From Deterministic ODEs to Dynamic Structural Causal Models

Paul K. Rubenstein[*]
Department of Engineering
University of Cambridge
pkr23@cam.ac.uk

Stephan Bongers
Informatics Institute
University of Amsterdam
S.R.Bongers@uva.nl

Bernhard Schölkopf
Max-Planck Institute for
Intelligent Systems, Tübingen
bs@tue.mpg.de

Joris M. Mooij
Informatics Institute
University of Amsterdam
J.M.Mooij@uva.nl

Abstract

Structural Causal Models are widely used in causal modelling, but how they relate to other modelling tools is poorly understood. In this paper we provide a novel perspective on the relationship between Ordinary Differential Equations and Structural Causal Models. We show how, under certain conditions, the asymptotic behaviour of an Ordinary Differential Equation under non-constant interventions can be modelled using Dynamic Structural Causal Models. In contrast to earlier work, we study not only the effect of interventions on equilibrium states; rather, we model asymptotic behaviour that is *dynamic* under interventions that vary in time, and include as a special case the study of static equilibria.

1 INTRODUCTION

Ordinary Differential Equations (ODEs) provide a universal language to describe deterministic systems via equations that determine how variables change in time as a function of other variables. They provide an immensely popular and highly successful modelling framework, with applications in many diverse disciplines, such as physics, chemistry, biology, and economy. They are *causal* in the sense that at least in principle they allow us to reason about interventions: any external intervention in a system—e.g., moving an object by applying a force—can be modelled using modified differential equations by, for instance, including suitable forcing terms. In practice, of course, this may be arbitrarily difficult.

Structural Causal Models (SCMs, also known as Structural Equation Models) are another language capable of

describing causal relations and interventions and have been widely applied in the social sciences, economics, genetics and neuroscience (Pearl, 2009; Bollen, 2014). One of the successes of SCMs over other causal frameworks such as causal Bayesian networks, for instance, has been their ability to express cyclic causal models (Spirtes, 1995; Mooij et al., 2011; Hyttinen et al., 2012; Voortman et al., 2010; Lacerda et al., 2008; Bongers et al., 2018).

We view SCMs as an intermediate level of description between the highly expressive differential equation models and the probabilistic, non-causal models typically used in machine learning and statistics. This intermediate level of description ideally retains the benefits of a data-driven statistical approach while still allowing a limited set of causal statements about the effect of interventions. While it is well understood how an SCM induces a statistical model (Bongers et al., 2018), much less is known about how a differential equation model—our most fundamental level of modelling—can imply an SCM in the first place. This is an important question because if we are to have models of a system on different levels of complexity, we should understand how they relate and the conditions under which they are consistent with one another.

Indeed, recent work has begun to address the question of how SCMs arise naturally from more fundamental models by showing how, under strong assumptions, SCMs can be derived from an underlying discrete time difference equation or continuous time ODE (Iwasaki and Simon, 1994; Dash, 2005; Lacerda et al., 2008; Voortman et al., 2010; Mooij et al., 2013; Sokol and Hansen, 2014). With the exception of (Voortman et al., 2010) and (Sokol and Hansen, 2014), each of these methods assume that the dynamical system comes to a static equilibrium that is independent of initial conditions, with the derived SCM describing how this equilibrium changes under intervention. More recently, the more general case in which the equilibrium state may depend on the initial conditions has been addressed (Bongers and Mooij, 2018; Blom and Mooij, 2018).

---

[*]Also affiliated with Max Planck Institute for Intelligent Systems, Tübingen.

If the assumption that the system reaches a static equilibrium is reasonable for a particular system under study, the SCM framework can be useful. Although the derived SCM then lacks information about the (possibly rich) transient dynamics of the system, if the system equilibrates quickly then the description of the system as an SCM may be a more convenient and compact representation of the causal structure of interest. By making assumptions on the dynamical system and the interventions being made, the SCM effectively allows us to reason about a 'higher level' qualitative description of the dynamics—in this case, the equilibrium states.

There are, however, two major limitations that stem from the equilibrium assumption. First, for many dynamical systems the assumption that the system settles to a unique equilibrium, either in its observational state or under intervention, may be a bad approximation of the actual system dynamics. Second, this framework is only capable of modelling interventions in which a subset of variables are clamped to fixed values (*constant* interventions). Even for rather simple physical systems such as a forced damped simple harmonic oscillator, these assumptions are violated.

Motivated by these observations, the work presented in this paper tries to answer the following questions: (i) Can the SCM framework be extended to model systems that do not converge to an equilibrium? (ii) If so, what assumptions need to be made on the ODE and interventions so that this is possible? Since SCMs are used in a variety of situations in which the equilibrium assumption does not necessarily hold, we view these questions as important in order to understand when they are indeed theoretically grounded as modelling tools. The main contribution of this paper is to show that the answer to the first question is 'Yes' and to provide sufficient conditions for the second. We do this by extending the SCM framework to encompass time-dependent dynamics and interventions and studying how such objects can arise from ODEs. We refer to this as a *Dynamic SCM (DSCM)* to distinguish it from the static equilibrium case for the purpose of exposition, but note that this is conceptually the same as an SCM on a fundamental level. Our construction draws inspiration from the approach of Mooij et al. (2013), that was recently generalized to also incorporate the stochastic setting (Bongers and Mooij, 2018). Here, we adapt the approach by replacing the static equilibrium states by continuous-time *trajectories*, considering two trajectories as equivalent if they do not differ asymptotically.

Note that whilst this paper applies a causal perspective to the study of dynamical systems, the goal of this paper is not to derive a learning algorithm which can be applied to time series data. In this sense, we view our main results as 'orthogonal' to methods such as Granger causality (Granger, 1969) and difference-in-differences (Card and Krueger, 1993) which aim to infer causal effects given time-series observations of a system. We envision that DSCMs may be used for causal analysis of dynamical systems that undergo periodic motion. Although these systems have been mostly ignored so far in the field of causal discovery, they have been studied extensively in the field of control theory. Some examples of systems that naturally exhibit oscillatory stationary states and where our framework may be applicable are EEG signals, circadian signals, seasonal influences, chemical oscillations, electric circuits, aerospace vehicles, and satellite control. We refer the reader to (Bittanti and Colaneri, 2009) for more details on these application areas from the perspective of periodic control theory.

Since the DSCM derived for a simple harmonic oscillator (see Example 4) is already quite complex, we leave the task of deriving methods that estimate the parameters from data for future work. Rather, our current work presents a first necessary theoretical step that needs to be done before applications of this theory can be developed, enabling the development of data-driven causal discovery and prediction methods for oscillatory systems, and possibly even more general systems, down the road.

The remainder of this paper is organised as follows. In Section 2, we introduce notation to describe ODEs. In Section 3, we describe how to apply the notion of an intervention on an ODE to the dynamic case. In Section 4, we define regularity conditions on the asymptotic behaviour of an ODE under a set of interventions. In Section 5, we present our main result: subject to conditions on the dynamical system and interventions being modelled, a *Dynamic SCM* can be derived that allows one to reason about how the asymptotic dynamics change under interventions on variables in the system. We conclude in Section 6.

2 ORDINARY DIFFERENTIAL EQUATIONS

Let $\mathcal{I} = \{1, \ldots, D\}$ be a set of variable labels. Consider time-indexed variables $X_i(t) \in \mathcal{R}_i$ for $i \in \mathcal{I}$, where $\mathcal{R}_i \subseteq \mathbb{R}$ and $t \in \mathbb{R}_{\geq 0} = [0, \infty)$. For $I \subseteq \mathcal{I}$, we write $\mathbf{X}_I(t) \in \prod_{i \in I} \mathcal{R}_i$ for the tuple of variables $(X_i(t))_{i \in I}$. By an ODE $\mathcal{D}$, we mean a collection of D coupled ordinary differential equations with initial conditions $\mathbf{X}_0^{(k)}$:

$$\mathcal{D} : \begin{cases} f_i(X_i, \mathbf{X}_{\mathtt{pa}(i)})(t) = 0, & X_i^{(k)}(0) = (\mathbf{X}_0^{(k)})_i, \\ & 0 \leq k \leq n_i - 1, \quad i \in \mathcal{I}, \end{cases}$$

where the ith differential equation determines the evolution of the variable X_i in terms of $\mathbf{X}_{\mathtt{pa}(i)}$, where $\mathtt{pa}(i) \subseteq \mathcal{I}$ are the *parents of* i, and X_i itself, and where

n_i is the order of the highest derivative $X_i^{(k)}$ of X_i that appears in equation i. Here, f_i is a functional that can include time-derivatives of its arguments. We think of the ith differential equation as modelling the *causal mechanism* that determines the dynamics of the effect X_i in terms of its direct causes $\mathbf{X}_{\text{pa}(i)}$.

One possible way to write down an ODE is to canonically decompose it into a collection of first order differential equations, such as is done in Mooij et al. (2013). We choose to present our ODEs as "one equation per variable" rather than splitting up the equations due to complications that would otherwise occur when considering time-dependent interventions (cf. Section 3.3).

Example 1. *Consider a one-dimensional system of D particles of mass m_i ($i = 1, \ldots, D$) with positions X_i coupled by springs with natural lengths l_i and spring constants k_i, where the ith spring connects the ith and $(i + 1)$th masses and the outermost springs have fixed ends (see Figure 1a). Assume further that the ith mass undergoes linear damping with coefficient b_i.*

Denoting by $\dot{X}_i$ and $\ddot{X}_i$ the first and second time derivatives of X_i respectively, the equation of motion for the ith variable is given by

$$
\begin{aligned}
m_i \ddot{X}_i(t) =& k_i[X_{i+1}(t) - X_i(t) - l_i] \\
& - k_{i-1}[X_i(t) - X_{i-1}(t) - l_{i-1}] - b_i \dot{X}_i(t)
\end{aligned}
$$

where we take $X_0 = 0$ and $X_D = L$ to be the fixed positions of the end springs. For the case that $D = 2$, we can write the system of equations as:

$$
\mathcal{D} : \begin{cases}
0 = m_1 \ddot{X}_1(t) + b_1 \dot{X}_1(t) + (k_1 + k_0)X_1(t) \\
\quad\quad - k_1 X_2(t) - k_0 l_0 + k_1 l_1 \,, \\
\\
0 = m_2 \ddot{X}_2(t) + b_2 \dot{X}_2(t) + (k_2 + k_1)X_2(t) \\
\quad\quad - k_2 L - k_1 X_1(t) - k_2 l_1 + k_2 l_2 \,, \\
\\
X_i^{(k)}(0) = (\mathbf{X}_0^{(k)})_i \quad k \in \{0, 1\}, \ i \in \{1, 2\} \,.
\end{cases}
$$

We can represent the functional dependence structure between variables implied by the functions f_i with a graph, in which variables are nodes and arrows point $X_j \longrightarrow X_i$ if $j \in \text{pa}(i)$. Self loops $X_i \longrightarrow X_i$ exist if $X_i^{(k)}$ appears in the expression of f_i for more than one value of k. This is illustrated for the system described in Example 1 in Figure 1b.

3 INTERVENTIONS ON ODES

We interpret ODEs as *causal* models. In particular, we consider the graph expressing the functional dependence structure to be the causal graph of the system, with an edge between X_i and X_j iff X_i is a direct cause of X_j (in the context of all variables $\mathbf{X}_{\mathcal{I}}$). In this section, we will formalize this causal interpretation by studying interventions on the system.

3.1 TIME-DEPENDENT PERFECT INTERVENTIONS

Usually in the causality literature, by a *perfect intervention* it is meant that a variable is clamped to take a specific given value. The natural analogue of this in the time-dependent case is a perfect intervention that forces a variable to take a particular *trajectory*. That is, given a subset $I \subseteq \mathcal{I}$ and a function $\boldsymbol{\zeta}_I : \mathbb{R}_{\geq 0} \longrightarrow \prod_{i \in I} \mathcal{R}_i$, we can intervene on the subset of variables $\mathbf{X}_I$ by forcing $\mathbf{X}_I(t) = \boldsymbol{\zeta}_I(t) \ \forall t \in \mathbb{R}_{\geq 0}$. Using Pearl's do-calculus notation (Pearl, 2009) and for brevity omitting the t, we write $\text{do}(\mathbf{X}_I = \boldsymbol{\zeta}_I)$ for this intervention. Such interventions are more general objects than those of the equilibrium or time-independent case, but in the specific case that we restrict ourselves to constant trajectories the two notions coincide.

3.2 SETS OF INTERVENTIONS

Recall that when modelling equilibrating dynamical systems under constant interventions, the set of interventions modelled coincides with the asymptotic behaviour of the system. We will generalise this relation to non-equilibrating behaviour.

The Dynamic SCMs that we will derive will describe the asymptotic dynamics of the ODE and how they change under different interventions. If we want to model 'all possible interventions', then the resulting asymptotic dynamics that can occur are arbitrarily complicated. The idea is to fix a simpler set of interventions and derive an SCM that models only these interventions, resulting in a model that is simpler than the original ODE but still allows us to reason about interventions we are interested in. In the examples in this paper, we restrict ourselves to periodic or quasi-periodic interventions, but the results hold for more general sets of interventions that satisfy the stability definitions presented later.

We need to define some notation to express the sets of interventions and the set of system responses to these interventions that we will model. Since interventions correspond to forcing variables to take some trajectory, we describe notation for defining sets of trajectories: For $I \subseteq \mathcal{I}$, let Dyn_I be a set of trajectories in $\prod_{i \in I} \mathcal{R}_i$. Let $\text{Dyn} = \cup_{I \in \mathcal{P}(\mathcal{I})} \text{Dyn}_I$ (where $\mathcal{P}(\mathcal{I})$ is the power set of $\mathcal{I}$ i.e., the set of all subsets of $\mathcal{I}$). Thus, an element $\boldsymbol{\zeta}_I \in \text{Dyn}_I$ is a function $\mathbb{R}_{\geq 0} \longrightarrow \prod_{i \in I} \mathcal{R}_i$, and Dyn consists of such functions for different $I \subseteq \mathcal{I}$. The main idea is that we want both the interventions and the system
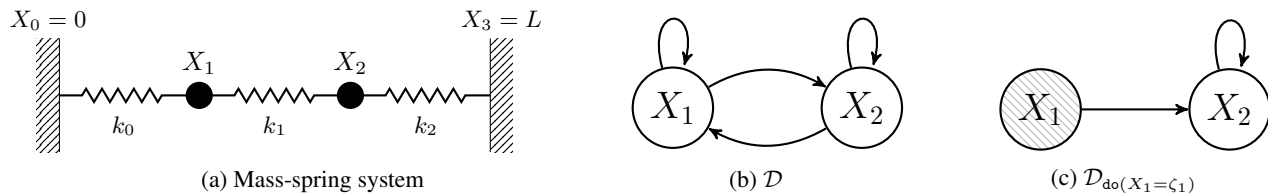
Figure 1: (a) The mass-spring system of Example 1 with $D = 2$; (b–c) graphs representing the causal structure of the mass-spring system for (b) the observational system, (c) after the intervention on variable X_1 described in Example 2. As a result of the intervention, X_1 is not causally influenced by any variable, while the causal mechanism of X_2 remains unchanged.

responses to be elements of Dyn; in other words, the set of possible system responses should be large enough to contain all interventions that we would like to model, and in addition, all responses of the system to those interventions. The reader might wonder why we do not simply take the set of *all* possible trajectories, but that set would be so large that it would not be practical for modeling purposes.[1]

Since our goal will be to derive a causal model that describes the relations between components (variables) of the system, we will need the following definition in Section 5.

Definition 1. *A set of trajectories* Dyn *is **modular** if, for any* $\{i_1, \ldots, i_n\} = I \subseteq \mathcal{I}$,

$$\zeta_I \in \text{Dyn} \iff \zeta_{i_k} \in \text{Dyn} \quad \forall k \in \{1, \ldots, n\}.$$

This should be interpreted as saying that admitted trajectories of single variables can be combined arbitrarily into admitted trajectories of the whole system (and *vice versa*, admitted system trajectories can be decomposed into trajectores of individual variables), and in addition, that interventions on each variable can be made independently and combined in any way.[2] This is not to say that all such interventions must be physically possible to implement in practice. Rather, this means that the mathematical model we derive should allow one to *reason* about all such interventions. Not all sets of trajectories Dyn are modular; in the following sections we will assume that the

sets of trajectories we are considering *are* for the purposes of constructing the Dynamic SCMs. Some examples of trivially modular sets of trajectories are: (i) all static (i.e., time-independent) trajectories, corresponding to (Mooij et al., 2013); (ii) all continuously-differentiable trajectories that differ asymptotically; (iii) all periodic motions. The latter is the running example in this paper.

3.3 DESCRIBING INTERVENTIONS ON ODEs

We can realise a perfect intervention by replacing the equations of the intervened variables with new equations that fix them to take the specified trajectories:[3]

$$\mathcal{D}_{\text{do}(\mathbf{X}_I = \zeta_I)}:$$
$$\begin{cases} f_i(X_i, \mathbf{X}_{\text{pa}(i)})(t) = 0, & X_i^{(k)}(0) = (\mathbf{X}_0^{(k)})_i, \\ 0 \le k \le n_i - 1, & i \in \mathcal{I} \setminus I, \\ X_i(t) - \zeta_i(t) = 0, & i \in I. \end{cases}$$

This procedure is analogous to the notion of intervention in an SCM. In reality, this corresponds to decoupling the intervened variables from their usual causal mechanism by forcing them to take a particular value, while leaving the non-intervened variables' causal mechanisms unaffected.

Perfect interventions will not generally be realisable in the real world. In practice, an intervention on a variable would correspond to altering the differential equation governing its evolution by adding extra forcing terms; perfect interventions could be realised by adding forcing terms that push the variable towards its target value at each instant in time, and considering the limit as these forcing terms become infinitely strong so as to dominate the usual causal mechanism determining the evolution of the variable.

Example 2 (continued). *Consider the mass-spring system described in Example 1. If we were to intervene on*

---

[1]For example, one might want to parameterize the set of trajectories in order to learn the model from data. Without any restriction on the smoothness of the trajectories, the problem of estimating a trajectory from data becomes ill-posed. Secondly, since we would like to identify trajectories that are asymptotically identical in order to focus the modeling efforts on the *asymptotic* behaviour of the system, we will only put a single trajectory into Dyn to represent all trajectories that are asymptotically identical to that trajectory, but whose transient dynamics may differ.

[2]This is related to notions that have been discussed in the literature under various headings, for instance autonomy and invariance (Pearl, 2009).

[3]Note that in the intervened ODE, the initial conditions of the intervened variables do not need to be specified explicitly as for the other variables, since they are implied by considering $t = 0$.

the system to force the mass X_1 to undergo simple harmonic motion, we could express this as a change to the system of differential equations as:

$$\mathcal{D}_{\mathrm{do}(X_1(t)=l_1+A\cos(\omega t))} :$$
$$\begin{cases} 0 = X_1(t) - l_1 - A\cos(\omega t)\,, \\[1em] 0 = m_2 \ddot{X}_2(t) + b_2 \dot{X}_2(t) + (k_2 + k_1)X_2(t) \\ \qquad -k_2 L - k_1 X_1(t) - k_2 l_1 + k_2 l_2\,, \\[1em] X_2^{(k)}(0) = (\mathbf{X}_0^{(k)})_2 \quad k \in \{0,1\}. \end{cases}$$

This induces a change to the graphical description of the causal relationships between the variables. We break any incoming arrows to any intervened variable, including self loops, as the intervened variables are no longer causally influenced by any other variable in the system. See Figure 1c for the graph corresponding to the intervened ODE in Example 2.

4 DYNAMIC STABILITY

A crucial assumption of Mooij et al. (2013) was that the systems considered were *stable* in the sense that they would converge to unique stable equilibria (if necessary, also after performing a constant intervention). This made them amenable to study by considering the $t \longrightarrow \infty$ limit in which any complex but transient dynamical behaviour would have decayed. The SCMs derived would allow one to reason about the asymptotic equilibrium states of the systems after interventions. Since we want to consider non-constant asymptotic dynamics, this is not a notion of stability that is fit for our purposes.

Instead, we define our stability with reference to a set of trajectories. We will use $\mathrm{Dyn}_{\mathcal{I}}$ for this purpose. Recall that elements of $\mathrm{Dyn}_{\mathcal{I}}$ are trajectories for all variables in the system. To be totally explicit, we can think of an element $\boldsymbol{\eta} \in \mathrm{Dyn}_{\mathcal{I}}$ as a function

$$\boldsymbol{\eta} : \quad \mathbb{R}_{\geq 0} \longrightarrow \mathcal{R}_{\mathcal{I}}$$
$$t \mapsto (\eta_1(t), \eta_2(t), \ldots, \eta_D(t))$$

where $\eta_i(t) \in \mathcal{R}_i$ is the state of the ith variable X_i at time t. Note that $\mathrm{Dyn}_{\mathcal{I}}$ is not a single fixed set, independent of the situation we are considering. We can choose $\mathrm{Dyn}_{\mathcal{I}}$ depending on the ODE $\mathcal{D}$ under consideration, and the interventions that we may wish to make on it.

Informally, stability in this paper means that the asymptotic dynamics of the dynamical system converge to a unique element of $\mathrm{Dyn}_{\mathcal{I}}$, independent of initial condition. If $\mathrm{Dyn}_{\mathcal{I}}$ is in some sense simple, we can simply characterise the asymptotic dynamics of the system under study. The following definitions of stability extend those

of Mooij et al. (2013) to allow for non-constant trajectories in $\mathrm{Dyn}_{\mathcal{I}}$, and coincide with them in the case that $\mathrm{Dyn}_{\mathcal{I}}$ consists of all constant trajectories in $\mathcal{R}_{\mathcal{I}}$.

Definition 2. *The ODE $\mathcal{D}$ is **dynamically stable with reference to** $\mathrm{Dyn}_{\mathcal{I}}$ if there exists a unique $\boldsymbol{\eta}_{\emptyset} \in \mathrm{Dyn}_{\mathcal{I}}$ such that $\mathbf{X}_{\mathcal{I}}(t) = \boldsymbol{\eta}_{\emptyset}(t) \,\forall t$ is a solution to $\mathcal{D}$ and that for any initial condition, the solution $\mathbf{X}_{\mathcal{I}}(t) \to \boldsymbol{\eta}_{\emptyset}(t)$ as $t \to \infty$.*[4]

We use a subscript $\emptyset$ to emphasise that $\boldsymbol{\eta}_{\emptyset}$ describes the asymptotic dynamics of $\mathcal{D}$ without any intervention. Observe that $\mathrm{Dyn}_{\mathcal{I}}$ could consist of the single element $\boldsymbol{\eta}_{\emptyset}$ in this case. The requirement that this hold for all initial conditions can be relaxed to hold for all initial conditions except on a set of measure zero, but that would mean that the proofs later on require some more technical details. For the purpose of exposition, we stick to this simpler case.

Example 3. *Consider a single mass on a spring that is undergoing simple periodic forcing and is underdamped. Such a system could be expressed as a single (parent-less) variable with ODE description:*

$$\mathcal{D} : \begin{cases} m\ddot{X}_1(t) + b\dot{X}_1(t) + k(X_1(t) - l) \\ \qquad = F\cos(\omega t + \phi)\,, \\[1em] X_1^{(k)}(0) = (X_0^{(k)}) \quad k \in \{0,1\}\,. \end{cases}$$

The solution to this differential equation is

$$X_1(t) = r(t) + l + A\cos(\omega t + \phi') \tag{1}$$

where $r(t)$ decays exponentially quickly (and is dependent on the initial conditions) and A and ϕ' depend on the parameters of the equation of motion (but not on the initial conditions).

Therefore such a system would be dynamically stable with reference to (for example)

$$\mathrm{Dyn}_{\mathcal{I}} = \{l + A\cos(\omega t + \phi') : A \in \mathbb{R},\ \phi' \in [0,2\pi)\}.$$

Remark 1. We use a subscript $\boldsymbol{\zeta}_I$ to emphasise that $\boldsymbol{\eta}_{\boldsymbol{\zeta}_I}$ describes the asymptotic dynamics of $\mathcal{D}$ after performing the intervention $\mathrm{do}(\mathbf{X}_I = \boldsymbol{\zeta}_I)$. Observe that $\mathrm{Dyn}_{\mathcal{I}}$ could consist only of the single element $\boldsymbol{\eta}_{\boldsymbol{\zeta}_I}$ and the above definition would be satisfied. But then the original ODE wouldn't be dynamically stable with reference to $\mathrm{Dyn}_{\mathcal{I}}$, nor would other intervened versions of $\mathcal{D}$. This motivates the following definition, extending dynamic stability to sets of intervened systems.

---

[4]The convergence we refer to here is the usual asymptotic convergence of real-valued functions, i.e., for $f : [0,\infty) \to \mathbb{R}^d$, $g : [0,\infty) \to \mathbb{R}^d$ we have that $f \to g$ iff for every $\epsilon > 0$ there is a $T \in [0,\infty)$ such that $|f(t) - g(t)| < \epsilon$ for all $t \in [T,\infty)$.

Definition 3. *Let* Traj *be a set of trajectories. We say that the pair* $(\mathcal{D}, \mathtt{Traj})$ *is **dynamically stable with reference to** $\mathtt{Dyn}_{\mathcal{I}}$ if, for any $\zeta_I \in \mathtt{Traj}$, $\mathcal{D}_{\mathrm{do}(\mathbf{X}_I = \zeta_I)}$ is dynamically stable with reference to $\mathtt{Dyn}_{\mathcal{I}}$.*

Example 3 (continued). *Suppose we are interested in modelling the effect of changing the forcing term, either in amplitude, phase or frequency. We introduce a second variable X_2 to model the forcing term:*

$$
\mathcal{D} : \begin{cases}
\begin{aligned}
0 &= f_1(X_1, X_2)(t) \\
&= m\ddot{X}_1(t) + b\dot{X}_1(t) + k(X_1(t) - l) - X_2(t), \\
0 &= f_2(X_2)(t) \\
&= X_2(t) - F_0 \cos(\omega_0 t + \phi_0), \\
X_1^{(k)}(0) &= (\mathbf{X}_0^{(k)})_1, \quad k \in \{0, 1\}.
\end{aligned}
\end{cases}
$$

If we want to change the forcing term that we apply to the mass, we can interpret this as performing an intervention on X_2. We could represent this using the notation we have developed as

$$
\mathtt{Dyn}_{\{2\}} = \{\zeta_2(t) = F_2 \cos(\omega t + \phi_2) :
$$
$$
F_2, \omega \in \mathbb{R}, \ \phi_2 \in [0, 2\pi)\}.
$$

For any intervention $\zeta_2 \in \mathtt{Dyn}_{\{2\}}$, the dynamics of X_1 in $\mathcal{D}_{\mathrm{do}(X_2 = \zeta_2)}$ will be of the form (1). Therefore $(\mathcal{D}, \mathtt{Dyn}_{\{2\}})$ will be dynamically stable with reference to

$$
\mathtt{Dyn}_{\mathcal{I}} = \Big\{ \boldsymbol{\zeta}(t) = (l + F_1 \cos(\omega t + \phi_1), F_2 \cos(\omega t + \phi_2))
$$
$$
: F_1, F_2, \omega \in \mathbb{R}, \ \phi_1, \phi_2 \in [0, 2\pi) \Big\}.
$$

The independence of initial conditions for Example 3 is illustrated in Figure 2.

Note that if $(\mathcal{D}, \mathtt{Traj})$ is dynamically stable with reference to $\mathtt{Dyn}_{\mathcal{I}}$, and $\mathtt{Dyn}'_{\mathcal{I}} \supseteq \mathtt{Dyn}_{\mathcal{I}}$ is a larger set of trajectories that still satisfies the uniqueness condition in the definition of dynamic stability,[5] then $(\mathcal{D}, \mathtt{Traj})$ is dynamically stable with reference to $\mathtt{Dyn}'_{\mathcal{I}}$.

5 DYNAMIC STRUCTURAL CAUSAL MODELS

A deterministic SCM $\mathcal{M}$ is a collection of structural equations, the ith of which defines the value of variable X_i

---

[5] Namely: $\forall \boldsymbol{\zeta}_I \in \mathtt{Traj}$, $\exists! \, \boldsymbol{\eta}_{\boldsymbol{\zeta}_I} \in \mathtt{Dyn}'_{\mathcal{I}}$ such that under $\mathcal{D}_{\mathrm{do}(\mathbf{X}_I = \boldsymbol{\zeta}_I)}$ and for any initial condition, $X_{\mathcal{I}}(t) \to \boldsymbol{\eta}_{\boldsymbol{\zeta}_I}(t)$ as $t \to \infty$. Assuming that $(\mathcal{D}, \mathtt{Traj})$ is dynamically stable with reference to $\mathtt{Dyn}_{\mathcal{I}}$, a sufficient condition for this is that none of the elements in $\mathtt{Dyn}'_{\mathcal{I}} \setminus \mathtt{Dyn}_{\mathcal{I}}$ are asymptotically equal to any of the elements of $\mathtt{Dyn}_{\mathcal{I}}$. That is: $\forall \boldsymbol{\zeta} \in \mathtt{Dyn}_{\mathcal{I}}$, $\forall \boldsymbol{\zeta}' \in \mathtt{Dyn}'_{\mathcal{I}} \setminus \mathtt{Dyn}_{\mathcal{I}}$, $\boldsymbol{\zeta}(t) \not\to \boldsymbol{\zeta}'(t)$ as $t \to \infty$.

in terms of its parents. We extend this to the case that our variables do not take fixed values but rather represent entire trajectories.

Definition 4. *Let* $\mathtt{Dyn} = \bigcup_{I \subseteq \mathcal{I}} \mathtt{Dyn}_I$ *be a modular set of trajectories, where* $\mathtt{Dyn}_I \subseteq \mathcal{R}_I^{\mathbb{R}^{\geq 0}}$. *A deterministic Dynamic Structural Causal Model (DSCM) on the time-indexed variables $\mathbf{X}_{\mathcal{I}}$ taking values in* $\mathtt{Dyn}$ *is a collection of* structural equations

$$
\mathcal{M} : \big\{ \ X_i = F_i(\mathbf{X}_{\mathrm{pa}(i)}) \quad i \in \mathcal{I},
$$

where $\mathrm{pa}(i) \subseteq \mathcal{I} \setminus \{i\}$ *and each F_i is a map* $\mathtt{Dyn}_{\mathrm{pa}(i)} \longrightarrow \mathtt{Dyn}_i$ *that gives the trajectory of an effect variable in terms of the trajectories of its direct causes.*

The point of this paper is to show that, subject to restrictions on $\mathcal{D}$ and $\mathtt{Dyn}$, we can derive a DSCM that allows us to reason about the effect on the asymptotic dynamics of interventions using trajectories in $\mathtt{Dyn}$. 'Traditional' deterministic SCMs arise as a special case, where all trajectories are constant over time.

In an ODE, the equations f_i determine the causal relationship between the variable $X_i(t)$ and its parents $\mathbf{X}_{\mathrm{pa}(i)}(t)$ *at each instant* in time. In contrast, we think of the function F_i of the DSCM as a causal mechanism that determines the entire trajectory of X_i in terms of the trajectories of the variables $\mathbf{X}_{\mathrm{pa}(i)}$, integrating over the instantaneous causal effects over all time. In the case that $\mathtt{Dyn}$ consists of constant trajectories (and thus the instantaneous causal effects are constant over time), a DSCM reduces to a traditional deterministic SCM.

The rest of this section is laid out as follows. In Section 5.1 we define what it means to make an intervention in a DSCM. In Section 5.2 we show how, subject to certain conditions, a DSCM can be derived from a pair $(\mathcal{D}, \mathtt{Dyn})$. The procedure for doing this relies on intervening on all but one variable at a time. In Section 5.3, Theorem 2 states that the DSCM thus derived is capable of modelling the effect of intervening on arbitrary subsets of variables, even though it was constructed by considering the case that we consider interventions on exactly $D - 1$ variables. Theorem 3 and Corollary 1 in Section 5.4 prove that the notions of intervention in ODE and the derived DSCM coincide. Collectively, these theorems tell us that we can derive a DSCM that allows us to reason about the effects of interventions on the asymptotic dynamics of the ODE. Proofs of these theorems are provided in Section A of the Supplementary Material.

5.1 INTERVENTIONS IN A DSCM

Interventions in (D)SCMs are realized by replacing the structural equations of the intervened variables. Given
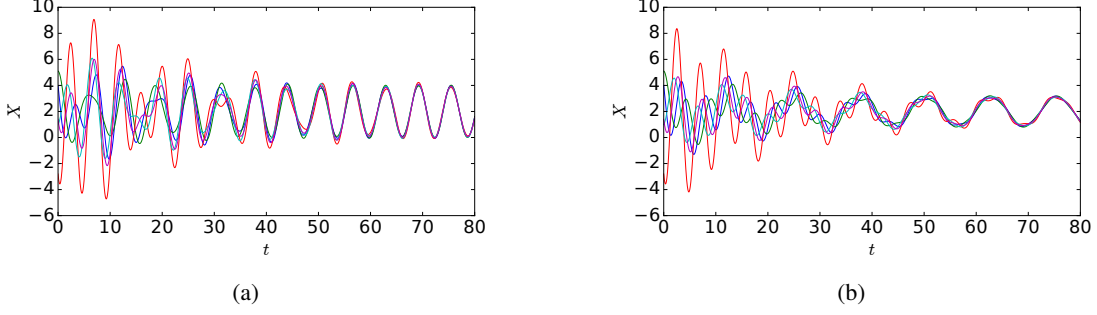
Figure 2: Simulations from the forced simple harmonic oscillator in Example 3 showing the evolution of X_1 with different initial conditions for different forcing terms (interventions on X_2). The parameters used were $m = 1, k = 1, l = 2, F = 2, b = 0.1$, with (a) $\omega = 3$ and (b) $\omega = 2$. Dynamic stability means that asymptotic dynamics are independent of initial conditions, and the purpose of the DSCM is to quantify how the asymptotic dynamics change under intervention.

$\zeta_I \in \mathtt{Dyn}_I$ for some $I \subseteq \mathcal{I}$, the intervened DSCM $\mathcal{M}_{\mathrm{do}(\mathbf{X}_I = \zeta_I)}$ can be written:

$$\mathcal{M}_{\mathrm{do}(\mathbf{X}_I = \zeta_I)} : \begin{cases} X_i & = F_i(\mathbf{X}_{\mathrm{pa}(i)}) & i \in \mathcal{I} \setminus I, \\ X_i & = \zeta_i & i \in I. \end{cases}$$

The causal mechanisms determining the non-intervened variables are unaffected, so their structural equations remain the same. The intervened variables are decoupled from their usual causal mechanisms and are forced to take the specified trajectory.

5.2 DERIVING DSCMs FROM ODEs

In order to derive a DSCM from an ODE, we require the following consistency property between the asymptotic dynamics of the ODE and the set of interventions.

Definition 5 (Structural dynamic stability). *Let* $\mathtt{Dyn}$ *be modular. The pair* $(\mathcal{D}, \mathtt{Dyn})$ *is **structurally dynamically stable** if* $(\mathcal{D}, \mathtt{Dyn}_{\mathcal{I} \setminus \{i\}})$ *is dynamically stable with reference to* $\mathtt{Dyn}_{\mathcal{I}}$ *for all* i.

This means that for any intervention trajectory $\zeta_{\mathcal{I} \setminus \{i\}} \in \mathtt{Dyn}_{\mathcal{I} \setminus \{i\}}$, the asymptotic dynamics of the intervened ODE $\mathcal{D}_{\mathrm{do}(\mathbf{X}_{\mathcal{I} \setminus \{i\}} = \zeta_{\mathcal{I} \setminus \{i\}})}$ are expressible uniquely as an element of $\mathtt{Dyn}_{\mathcal{I}}$. Since $\mathtt{Dyn}$ is modular, the asymptotic dynamics of the non-intervened variable can be realised as the trajectory $\zeta_i \in \mathtt{Dyn}_i$, and thus $\mathtt{Dyn}$ is rich enough to allow us to make an intervention which forces the non-intervened variable to take this trajectory. This is a crucial property that allows the construction of the structural equations. In the particular case that $\mathtt{Dyn}$ consists of all constant trajectories, structural dynamic stability means that after any intervention on all-but-one-variable, the non-intervened variable settles to a unique equilibrium. In the language of Mooij et al. (2013), this would imply that the ODE is *structurally stable*.

It should be noted that $(\mathcal{D}, \mathtt{Dyn})$ being structurally dynamically stable is a strong assumption in general. If $\mathtt{Dyn}$ is too small,[6] then it may be possible to find a larger set $\mathtt{Dyn}' \supset \mathtt{Dyn}$ such that $(\mathcal{D}, \mathtt{Dyn}')$ *is* structurally dynamically stable. The procedure described in this section describes how to derive a DSCM capable of modelling all interventions in $\mathtt{Dyn}'$, which can thus be used to model interventions in $\mathtt{Dyn}$.

Henceforth, we use the notation $I_i = \mathcal{I} \setminus \{i\}$ for brevity. Suppose that $(\mathcal{D}, \mathtt{Dyn})$ is structurally dynamically stable. We can **derive structural equations** $F_i : \mathtt{Dyn}_{\mathrm{pa}(i)} \longrightarrow \mathtt{Dyn}_i$ to describe the asymptotic dynamics of children variables as functions of their parents as follows. Pick $i \in \mathcal{I}$. The variable X_i has parents $\mathbf{X}_{\mathrm{pa}(i)}$. Since $\mathtt{Dyn}$ is modular, for any configuration of parent dynamics $\boldsymbol{\eta}_{\mathrm{pa}(i)} \in \mathtt{Dyn}_{\mathrm{pa}(i)}$ there exists $\zeta_{I_i} \in \mathtt{Dyn}_{I_i}$ such that $(\zeta_{I_i})_{\mathrm{pa}(i)} = \boldsymbol{\eta}_{\mathrm{pa}(i)}$.

By structural dynamic stability, the system $\mathcal{D}_{\mathrm{do}(\mathbf{X}_{I_i} = \zeta_{I_i})}$ has asymptotic dynamics specified by a unique element $\boldsymbol{\eta} \in \mathtt{Dyn}_{\mathcal{I}}$, which in turn defines a unique element $\eta_i \in \mathtt{Dyn}_i$ specifying the asymptotic dynamics of variable X_i since $\mathtt{Dyn}$ is modular.

Theorem 1. *Suppose that* $(\mathcal{D}, \mathtt{Dyn})$ *is structurally dynamically stable. Then the functions*

$$F_i : \mathtt{Dyn}_{\mathrm{pa}(i)} \to \mathtt{Dyn}_i : \boldsymbol{\eta}_{\mathrm{pa}(i)} \mapsto \eta_i$$

constructed as above are well-defined.

Given the structurally dynamically stable pair $(\mathcal{D}, \mathtt{Dyn})$ we define the derived DSCM

$$\mathcal{M}_{\mathcal{D}} : \{ \quad X_i = F_i(\mathbf{X}_{\mathrm{pa}(i)}) \quad i \in \mathcal{I},$$

---

[6]For example, if $\mathtt{Dyn}$ is not modular or represents interventions on only a subset of the variables.

where the $F_i : \mathtt{Dyn}_{\mathtt{pa}(i)} \to \mathtt{Dyn}_i$ are defined as above. Note that structural dynamic stability was a crucial property that ensured $F_i(\mathtt{Dyn}_{\mathtt{pa}(i)}) \subseteq \mathtt{Dyn}_i$. If $(\mathcal{D}, \mathtt{Dyn})$ is not structurally dynamically stable, we cannot build structural equations in this way.

We provide next an example of a DSCM for the mass-spring system of Example 1 with $D = 2$. The derivation of this for the general case of arbitrarily many masses is included in the Supplementary Material.

Example 4. *Consider the system $\mathcal{D}$ governed by the differential equation of Example 1 with $D = 2$. Let $\mathtt{Dyn}_{\{1,2\}}$ be the modular set of trajectories with*

$$\mathtt{Dyn}_{\{i\}} = \left\{ \sum_{j=1}^{\infty} A_i^j \cos(\omega_i^j t + \phi_i^j) \ : \right.$$
$$\left. w_i^j, \phi_i^j, A_i^j \in \mathbb{R}, \sum_{j=1}^{\infty} |A_i^j| < \infty \right\}$$

for $i = 1, 2$, where for each i it holds that $\sum_{j=1}^{\infty} |A_i^j| < \infty$ (so that the series is absolutely convergent). Then $(\mathcal{D}, \mathtt{Dyn}_{\{1,2\}})$ is structurally dynamically stable and admits the following DSCM.

$$\mathcal{M} : \begin{cases} X_1 &= F_1(X_2) \\ X_2 &= F_2(X_1) \end{cases}$$

where, writing $C_1^j = [k_1 + k_2 - m_1(\omega_2^j)^2]^2$ and $C_2^j = [k_1 + k_2 - m_2(\omega_1^j)^2]^2$, the functionals F_1 and F_2 are given by Equations 2 and 3 overleaf.

5.3 SOLUTIONS OF A DSCM

Theorem 1 states that we can construct a DSCM by the described procedure. We constructed each equation by intervening on $D - 1$ variables at a time. The result of this section states that the DSCM can be used to correctly model interventions on *arbitrary* subsets of variables. We say that $\boldsymbol{\eta}_{\mathcal{I}} \in \mathtt{Dyn}_{\mathcal{I}}$ is a *solution* of $\mathcal{M}$ if $\eta_i = F_i(\boldsymbol{\eta}_{\mathtt{pa}(i)}) \ \forall i \in \mathcal{I}$.

Theorem 2. *Suppose that $(\mathcal{D}, \mathtt{Dyn})$ is structurally dynamically stable. Let $I \subseteq \mathcal{I}$, and let $\boldsymbol{\zeta}_I \in \mathtt{Dyn}_I$. Then $\mathcal{D}_{do(\mathbf{X}_I = \boldsymbol{\zeta}_I)}$ is dynamically stable if and only if the intervened SCM $\mathcal{M}_{(\mathcal{D}_{do(\mathbf{X}_I = \boldsymbol{\zeta}_I)})}$ has a unique solution. If there is a unique solution, it coincides with the element of $\mathtt{Dyn}_{\mathcal{I}}$ describing the asymptotic dynamics of $\mathcal{D}_{do(\mathbf{X}_I = \boldsymbol{\zeta}_I)}$.*

Remark 2. We could also take $I = \emptyset$, in which case the above theorem applies to just $\mathcal{D}$.

5.4 CAUSAL REASONING IS PRESERVED

We have defined ways to model interventions in both ODEs and DSCMs. The following theorem and its immediate corollary proves that these notions of intervention coincide, and hence that DSCMs provide a representation to reason about the asymptotic behaviour of the ODE under interventions in $\mathtt{Dyn}$. A consequence of these results is that the diagram in Figure 3 commutes.

Theorem 3. *Suppose that $(\mathcal{D}, \mathtt{Dyn})$ is structurally dynamically stable. Let $I \subseteq \mathcal{I}$ and let $\boldsymbol{\zeta}_I \in \mathtt{Dyn}_I$. Then $\mathcal{M}_{(\mathcal{D}_{do(\mathbf{X}_I = \boldsymbol{\zeta}_I)})} = (\mathcal{M}_{\mathcal{D}})_{do(\mathbf{X}_I = \boldsymbol{\zeta}_I)}$.*

Corollary 1. *Suppose additionally that $J \subseteq \mathcal{I} \setminus I$ and let $\boldsymbol{\zeta}_J \in \mathtt{Dyn}_J$. Then*

$$\left(\mathcal{M}_{(\mathcal{D}_{do(\mathbf{X}_I = \boldsymbol{\zeta}_I)})} \right)_{do(\mathbf{X}_J = \boldsymbol{\zeta}_J)} = (\mathcal{M}_{\mathcal{D}})_{do(\mathbf{X}_I = \boldsymbol{\zeta}_I, \mathbf{X}_J = \boldsymbol{\zeta}_J)} .$$

To summarise, Theorems 1–3 and Corollary 1 collectively state that if $(\mathcal{D}, \mathtt{Dyn})$ is dynamically structurally stable then it is possible to derive a DSCM that allows us to reason about the asymptotic dynamics of the ODE under any possible intervention in $\mathtt{Dyn}$.

5.5 RELATION TO ODEs AND DYNAMIC BAYESIAN NETWORKS

An ODE is capable of modelling arbitrary interventions on the system it describes. At the cost of only modelling a restricted set of interventions, a DSCM can be derived which describes the asymptotic behaviour of the system under these interventions. This may be desirable in cases for which transient behaviour is not important.

We now compare DSCMs to Dynamic Bayesian Networks (DBNs), an existing popular method for causal modelling of dynamical systems (Koller and Friedman, 2009). DBNs are essentially Markov chains, and thus are appropriate for discrete-time systems. When the discrete-time Markov assumption holds, DBNs are a powerful tool capable of modelling arbitrary interventions. However, approximations must be made whenever these assumptions do not hold. In particular, a continuous system must be approximately discretised in order to be modelled by a DBN (Sokol and Hansen, 2014).

By using the Euler method for numerically solving ODEs, we can make such an approximation to derive a DBN describing the system in Example 1, leading to the discrete time equation given in (8) the Supplementary Material. For DBNs, the main choice to be made is how fine the temporal discretisation should be. The smaller the value of Δ, the better the discrete approximation will be. Even if there is a natural time-scale on which measurements can be made, choosing a finer discretisation than this will provide a better approximation to the behaviour of the true system. The choice of Δ should reflect the natural timescales of the interventions to be considered too; for example, it is not clear how one would model the intervention $do\left(X_1(t) = \cos\left(\frac{2\pi t}{\Delta}\right)\right)$ with a discretisation length Δ. Another notable disadvantage of DBNs is that the
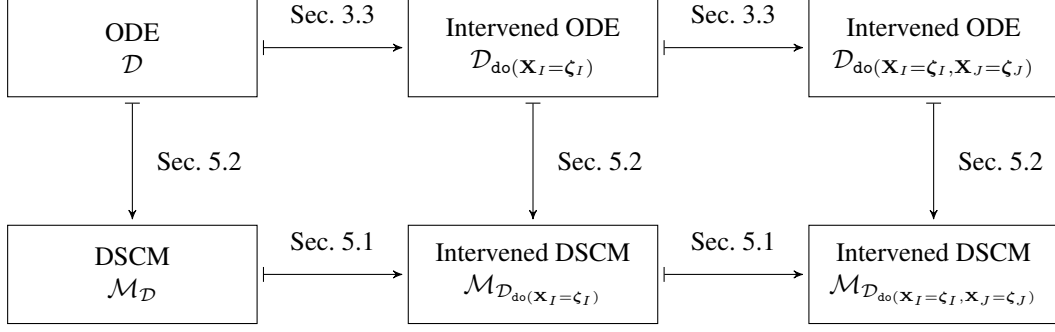
Figure 3: Top-to-bottom arrows: Theorems 1 and 2 together state that if $(\mathcal{D}, \mathtt{Dyn})$ is structurally dynamically stable then we can construct a DSCM to describe the asymptotic behaviour of $\mathcal{D}$ under different interventions in the set $\mathtt{Dyn}$. Left-to-right arrows: Both ODEs and DSCMs are equipped with notions of intervention. Theorem 3 and Corollary 1 say that these two notions of intervention coincide, and thus the diagram commutes.

$$F_1\left(\sum_{j=1}^{\infty} A_2^j \cos(\omega_2^j t + \phi_2^j)\right) = \frac{-k_1 l_1}{k_1 + k_0} + \sum_{j=1}^{\infty} \frac{k_1 A_2^j}{\sqrt{C_1^j + b_1 m_1 (\omega_2^j)^2}} \cos\left(\omega_2^j t + \phi_2^j - \arctan\left[\frac{b_1 \omega_2^j}{C_1^j}\right]\right) \tag{2}$$

$$F_2\left(\sum_{j=1}^{\infty} A_1^j \cos(\omega_1^j t + \phi_1^j)\right) = \frac{k_1 l_1 - k_2 l_2}{k_1 + k_2} + \frac{k_2 L}{k_2 + k_3} + \sum_{j=1}^{\infty} \frac{k_1 A_1^j}{\sqrt{C_2^j + b_2 m_2 (\omega_1^j)^2}} \cos\left(\omega_1^j t + \phi_1^j - \arctan\left[\frac{b_2 \omega_1^j}{C_2^j}\right]\right) \tag{3}$$

Figure 4: Equations giving the structural equations for the DSCM describing the mass-spring system of Example 4

computational cost of learning and inference increases for smaller Δ, where computational cost becomes infinitely large in the limit $\Delta \to 0$.

In contrast, the starting point for DSCMs is to fix a convenient set of interventions we are interested in modelling. If a DSCM containing these interventions exists, it will model the asymptotic behaviour of the system under each of these interventions *exactly*, rather than approximately modelling the transient and asymptotic behaviour as in the case of a DBN. Computational cost does not relate inversely to accuracy as for DBNs, but depends on the chosen representation of the set of admitted interventions.

6 DISCUSSION AND FUTURE WORK

The main contribution of this paper is to show that the SCM framework can be applied to reason about time-dependent interventions on an ODE in a dynamic setting. In particular, we showed that if an ODE is sufficiently well-behaved under a set of interventions, a DSCM can be derived that captures how the asymptotic dynamics change under these interventions. This is in contrast to previous approaches to connecting the language of ODEs with the SCM framework, which used SCMs to describe the stable (constant-in-time) equilibria of the ODE and

how they change under intervention.

We identify three possible directions in which to extend this work in the future. The first is to properly understand how learning DSCMs from data could be performed. This is important if DSCMs are to be used in practical applications. Challenges to be addressed include finding practical parameterizations of DSCMs, the presence of measurement noise in the data and the fact that time-series data are usually sampled at a finite number of points in time. The second is to relax the assumption that the asymptotic dynamics are *independent of initial conditions*, as was done recently for the static equilibrium scenario by Blom and Mooij (2018). The third extension is to move away from deterministic systems and consider Random Differential Equations (Bongers and Mooij, 2018), thereby allowing to take into account model uncertainty, but also to include systems that may be inherently stochastic.

References

J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, NY, 2nd edition, 2009.

Kenneth A Bollen. *Structural equations with latent variables*. John Wiley & Sons, 2014.

Peter Spirtes. Directed cyclic graphical representations of feedback models. In *Proceedings of the Eleventh conference on Uncertainty in Artificial Intelligence (UAI 1995)*, pages 491–498, 1995.

Joris M Mooij, Dominik Janzing, Tom Heskes, and Bernhard Schölkopf. On causal discovery with cyclic additive noise models. In *Advances in Neural Information Processing Systems (NIPS 2011)*, pages 639–647, 2011.

Antti Hyttinen, Frederick Eberhardt, and Patrik O Hoyer. Learning linear cyclic causal models with latent variables. *The Journal of Machine Learning Research*, 13 (1):3387–3439, 2012.

Mark Voortman, Denver Dash, and Marek J Druzdzel. Learning why things change: the difference-based causality learner. In *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence (UAI 2010)*, 2010.

Gustavo Lacerda, Peter L Spirtes, Joseph Ramsey, and Patrik O Hoyer. Discovering cyclic causal models by independent components analysis. In *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence (UAI 2008)*, 2008.

Stephan Bongers, Jonas Peters, Bernhard Schölkopf, and Joris M. Mooij. Theoretical aspects of cyclic structural causal models. *arXiv.org preprint*, arXiv:1611.06221v2 [stat.ME], 2018.

Yumi Iwasaki and Herbert A Simon. Causality and model abstraction. *Artificial Intelligence*, 67(1):143–194, 1994.

Denver Dash. Restructuring dynamic causal systems in equilibrium. In *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics (AISTATS 2005)*, 2005.

J. Mooij, D. Janzing, and B. Schölkopf. From ordinary differential equations to structural causal models: the deterministic case. In *Proceedings of the Twenty-Ninth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI 2013)*, pages 440–448, 2013.

Alexander Sokol and Niels Richard Hansen. Causal interpretation of stochastic differential equations. *Electronic Journal of Probability*, 19(100):1–24, 2014.

Stephan Bongers and Joris M. Mooij. From random differential equations to structural causal models: the stochastic case. *arXiv.org preprint*, arXiv:1803.08784 [cs.AI], March 2018. URL https://arxiv.org/abs/1803.08784.

Tineke Blom and Joris M. Mooij. Generalized structural causal models. *arXiv.org preprint*, https://arxiv.org/abs/1805.06539 [cs.AI], May 2018. URL https://arxiv.org/abs/1805.06539.

Clive WJ Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438, 1969.

David Card and Alan B Krueger. Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania. Technical report, National Bureau of Economic Research, 1993.

Sergio Bittanti and Patrizio Colaneri. *Periodic systems: filtering and control*, volume 5108985. Springer Science & Business Media, 2009.

Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*. The MIT Press, 2009.