathcal{D}$ be a dynamical system describing n variables $\boldsymbol{X}$ taking value in $\mathbb{R}^n$ with initial values $\boldsymbol{X}_0$ taking value in $\mathbb{R}^n$, parameters $\boldsymbol{K}$ taking value in $\mathbb{R}^m$ and $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$ a probability distribution over the space of exogenous variables $\boldsymbol{E} = (\boldsymbol{K}, \boldsymbol{X}^0)$. Then there exists a GSCM $\mathcal{M}(\mathcal{D})$ so that for $I \in \mathcal{P}(\mathbb{I}_{\boldsymbol{X}})$ and $\boldsymbol{\xi}_I \in \mathbb{R}_I^n$:*

- *If $\mathcal{D}_{do(I,\boldsymbol{\xi}_I)}$ converges to an equilibrium $\boldsymbol{X}^*(I,\boldsymbol{\xi}_I)$ then $\big(\mathcal{M}(\mathcal{D})\big)_{do(I,\boldsymbol{\xi}_I)}$ has a solution, and all its solutions have the same distribution as $\boldsymbol{X}^*(I,\boldsymbol{\xi}_I)$.*

- *If $\mathcal{D}_{do(I,\boldsymbol{\xi}_I)}$ does not converge to an equilibrium then $\big(\mathcal{M}(\mathcal{D})\big)_{do(I,\boldsymbol{\xi}_I)}$ has no solutions.*
- *The following diagram commutes:*

$$
\begin{array}{ccc}
\mathcal{D} & \longmapsto & \mathcal{M}(\mathcal{D}) \\
\big\downarrow & & \big\downarrow \\
\mathcal{D}_{do(I,\boldsymbol{\xi}_I)} & \longmapsto & \big(\mathcal{M}(\mathcal{D})\big)_{do(I,\boldsymbol{\xi}_I)}
\end{array}
$$

Proof. Consider an intervention target $I \subseteq \mathcal{P}(\mathbb{I}_{\boldsymbol{X}})$, and let $\boldsymbol{C}_I$ be the set of all intervention values so that for $\boldsymbol{\xi}_I \in \boldsymbol{C}_I$ the dynamical system $\mathcal{D}_{\text{do}(I,\boldsymbol{\xi}_I)}$ converges to an equilibrium $\boldsymbol{X}^*(I,\boldsymbol{\xi}_I)$. Consider the generalized structural equation,

$$\boldsymbol{X}^*(I,\boldsymbol{X}_I)\mathbf{1}_{\boldsymbol{C}_I}(\boldsymbol{X}_I)+(\boldsymbol{X}+1)\big(1-\mathbf{1}_{\boldsymbol{C}_I}(\boldsymbol{X}_I)\big)-\boldsymbol{X} = \mathbf{0}.$$

If $\boldsymbol{X}_I \notin \boldsymbol{C}_I$ then this equation yields a contradiction. Under the intervention $\text{do}(I,\boldsymbol{\xi}_I)$, $\boldsymbol{X}_I = \boldsymbol{\xi}_I$. In that case, if $\boldsymbol{\xi}_I \in \boldsymbol{C}_I$ then $\boldsymbol{X} = \boldsymbol{X}^*(I,\boldsymbol{\xi}_I)$ is the only solution to these equations. Since $\boldsymbol{C}_I$ is a measurable set, the mapping $\boldsymbol{g}_I$ corresponding to the generalized structural equation is measurable.

The GSCM $\mathcal{M}(\mathcal{D}) := (\boldsymbol{\mathcal{X}}, G, \mathbb{P}_{\boldsymbol{\mathcal{E}}})$ where G consists of the causal constraints $\{(\boldsymbol{g}_I, A_I = \{I\}) : I \in \mathcal{P}(\mathbb{I}_{\boldsymbol{X}})\}$ satisfies the properties of the theorem by construction and by Lemma 1. $\square$

The result in Theorem 2 proves that a GSCM representation always *exists* for the equilibrium states of a dynamical system. Although the GSCM that we construct in the proof captures the causal semantics of the equilibrium states, it does not give a parsimonious representation of the system. In the previous section we sketched an alternative construction method, and showed how a parsimonious GSCM can be derived from the equations and constants of motion in the basic enzyme reaction.[10]

4 FUNCTIONAL LAWS

In this section, we further motivate the use of GSCMs by applying it to systems that are described by a *functional law*, which is a relation between variables that must hold under *all* interventions. We show that the causal constraint describing such a relation has *more* invariance under interventions than an ordinary structural equation, and that the advantage of a GSCM representation of a functional law compared to an SCM representation is two-fold. First of all its solutions never violate the

---

[10]In the supplementary material we apply this construction to two more dynamical systems (Lotka-Volterra model, chemical reaction network).

functional law under interventions, and second it gives a much more parsimonious description.

A functional law describes a relation between endogenous and possibly exogenous variables ($\boldsymbol{X}' \subseteq \boldsymbol{X}, \boldsymbol{E}' \subseteq \boldsymbol{E}$) that is invariant under any intervention on variables $\boldsymbol{X}$ in the system. Such a relation can be described by a causal constraint of the form

$$g(\boldsymbol{X}', \boldsymbol{E}') = 0, \qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}})$$

This ensures that the constraint is active under any intervention on the variables $\boldsymbol{X}$ in the system. Hence any solution of the intervened GSCM satisfies the functional law. We will demonstrate the advantage of this representation in Example 5.

Example 5. It is well-known that the resistance $R \in \mathbb{R}$, the potential $V \in \mathbb{R}$, and the current $I \in \mathbb{R}$ in an electrical circuit are related by Ohm's law (i.e. $V = IR$). Let $\boldsymbol{X} = (R, V, I)$ denote the variables in Ohm's law. In absence of any knowledge about the environment, this system can be represented by the GSCM $\mathcal{M} = (\boldsymbol{\mathcal{X}}, G, \mathbb{P}_{\emptyset})$, where the G consists of the causal constraint,

$$V - IR = 0, \qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}}). \qquad (18)$$

If we were to describe the same system using an SCM, then we would need the following three copies of the causal constraint in (18) as structural equations,

$$V = IR$$
$$I = \frac{V}{R}$$
$$R = \frac{V}{I}.$$

Indeed, considering interventions on two out of three variables leaves one with no choice for the structural equation of the third one. Furthermore, a simultaneous intervention on R, V, and I always has a solution in the SCM representation, even when this means that Ohm's law is violated. The GSCM representation typically does not have a solution under such an intervention (unless the target values satisfy the constraint).

Note that a functional law can be any relation that is invariant under all interventions. For example, a transformation of a (set of) variables describing a system to another (set of) variables describing the same system can be modeled as a functional law. We will illustrate such a transformation on the description of an ideal gas in Example 6.

Example 6. An ideal gas can be described at different levels. On the one hand, one can look at the *microscopic* level and model the speed of each particle in the system. On the other hand, it is perhaps more convenient

to model *macroscopic* quantities such as pressure P and temperature T. Physicists have shown that, for N particles of mass m in a fixed cuboid volume V, the mean squared speed $\bar{v}^2$ of the particles can be related to pressure and temperature by[11]

$$P = \frac{Nm\bar{v}^2}{3V}, \qquad (19)$$

$$T = \frac{m\bar{v}^2}{3k_B}. \qquad (20)$$

These *variable transformations* from a microscopic to a macroscopic description $\boldsymbol{X} = (P, T, \bar{v})$ of the ideal gas can be captured by the GSCM $\mathcal{M} = (\boldsymbol{\mathcal{X}}, G, \mathbb{P}_{\emptyset})$ where G consists of

$$3PV - Nm\bar{v}^2 = 0 \qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}}),$$
$$3k_B T - m\bar{v}^2 = 0 \qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}}),$$

and N and V are treated as constants. The solutions to this GSCM will always satisfy the ideal gas law (i.e. $PV = Nk_b T$). In an SCM representation of this transformed system, first of all we would need to include the structural equations in (19) and (20). For $\bar{v}^2$ we need to make sure that it is *both* a function of P and T, for instance by enforcing the constraints by a self-loop,

$$\bar{v}^2 = \bar{v}^2 + \mathbf{1}_{\{3PV - Nm\bar{v}^2 \neq 0\}} + \mathbf{1}_{\{3k_B T - m\bar{v}^2 \neq 0\}}.$$

Note that if we would treat $\bar{v}$ as a latent variable, the system can be described by a single causal constraint

$$PV - Nk_B T = 0, \qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}}).$$

In analogy to the system in Example 5, if we would try to capture this in an SCM, we would end up with multiple copies of this constraint, and a simultaneous intervention on P and T would typically lead to a solution that does not satisfy the ideal gas law. Therefore, the GSCM representation of functional laws like the ideal gas law is more parsimonious, more natural, and more accurate than any SCM representation.

5 CONCLUSION

While structural causal models form a very popular modeling framework in many applied sciences, we have shown that they are neither powerful enough to model the rich equilibrium behavior of simple dynamical systems such as the basic enzyme reaction, nor simple functional laws of nature like the ideal gas law. In order to represent the causal semantics of such systems, we introduced *generalized* structural causal models and proved

---

[11]This can be shown by using the kinetic theory of gases and the equipartition theorem [17].

that they do completely capture the stationary behavior of dynamical systems that converge to equilibrium and causal semantics of functional laws.

The main idea of generalized structural causal models is that they are more flexible than SCMs because they specify a set of probability distributions on the space of endogenous variables by causal constraints instead of by structural equations, where the validity of a causal constraint after perfect interventions can be more flexibly specified than the validity of structural equations.

We illustrated the power and flexibility of our approach on the simple basic enzyme reaction, one of the fundamental building blocks of biochemistry. We showed that this system has rich stationary behavior and a dependence on initial values that cannot be represented by an SCM. We demonstrated how the causal constraints can be derived naturally from the equations and constants of motion of the corresponding ODEs and initial conditions, thereby completely representing the system's stationary behavior by a parsimonious GSCM. We used our framework to identify a non-trivial marginal model which is equivalent to an SCM and can be analyzed using existing methods.

We further motivated generalized structural models by considering their application to systems which can be described by functional laws such as the ideal gas law. We pointed out that any solution of an (intervened) GSCM that includes the ideal gas law, must satisfy this law. We highlighted the difference with SCMs, which may have solutions that do not satisfy the ideal gas law.

We believe that the examples we have presented here form a compelling motivation to introduce GSCMs and to investigate the properties of these causal models in more detail in future work. We can only speculate on the impact of this work, but we are left wondering whether the standard starting point in causal discovery—that the data-generating process can be accurately modeled with an SCM—is tenable in the context of biochemical systems, considering that even very simple biochemical systems (a single enzyme reaction) already violate this assumption. An interesting starting point for future work would be to investigate how and under which conditions GSCMs can be learned from observational and interventional data.

Acknowledgements

References

[1] M. Weber, "On the Incompatibility of Dynamical Biological Mechanisms and Causal Graphs," *Philosophy of Science*, vol. 83, no. 5, pp. 959–971, 2016.

[2] Y. Iwasaki and H. A. Simon, "Causality and model abstraction," *Artificial Intelligence*, vol. 67, no. 1, pp. 143–194, 1994.

[3] M. Voortman, D. Dash, and M. J. Druzdzel, "Learning why things change: The difference-based causality learner," in *Proceedings of the Twenty-Sixth Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, 2010.

[4] F. M. Fisher, "A Correspondence Principle for Simultaneous Equation Models," *Econometrica*, vol. 38, no. 1, pp. 73–92, 1970.

[5] D. Dash, "Restructuring Dynamic Causal Systems in Equilibrium," in *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics (AISTATS 2005)*, 2005.

[6] J. M. Mooij, D. Janzing, and B. Schölkopf, "From Ordinary Differential Equations to Structural Causal Models: the deterministic case," in *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI-13)*, pp. 440–448, 2013.

[7] J. Pearl, *Causality : models, reasoning, and inference*. Cambridge University Press, 2000.

[8] S. Bongers, J. Peters, B. Schölkopf, and J. M. Mooij, "Structural causal models: Cycles, marginalizations, exogenous reparametrizations and reductions," *arXiv.org preprint*, vol. arXiv:1611.06221 [stat.ME], Nov. 2016.

[9] P. Forré and J. M. Mooij, "Markov properties for graphical models with cycles and latent variables," *arXiv.org preprint*, vol. arXiv:1710.08775 [math.ST], 2017.

[10] A. Hyttinen, F. Eberhardt, and P. O. Hoyer, "Learning Linear Cyclic Causal Models with Latent Variables," *Journal of Machine Learning Research*, vol. 13, no. 1, pp. 3387–3439, 2001.

[11] G. Lacerda, P. L. Spirtes, J. Ramsey, and P. O. Hoyer, "Discovering Cyclic Causal Models by Independent Components Analysis," *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence*, 2008.

[12] J. M. Mooij, D. Janzing, T. Heskes, and B. Schölkopf, "On causal discovery with cyclic additive noise models," *Advances in Neural Information Processing Systems*, pp. 639–647, 2011.

[13] P. K. Rubenstein, S. Weichwald, S. Bongers, J. M. Mooij, D. Janzing, M. Grosse-Wentrup, and B. Schölkopf, "Causal Consistency of Structural Equation Models," in *Proceedings of the 33rd Annual Conference on Uncertainty in Artificial Intelligence (UAI-17)*, 2017.

[14] X. Han and P. E. Kloeden, *Random Ordinary Differential Equations and Their Numerical Solution*. Springer Singapore, 2017.

[15] J. D. Murray, *Mathematical Biology I: An Introduction*. Springer-Verlag New York, third edit ed., 2002.

[16] I. Belgacem and J.-L. Gouzé, "Global Stability of Full Open Reversible Michaelis-Menten Reactions," *IFAC Proceedings Volumes*, vol. 45, no. 15, pp. 591–596, 2012.

[17] S. J. Blundell and K. M. Blundell, *Concepts in Thermal Physics*. Oxford University Press, oct 2009.

Supplementary Material

A Basic Enzyme Reaction

In this section we show the additional results, concerning the basic enzyme reaction, that were discussed in the main paper. First we discuss the fixed points of the basic enzyme reaction. Then we show that the systems converges to its fixed point whenever it exists. Finally, we derive a simple marginal model from the GSCM representation of the basic enzyme reaction.

A.1 Fixed points

The fixed points of the basic enzyme reaction, for all intervened systems, are given in Table 2. For any intervention, these are obtained by solving the system of equations that one gets by considering the causal constraints in the GSCM in (12) to (17) that are active under that specific intervention. That is, we take all equations for which the intervention is in the activation set.

Table 2: Fixed points of the basic enzyme reaction, where $y = \frac{1}{2}\sqrt{(e_0 - s_0)^2 + 4\frac{k_0(k_{-1}+k_2)}{k_1 k_2}}$.

intervention	S	C	E	P
none	$\frac{k_0+k_{-1}\frac{k_0}{k_2}}{k_1(e_0+c_0-\frac{k_0}{k_2})}$	$\frac{k_0}{k_2}$	$e_0 + c_0 - \frac{k_0}{k_2}$	$\frac{k_0}{k_3}$
$\mathrm{do}(S = s)$	s	$\frac{k_1 s(e_0+c_0)}{k_{-1}+k_2+k_1 s}$	$\frac{(k_{-1}+k_2)(e_0+c_0)}{k_{-1}+k_2+k_1 s}$	$\frac{k_2}{k_3}\frac{k_1 s(e_0+c_0)}{k_{-1}+k_2+k_1 s}$
$\mathrm{do}(C = c),\ c = \frac{k_0}{k_2}$	$\frac{(s_0-e_0)}{2}+y$	c	$\frac{-(s_0-e_0)}{2}+y$	$\frac{k_2}{k_3}c$
$\mathrm{do}(C = c),\ c \neq \frac{k_0}{k_2}$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$\mathrm{do}(E = e)$	$\frac{k_0+k_{-1}\frac{k_0}{k_2}}{k_1 e}$	$\frac{k_0}{k_2}$	e	$\frac{k_0}{k_3}$
$\mathrm{do}(P = p)$	$\frac{k_0+k_{-1}\frac{k_0}{k_2}}{k_1(e_0+c_0-\frac{k_0}{k_2})}$	$\frac{k_0}{k_2}$	$e_0 + c_0 - \frac{k_0}{k_2}$	p
$\mathrm{do}(S = s, C = c)$	s	c	$\frac{k_{-1}+k_2}{k_1}\frac{c}{s}$	$\frac{k_2}{k_3}c$
$\mathrm{do}(S = s, E = e)$	s	$\frac{k_1}{k_{-1}+k_2}se$	e	$\frac{k_2}{k_3}\frac{k_1}{k_{-1}+k_2}se$
$\mathrm{do}(S = s, P = p)$	s	$\frac{k_1 s(e_0+c_0)}{k_{-1}+k_2+k_1 s}$	$\frac{(k_{-1}+k_2)(e_0+c_0)}{k_{-1}+k_2+k_1 s}$	p
$\mathrm{do}(C = c, E = e)$	$\frac{k_0+k_{-1}c}{k_1 e}$	c	e	$\frac{k_2}{k_3}c$
$\mathrm{do}(C = c, P = p),\ c = \frac{k_0}{k_2}$	$\frac{(s_0-e_0)}{2}+y$	c	$\frac{-(s_0-e_0)}{2}+y$	p
$\mathrm{do}(C = c, P = p),\ c \neq \frac{k_0}{k_2}$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$\mathrm{do}(E = e, P = p)$	$\frac{k_0+k_{-1}\frac{k_0}{k_2}}{k_1 e}$	$\frac{k_0}{k_2}$	e	p
$\mathrm{do}(S = s, C = c, E = e)$	s	c	e	$\frac{k_2}{k_3}c$
$\mathrm{do}(S = s, C = c, P = p)$	s	c	$\frac{k_{-1}+k_2}{k_1}\frac{c}{s}$	p
$\mathrm{do}(S = s, E = e, P = p)$	s	$\frac{k_1}{k_{-1}+k_2}se$	e	p
$\mathrm{do}(C = c, E = e, P = p)$	$\frac{k_0+k_{-1}c}{k_1 e}$	c	e	p
$\mathrm{do}(S = s, C = c, E = e, P = p)$	s	c	e	p

A.2 Convergence results for the basic enzyme reaction

In this section, we show that the basic enzyme reaction always converges to its fixed point, as long as it exists. We also show that the intervened basic enzyme reaction has the same property. To prove this result we rely on both explicit calculations and a convergence property of so-called cooperative systems that we obtained from [16]. To prove convergence for the observed system and the system after interventions on P and E, we use the latter technique. Convergence to the equilibrium solution after interventions on S and C can be shown by explicit calculation. The convergence results for combinations of interventions can be obtained by a trivial extension of the arguments that were used in the other cases.

A.2.1 Cooperativity in the basic enzyme reaction

To show that the basic enzyme reaction converges to a unique equilibrium, if it exists, we first state a result that we obtained from [16]: cooperative systems as in Definition 6 have the attractive convergence property in Proposition 1.

Definition 6. A system of ODEs $\dot{X}$ is cooperative if the Jacobian matrix has non-negative off-diagonal elements, or there exists an integer k such that the Jabobian has $(k \times k)$ and $(n - k) \times (n - k)$ main diagonal matrices with nonnegative off-diagonal entries and the rectangular off-diagonal submatrices have non-positive entries.

Proposition 1. Let $\dot{X} = f(X)$ be a cooperative system with a fixed point x^*. If there exist two points $x_{\min}, x_{\max} \in \mathcal{X}$ such that $x_{\min} \leq x^* \leq x_{\max}$ and $f(x_{\min}) \geq 0$ and $f(x_{\max}) \leq 0$, then the hyperrectangle betweeen $x_{\min}$ and $x_{\max}$ is invariant[12] and for almost all initial conditions inside this rectangle the solution converges to x^*.

A.2.2 Convergence of the observed system

Recall that the dynamics of the basic enzyme reaction are given by

$$\dot{S}(t) = k_0 - k_1 S(t)E(t) + k_{-1}C(t), \tag{21}$$

$$\dot{E}(t) = -k_1 S(t)E(t) + (k_{-1} + k_2)C(t), \tag{22}$$

$$\dot{C}(t) = k_1 S(t)E(t) - (k_{-1} + k_2)C(t), \tag{23}$$

$$\dot{P}(t) = k_2 C(t) - k_3 P(t), \tag{24}$$

$$S(0) = s_0, \quad E(0) = e_0, \quad C(0) = c_0, \quad P(0) = p_0, \tag{25}$$

where $x_0 = (s_0, e_0, c_0, p_0)$ are the *initial conditions* of the system.

The analysis in [16] of the basic enzyme reaction makes use of Proposition 1, but also includes feedback from P to C. In this section, we repeat their analysis on our sligthly different model. Note that the arguments given in this section can also be applied to the system where P is intervened upon.

We start by rewriting the system of ODEs in equation (21) to (24), by using the fact that $\dot{E}(t) + \dot{C}(t) = 0$ so that $E(t) = e_0 + c_0 - C(t)$:

$$\dot{S}(t) = k_0 - k_1 S(t)(e_0 + c_0 - C(t)) + k_{-1}C(t), \tag{26}$$

$$\dot{C}(t) = k_1 S(t)(e_0 + c_0 - C(t)) - (k_{-1} + k_2)C(t), \tag{27}$$

$$\dot{P}(t) = k_2 C(t) - k_3 P(t). \tag{28}$$

Cooperativity The corresponding Jacobian matrix is given by,

$$J(S, C, P) = \begin{pmatrix} -k_1(e_0 + c_0 - C(t)) & k_{-1} + k_1 S(t) & 0 \\ k_1(e_0 + c_0 - C(t)) & -(k_{-1} + k_2) - k_1 S(t) & 0 \\ 0 & k_2 & -k_3 \end{pmatrix}. \tag{29}$$

Since all off-diagonal elements in the Jacobian matrix are nonnegative, the observational system is a cooperative system by Definition 6.

Convergence From Table 2 we find that the observed system has a unique (positive) fixed point as long as $e_0 + c_0 > \frac{k_0}{k_2}$. We want to use Proposition 1 to show that the system converges to this fixed point, so we need to find $x_{\min}$ and $x_{\max}$ so that all three derivatives are nonnegative and nonpositive respectively.

For $x_{\min} = (0, 0, 0)$, then $\dot{S} = k_0 > 0$ and $\dot{C} = \dot{P} = 0$ so all derivatives are nonnegative. The upper vertex must be

---

[12]An invariant set is a set with the property that once a trajectory of a dynamical set enters it, it cannot leave.

chosen so that all derivative are non-positive:

$$\dot{S} \leq 0 \iff S \geq \frac{k_0 + k_{-1}C}{k_1(e_0 + c_0 - C)},$$

$$\dot{C} \leq 0 \iff S \geq \frac{(k_{-1} + k_2)C}{k_1(e_0 + c_0 - C)},$$

$$\dot{P} \leq 0 \iff P \geq \frac{k_2}{k_3}C.$$

The basic enzyme reaction only has a fixed point as long as $C < e_0 + c_0$ (otherwise $\dot{S}(t) > 0$). If we let C approach $e_0 + c_0$, then the inequality constraints on the derivatives are satisfied as S and P go to infinity. More formally we can choose

$$\boldsymbol{x}_{\max} = \left(S = \max \left(\frac{k_0 + k_{-1}C}{k_1(e_0 + c_0 - C)}, \frac{(k_{-1} + k_2)C}{k_1(e_0 + c_0 - C)}\right), C = e_0 + c_0 - \epsilon, P = \frac{k_2}{k_3}C + \frac{1}{\epsilon}\right).$$

When ϵ approaches zero, both S and P go to infinity and all derivatives are nonpositive. Hence, by Proposition 1, the system converges to its fixed point for almost all valid initial values of S, C, and P (for which the fixed point exists).

A.2.3 Intervention on E

Similarly, we can also show that the system where E is targeted by an intervention that sets it equal to e, converges to the (unique) equilibrium in Table 2. The intervened system of ODEs is given by

$$\dot{S} = k_0 - k_1 e S + k_{-1}C,$$

$$\dot{C} = k_1 e S - (k_{-1} + k_2)C,$$

$$\dot{P} = k_2 C - k_3 P.$$

The Jacobian is given by

$$J(S, C, P) = \begin{pmatrix} -k_1 e & k_{-1} & 0 \\ k_1 e & -(k_{-1} + k_2) & 0 \\ 0 & k_2 & -k_3 \end{pmatrix}. \tag{30}$$

Since all off-diagonal elements are nonnegative this is a cooperative system by Definition 6.

All derivatives are nonnegative at the point $(S, C, P) = (0, 0, 0)$, and all derivatives are nonpositive at the point (s, c, p) where

$$s = \max \left(\frac{k_{-1}c + k_0}{k_1 e}, \frac{(k_{-1} + k_2)c}{k_1 e}\right),$$

$$p = \frac{k_2}{k_3}c,$$

where $c \to \infty$. We then apply Proposition 1 to show that the intervened system converges to the equilibrium value from all valid initial values.

A.2.4 Intervention on S

We show that the system converges to the equilibrium solution after an intervention on S by explicit calculation. The intervened system of ODEs is given by

$$\dot{S}(t) = 0,$$

$$\dot{E}(t) = -k_1 s E(t) + (k_{-1} + k_2)C(t),$$

$$\dot{C}(t) = k_1 s E(t) - (k_{-1} + k_2)C(t),$$

$$\dot{P}(t) = k_2 C(t) - k_3 P(t).$$

Since $\dot{C}(t) + \dot{E}(t) = 0$, we can write $E(t) = e_0 + c_0 - C(t)$, resulting in the following differential equation

$$\dot{C}(t) = k_1 s(e_0 + c_0 - C(t)) - (k_{-1} + k_2)C(t), \tag{31}$$
$$= -(k_1 s + k_{-1} + k_2)C(t) + k_1 s(e_0 + c_0). \tag{32}$$

We take the limit $t \to \infty$ of the solution to the initial value problem to obtain

$$C^* = \lim_{t \to \infty} \frac{k_1 s(e_0 + c_0)}{(k_1 s + k_1 + k_2)} + e^{-(k_1 s + k_{-1} + k_2)t} = \frac{k_1 s(e_0 + c_0)}{(k_1 s + k_{-1} + k_2)}. \tag{33}$$

The result for E follows from the fact that $E(t) = e_0 + c_0 - C(t)$. The result for P follows by explicitly solving the differential equation and taking the limit $t \to \infty$.

A.2.5   Intervention on C

There is no equilibrium solution when the intervention targeting C does not have value $\frac{k_0}{k_2}$, as can be seen from Table 2. To show that the system converges when the equilibrium solution exists, we can explicitly solve the initial value problem and take the limit $t \to \infty$. The intervened system of ODEs after an intervention $\text{do}(C = \frac{k_0}{k_2})$ is given by

$$\dot{S}(t) = -k_1 S(t)E(t) + (k_{-1} + k_2)\frac{k_0}{k_2} = -k_1 S(t)E(t) + k,$$

$$\dot{E}(t) = -k_1 S(t)E(t) + (k_{-1} + k_2)\frac{k_0}{k_2} = -k_1 S(t)E(t) + k,$$

$$\dot{C}(t) = 0,$$

$$\dot{P}(t) = k_0 - k_3 P(t),$$

where we set $k = (k_{-1} + k_2)\frac{k_0}{k_2}$ for brevity.

The initival value problem for P can be solved explicitly, and by taking the limit $t \to \infty$ we obtain

$$P^* = \lim_{t \to \infty} P(t) = \lim_{t \to \infty} \frac{k_0}{k_3} + c \cdot e^{-k_3 t} = \frac{k_0}{k_3},$$

which is the same as the equilibrium solution in Table 2.

The solution for S is more involved. First we substitute $E(t) = S(t) - (s_0 - e_0)$ (since $\dot{S}(t) - \dot{E}(t) = 0$) which gives us the following differential equation

$$\dot{S}(t) = -k_1 S(t)(S(t) - (s_0 - e_0)) + k = -k_1 S(t)^2 + (s_0 - e_0)k_1 S(t) + k.$$

To solve this differential equation we first divide both sides by $(-k_1(S(t))^2 + (s_0 - e_0)k_1 S(t) + k)$, and integrate both sides with respect to t,

$$\int \frac{dS(t)/dt}{-k_1 S(t)^2 + (s_0 - e_0)k_1 S(t) + k} dt = \int 1 dt \tag{34}$$

$$\int \frac{dS(t)}{-k_1 S(t)^2 + (s_0 - e_0)k_1 S(t) + k} = (t + c) \tag{35}$$

To evaluate the left-hand side of this equation we want to apply the following standard integral:

$$\int \frac{1}{ax^2 + bx + c} dx = \begin{cases} -\frac{2}{\sqrt{b^2 - 4ac}} \tanh^{-1}\left(\frac{2ax + b}{\sqrt{b^2 - 4ac}}\right) + C, & \text{if } |2ax + b| < \sqrt{b^2 - 4ac}, \\ -\frac{2}{\sqrt{b^2 - 4ac}} \coth^{-1}\left(\frac{2ax + b}{\sqrt{b^2 - 4ac}}\right) + C, & \text{else.} \end{cases} \tag{36}$$

for $b^2 - 4ac > 0$. We first check the condition:

$$b^2 - 4ac = (s_0 - e_0)^2 k_1^2 + 4k_1 k > 0.$$

15

We now take the first solution to the standard integral (the second solution gives the same limiting result for S, as we will see later on). We apply the first solution in (36) to (35) to obtain

$$\frac{2\tanh^{-1}\left(\frac{2k_1 S(t)-(s_0-e_0)k_1}{\sqrt{4k_1 k+(s_0-e_0)^2 k_1^2}}\right)}{\sqrt{4k_1 k+(s_0-e_0)^2 k_1^2}} = t+c \tag{37}$$

$$\tanh^{-1}\left(\frac{2k_1 S(t)-(s_0-e_0)k_1}{\sqrt{4k_1 k+(s_0-e_0)^2 k_1^2}}\right) = \frac{1}{2}(t+c)\sqrt{4k_1 k+(s_0-e_0)^2 k_1^2} \tag{38}$$

$$\frac{2k_1 S(t)-(s_0-e_0)k_1}{\sqrt{4k_1 k+(s_0-e_0)^2 k_1^2}} = \tanh\left(\frac{1}{2}(t+c)\sqrt{4k_1 k+(s_0-e_0)^2 k_1^2}\right), \tag{39}$$

Solving (39) for S gives,

$$S(t) = \frac{1}{2k_1}\left(\tanh\left(\frac{1}{2}(t+c)\sqrt{4k_1 k+(s_0-e_0)^2 k_1^2}\right)\sqrt{4k_1 k+(s_0-e_0)^2 k_1^2}+k_1(s_0-e_0)\right).$$

By taking the limit $t\to\infty$, plugging in $k=(k_{-1}+k_2)\frac{k_0}{k_2}$, and rewriting we obtain the equilibrium solution in Table 2:

$$\lim_{t\to\infty} S(t) = \frac{k_1(s_0-e_0)+\sqrt{4k_1 k+(s_0-e_0)^2 k_1^2}}{2k_1}$$

$$= \frac{k_1(s_0-e_0)+\sqrt{4k_1(k_{-1}+k_2)\frac{k_0}{k_2}+(s_0-e_0)^2 k_1^2}}{2k_1}$$

$$= \frac{1}{2}\left((s_0-e_0)+\sqrt{(s_0-e_0)^2+4\frac{k_0(k_{-1}+k_2)}{k_1 k_2}}\right).$$

Note that if we take the second solution to the standard integral in (36), then we would have ended up with the same solution for $S(t)$ with $\tanh$ replaced by $\coth$, but the limit $\lim_{t\to\infty} S(t)$ would still be the same.

The solution for E follows from the fact that $E(t) = S(t) - (s_0 - e_0)$. The solutions for all joint interventions were found by combining the arguments that were given for the single interventions.

A.3 Marginal model

In the paper we presented a marginal model for the basic enzyme reaction. Here we show how it can be derived from the causal constraints in the GSCM, which are given by

$$k_0 + k_{-1}C - k_1 SE = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus S}), \tag{40}$$

$$k_1 SE - (k_{-1}+k_2)C = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus C}), \tag{41}$$

$$-k_1 SE + (k_{-1}+k_2)C = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus E}), \tag{42}$$

$$k_2 C - k_3 P = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus P}), \tag{43}$$

$$C + E - (c_0 + e_0) = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus\{C,E\}}), \tag{44}$$

$$S - E - (s_0 - e_0) = 0, \qquad\qquad \{\mathbb{I}_C, (\mathbb{I}_C \cup \mathbb{I}_P)\}. \tag{45}$$

We obtain the marginal model as follows:

1. Reduce the number of variables that can be targeted by an intervention: $\boldsymbol{X}' = \{S, E, P\}$.

2. Rewrite the causal constraint in (41) to $C = \frac{k_1 SE}{k_{-1}+k_2}$. Note that this equation holds under any intervention in $\mathcal{P}(\mathbb{I}_{\boldsymbol{X}'}) = \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus C})$. Then substitute this expression for C into equation (40) to obtain

$$\frac{k_0 + k_{-1}\frac{k_0}{k_2}}{k_1 E} - S = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}'\setminus S}),$$

where the activation set of the causal constraint is given by the intersection $\mathcal{P}(\mathbb{I}_{\boldsymbol{X} \backslash S}) \cap \mathcal{P}(\mathbb{I}_{\boldsymbol{X}'})$. Then substitute this expresion for C into equation (43) to obtain

$$\frac{k_2}{k_3} \frac{k_1 S E}{k_{-1} + k_2} - P = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}' \backslash P}),$$

where the activation set of the causal constraint is given by the intersection $\mathcal{P}(\mathbb{I}_{\boldsymbol{X} \backslash P}) \cap \mathcal{P}(\mathbb{I}_{\boldsymbol{X}'})$.

3. Rewrite the causal constraint in (44) to $C = e_0 + c_0 - E$ and note that this equation holds under interventions in $\mathcal{P}(\mathbb{I}_{\boldsymbol{X}' \backslash E})$. Then substitute this expression for C into equation (42) to obtain

$$\frac{(k_{-1} + k_2)(c_0 + e_0)}{k_{-1} + k_2 + k_1 S} - E = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}' \backslash E}),$$

where the activation set of the causal constraint is given by the intersection $\mathcal{P}(\mathbb{I}_{\boldsymbol{X} \backslash \{C, E\}}) \cap \mathcal{P}(\mathbb{I}_{\boldsymbol{X}' \backslash E})$.

This procedure results in the following marginal model

$$\frac{k_0 + k_{-1}\frac{k_0}{k_2}}{k_1 E} - S = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}' \backslash S}),$$

$$\frac{(k_{-1} + k_2)(c_0 + e_0)}{k_{-1} + k_2 + k_1 S} - E = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}' \backslash E}),$$

$$\frac{k_2}{k_3} \frac{k_1 S E}{k_{-1} + k_2} - P = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}' \backslash P}).$$

Because we kept track of the interventions under which each equation is active when we substituted C into the equations of other causal constraints, we preserved the causal structure of the model. That is, the marginal GSCM model has the same solutions as the original GSCM under interventions in $\mathcal{P}(\mathbb{I}_{\boldsymbol{X}'})$.

B Additional examples

B.1 A chemical reaction network

The basic enzyme reaction is not a special case where SCMs cannot fully represent the equilibrium states. The example in this section is intended to illustrate that the problem occurs more generally in chemical reaction networks.

Consider a chemical reaction where molecules X_1 and X_2 react to form a molecule X_3 at a rate k_1, and with rate k_2 the molecules X_3 are converted back to X_2. This reaction is described by the following reaction equations,

$$X_1 + X_2 \xrightarrow{k_1} X_3$$
$$X_3 \xrightarrow{k_2} X_2.$$
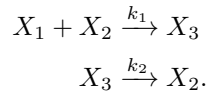
We assume that the concentration of $X_1(t)$ is kept constant[13].

The dynamics of the concentrations in this system as obtained by the law of mass-action are described by the following equations of motion and initial conditions,

$$\dot{X}_1 = 0$$
$$\dot{X}_2 = -k_1 X_1(t) X_2(t) + k_2 X_3(t)$$
$$\dot{X}_3 = k_1 X_1(t) X_2(t) - k_2 X_3(t),$$
$$X_1(0) = X_1^0, \qquad X_2(0) = X_2^0, \qquad X_3(0) = X_3^0.$$

We let $\boldsymbol{X}(t) = (X_1(t), X_2(t), X_3(t))$ be a set of endogenous variables taking value in $\mathbb{R}_+^3$,[14] and $\boldsymbol{E} = (X_1^0, X_2^0, X_3^0, k_1, k_2)$ a set of time-independent exogenous variables taking value in $\mathcal{E} = \mathbb{R}_+^5$ with a probability measure $\mathbb{P}_{\mathcal{E}}$.

---

[13]This can be accomplished by a process called chemostatting, where a permeable membrane allows one type of particle to move to and from a large buffer with a fixed concentration.

[14]$\mathbb{R}_+^3$ are the strictly positive real numbers.

Table 3: Solutions to the ODE of the chemical reaction network, under all interventions $I \in \mathcal{P}(\mathbb{I}_{\boldsymbol{X}})$

I	X_1	X_2	X_3
$\emptyset$	X_1^0	$(X_2^0 + X_3^0))\frac{k_2}{k_2+k_1 X_1^0}$	$(X_2^0 + X_3^0)\frac{k_1 X_1^0}{k_2+k_1 X_1^0}$
$\{1\}$	ξ_1	$(X_2^0 + X_3^0)\frac{k_2}{k_2+k_1 \xi_1}$	$(X_2^0 + X_3^0)\frac{k_1 \xi_1}{k_2+k_1 \xi_1}$
$\{2\}$	X_1^0	ξ_2	$\frac{k_1}{k_2}X_1^0\xi_2$
$\{3\}$	X_1^0	$\frac{k_2}{K_1}\frac{\xi_3}{X_1^0}$	ξ_3
$\{1,2\}$	ξ_1	ξ_2	$\frac{k_1}{k_2}xi_1\xi_2$
$\{1,3\}$	ξ_1	$\frac{k_2}{k_1}\frac{\xi_3}{\xi_1}$	ξ_3
$\{2,3\}$	X_1^0	ξ_2	ξ_3
$\{1,2,3\}$	ξ_1	ξ_2	ξ_3

Equilibrium solutions   The equilibrium solutions of this dynamical system under all possible intervention are displayed in Table 3. We can show that the system converges to these solutions by explicit calculation.

For the observed system note that $\dot{X}_2(t) + \dot{X}_3(t) = 0$ so that $X_2(t) = X_2^0 + X_3^0 - X_3(t)$, and $\dot{X}_1(t) = 0$ so that $X_1(t) = X_1^0$. Plugging this into $\dot{X}_2$ we find that

$$\dot{X}_2(t) = -k_1 X_1^0 X_2(t) + k_2(X_2^0 + X_3^0 - X_2(t)).$$

The solution to this differential equation is given by,

$$X_2(t) = \frac{k_2(X_2^0 + X_3^0)}{k_2 + k_1 X_1^0} + e^{-(k_1 X_1^0 + k_2)t}.$$

Taking the limit $t \to \infty$ indeed gives the equilibrium solution in Table 3. The equilibrium solution for X_3 follows from $X_3(t) = X_2^0 + X_3^0 - X_2(t)$.

After an intervention $\mathrm{do}(X_3 = \xi_3)$, we obtain the differential equation,

$$\dot{X}_2(t) = -k_1 X_1^0 X_2(t) + k_2\xi_3.$$

Solving this differential equation gives,

$$X_2(t) = \frac{k_2\xi_3}{k_1 X_1^0} + e^{-k_1 X_1^0 t}.$$

Taking the limit gives the equilibrium solution in Table 3. Convergence for the system under influence of other interventions follows similarly.

GSCM representation   To obtain the GSCM representation of the equilibrium states in this system, we set all equations of motions equal to zero. This leads to the following causal constraints,

$$-K_1 X_1 X_2 + K_2 X_3 = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus X_2}), \qquad\qquad (46)$$
$$K_1 X_1 X_2 - K_2 X_3 = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus X_3}). \qquad\qquad (47)$$

Since $\dot{X}_1 = 0$ and $\dot{X}_2 + \dot{X}_3 = 0$ for all t, the system also admits two constants of motion. The corresponding causal constraints are,

$$X_1 - X_1^0 = 0, \qquad\qquad \mathcal{P}(\mathbb{I}_{\boldsymbol{X}\setminus X_1}) \qquad\qquad (48)$$
$$X_2 + X_3 - (X_2^0 + X_3^0) = 0, \qquad\qquad \{\emptyset, \{1\}\} \qquad\qquad (49)$$

Let G denote the causal constraints in equations (46), (47), (48), and (49). These causal constraints imply a unique probability distribution on $\boldsymbol{X}$ under all interventions. The solutions to the GSCM correspond to the equilibrium solutions of the initial value problem, for each intervention. The GSCM $(\boldsymbol{X}, G, \mathbb{P}_{\boldsymbol{\varepsilon}})$ thus represents the causal semantics of the equilibrium states of the dynamical system.

Since an SCM can only have three equations, it cannot completely represent the equilibrium solutions of this system. This can be seen as follows. In order to capture the causal dependencies in the case that two out of three variables are intervened upon, an SCM must include the relations in (46) to (48). This leaves no room to include the dependence of X_2 and X_3 on the initial conditions, in the cases that both X_2 and X_3 are *not* targeted by an intervention.

B.2 Lotka-Volterra Predator-Prey Model

The Lotka-Volterra model describes the dynamics of population sizes of predator and prey species. It typically does not converge to a single equilibrium state but to steady-state oscillations. The dynamical system is given by

$$\frac{dX_1}{dt} = X_1(t)(\theta_{11} - \theta_{12}X_2(t))$$
$$\frac{dX_2}{dt} = -X_2(t)(\theta_{22} - \theta_{21}X_1)(t),$$

with initial values $X_1(0) = X_1^0 > 0, X_2(0) = X_2^0 > 0$ and rate parameters $\boldsymbol{\theta} > \mathbf{0}$. Let X_1 and X_2 be endogenous variables taking value in the positive reals and $\boldsymbol{E} = (X_1(0), X_2(0))$ background variables with probability distribution $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$, where $\boldsymbol{\mathcal{E}} = \mathbb{R}_+^2$.

First note that in the equilibrium state the equations of motion need to be equal to zero. This leads to the following causal constraints

$$X_1(\theta_{11} - \theta_{12}X_2) = 0 \qquad\qquad\qquad \{\{\emptyset\}, \{2\}\} \qquad\qquad (50)$$
$$X_2(\theta_{22} - \theta_{21}X_1) = 0 \qquad\qquad\qquad \{\{\emptyset\}, \{1\}\} \qquad\qquad (51)$$

The observed system admits a constant of motion through a nonlinear dependence between the equations of motion which is given by

$$\theta_{21}X_1 + \theta_{22}\log(X_1) - \theta_{12}X_2 + \theta_{11}\log(X_2) = -\theta_{21}X_1^0 + \theta_{22}\log(X_1^0) - \theta_{12}X_2^0 + \theta_{11}\log(X_2^0)$$

To represent this constant of motion in a GSCM, we consider the causal constraint

$$\theta_{21}X_1 + \theta_{22}\log(X_1) - \theta_{12}X_2 + \theta_{11}\log(X_2) - (-\theta_{21}X_1^0 + \theta_{22}\log(X_1^0) - \theta_{12}X_2^0 + \theta_{11}\log(X_2^0)) = 0, \quad \{\emptyset\}. \tag{52}$$

If an intervention on X_1 is such that $\theta_{22} - \theta_{21}X_1 = 0$, the system admits another constant of motion so that $X_1(t) = X_1(0)$ for all t. If the intervention is such that $\theta_{22} - \theta_{21}X_1 > 0$ then $X_2(t)$ tends to zero as $t \to \infty$. If the intervention on X_1 is such that $\theta_{22} - \theta_{21}X_1 < 0$, then the system only has a solution if $X_2(0) = 0$, otherwise $X_2(t)$ runs of to infinity. This behavior is captured by the combination of the causal constraints in (51) and (53). Analogously, we derive an additional causal constraint describing the effects of interventions on X_2 given in (54).

$$X_2 = X_2^0 \mathbf{1}_{\{\theta_{22} - \theta_{21}X_1 \le 0\}} \qquad\qquad\qquad \{\{1\}\} \qquad\qquad (53)$$
$$X_1 = X_1^0 \mathbf{1}_{\{\theta_{11} - \theta_{12}X_2 \ge 0\}} \qquad\qquad\qquad \{\{2\}\} \qquad\qquad (54)$$

The GSCM representation of the Lotka-Volterra model is obtained by taking G to be the causal constraints in equations (50) to (54), and $\mathcal{M} = (\boldsymbol{X}, G, \mathbb{P}_{\boldsymbol{E}})$.