arXiv:1611.06221v2 [stat.ME] 5 Aug 2018

Theoretical Aspects of Cyclic Structural Causal Models

Stephan Bongers, Jonas Peters, Bernhard Schölkopf, Joris M. Mooij

Abstract

Structural causal models (SCMs), also known as (non-parametric) structural equation models (SEMs), are widely used for causal modeling purposes. A large body of theoretical results is available for the special case in which cycles are absent (i.e., *acyclic* SCMs, also known as *recursive* SEMs). However, in many application domains cycles are abundantly present, for example in the form of feedback loops. In this paper, we provide a general and rigorous theory of cyclic SCMs. The paper consists of two parts: the first part gives a rigorous treatment of structural causal models, dealing with measure-theoretic and other complications that arise in the presence of cycles. In contrast with the acyclic case, in cyclic SCMs solutions may no longer exist, or if they exist, they may no longer be unique, or even measurable in general. We give several sufficient and necessary conditions for the existence of (unique) measurable solutions. We show how causal reasoning proceeds in these models and how this differs from the acyclic case. Moreover, we give an overview of the Markov properties that hold for cyclic SCMs. In the second part, we address the question of how one can marginalize an SCM (possibly with cycles) to a subset of the endogenous variables. We show that under a certain condition, one can effectively remove a subset of the endogenous variables from the model, leading to a more parsimonious marginal SCM that preserves the causal and counterfactual semantics of the original SCM on the remaining variables. Moreover, we show how the marginalization relates to the latent projection and to latent confounders, i.e. latent common causes.

Contents

1 Introduction

Structural causal models (SCMs), also known as (non-parametric) structural equation models (SEMs), are widely used for causal modeling purposes (Pearl, 2009; Spirtes et al., 2000). In these models, the causal relationships are expressed in the form of deterministic, functional relationships, and probabilities are introduced through the assumption that certain variables are exogenous latent random variables. SCMs arose out of certain causal models that were first introduced in genetics (Wright, 1921), econometrics (Haavelmo, 1943), electrical engineering (Mason, 1953, 1956), and the social sciences (Goldberger and Duncan, 1973; Duncan, 1975).

Acyclic SCMs, also known as recursive SEMs, form a special well-studied class of SCMs that are closely related to causal Bayesian networks (Pearl, 2009). These directed graphical models have relatively simple definitions in terms of various equivalent Markov properties, are computationally scalable and have several appealing statistical properties (Pearl, 1988; Lauritzen et al., 1990; Lauritzen, 1996). However, an important limitation of acyclic SCMs is that they cannot represent systems that involve cycles. For that purpose, cyclic SCMs (or non-recursive SEMs) form an appropriate model class (see e.g., Mooij et al., 2013; Bongers and Mooij, 2018). In contrast to the acyclic case, however, cyclic SCMs have enjoyed less attention and are not as well understood as their acyclic counterparts, although some progress has been made in the case of discrete (Pearl and Dechter, 1996; Neal, 2000) and linear models (Spirtes, 1994, 1995; Hyttinen et al., 2012). More recently, their Markov properties have been elucidated (Forré and Mooij, 2017) in more generality.

When allowing for cyclic functional relationships between variables, one encounters various technical complications. The structural equations of an acyclic SCM trivially have a unique solution. This *unique solvability* property ensures that the SCM gives rise to a unique, well-defined probability distribution on the variables. In the case of cycles, however, this property may be violated, and consequently, the SCM may have no solution at all, or may allow for multiple different probability distributions, which leads to ambiguity. Worse, even if one starts with a cyclic SCM that is uniquely solvable, performing an intervention on the SCM may lead to an intervened SCM that is not uniquely solvable. Furthermore, solutions need not be measurable in complete generality. Moreover, the causal interpretation of SCMs with multiple different probability distributions can be counter-intuitive, as the functional relations between variables no longer need to coincide with their causal relations. All these issues make cyclic SCMs a notoriously more complicated class of models to work with than the class of acyclic SCMs. In the first part of this paper we will give a general, rigorous treatment including both acyclic and cyclic SCMs, dealing with the technical complications that arise in the presence of cycles.

In the second part of this paper, we attempt to define the concept of *marginalization* for SCMs. Consider an SCM (possibly with cycles) that is defined on a large system, consisting of observable endogenous variables and latent exogenous variables. For example, such an SCM could be obtained by starting with an ordinary differential equation model and considering its equilibrium states (Mooij et al., 2013; Bongers and Mooij, 2018). Often, one is not interested in modeling the entire system, but would like to focus attention on a smaller subsystem instead, for example due to practical limitations on the observability of some of the endogenous variables in the model. The central question in this part of the paper is whether and how the SCM that represents the large system can be "projected down" to a more parsimonious SCM that represents the smaller subsystem.[1] This so-called *marginal SCM* representing the subsystem is defined on a subset of the original endogenous variables and should preserve the causal semantics of the variables of the subsystem. Marginalizing an SCM over a set of endogenous variables allows us to focus our causal modeling efforts on a part of the system and ignore the rest of the system. On a high level, this concept is analogous to marginalizing a probability distribution over a set of variables: this also reduces the description of a set of variables to a description of a subset of those, but the difference is that while the original meaning of marginalization is purely probabilistic, for SCMs it can in addition be interpreted causally.

We will define a marginalization operation that effectively removes a subset of the endogenous variables from the model, while preserving the probabilistic, causal and even counterfactual semantics. Intuitively, the idea is to think of the subset of endogenous variables of interest as a subsystem that can interact with the rest of the system. Under some suitable conditions, one can ignore the internals of this subsystem and treat it effectively as a "black box", which has a unique output for every possible input. We will show that this marginalization operation indeed preserves the causal and counterfactual semantics, i.e., all interventional and counterfactual distributions of the remaining endogenous variables induced by the original SCM are identical to those induced by its marginalization.

This paper is structured as follows. In Section 2 we will provide a formal definition of SCMs, define the (augmented) functional graph, describe interventions and counterfactuals, discuss the concept of solvability, consider various equivalence relations between SCMs, discuss their causal interpretation, and finally discuss the Markov properties. In Section 3, we will give a definition of a marginalization operation that is under certain conditions applicable to cyclic SCMs. We discuss several properties of this marginalization operation, define a marginalization operation on directed graphs and discuss their relation. We will give a causal interpretation of marginal SCMs and in the end discuss latent confounders. We wrap up and discuss the implications of this work in the last section.

[1] This question relates to the concept of "latent projection" (Verma, 1993; Tian, 2002) that has been formulated for DAG models and for ancestral graph Markov models (Richardson and Spirtes, 2002).

2 Structural causal models

2.1 Structural causal models and their solutions

In this section, we will start by formally defining structural causal models (SCMs) and their solutions. One usually considers solutions, in terms of random variables, to be part of an SCM. We will use a slightly different approach here that makes it easier to deal with cyclic SCMs. Our approach is to strip off the random variables from the definition of the SCM, which ensures that SCMs form a well-defined model class that is closed under any perfect intervention (see also Section 2.4). In Section 2.6, we will give conditions under which these SCMs induce a well-defined set of (measurable) random variables.

Definition 2.1.1 *A* structural causal model (SCM) *is a tuple*[2]

$$\mathcal{M} := \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$$

where

- $\mathcal{I}$ *is a finite index set of* endogenous variables*,*
- $\mathcal{J}$ *is a finite index set of* exogenous variables*,*
- $\boldsymbol{\mathcal{X}} = \prod_{i\in\mathcal{I}} \mathcal{X}_i$ *is the product of the codomains of the endogenous variables where each codomain* $\mathcal{X}_i$ *is a standard*[3] *measurable space,*
- $\boldsymbol{\mathcal{E}} = \prod_{j\in\mathcal{J}} \mathcal{E}_j$ *is the product of the codomains of the exogenous variables where each codomain* $\mathcal{E}_j$ *is a standard measurable space,*
- $\boldsymbol{f} : \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{X}}$ *is a measurable function that specifies the* causal mechanisms*,*
- $\mathbb{P}_{\boldsymbol{\mathcal{E}}} = \prod_{j\in\mathcal{J}} \mathbb{P}_{\mathcal{E}_j}$ *where* $\mathbb{P}_{\mathcal{E}_j}$ *is a probability measure on* $\mathcal{E}_j$ *for all* $j \in \mathcal{J}$*.*

Although it is common to assume the absence of cyclic functional relations (see Definition 2.2.4), we will make no such assumption here. When allowing for such cycles, it turns out that the most natural setting is obtained when allowing for self-loops as well. Under certain conditions, the existence of self-loops are obstructed. We will discuss this in more detail in Section 2.6.5.

In structural causal models, the functional relationships are expressed in terms of (deterministic) functional equations.

[2]We often use boldface for variables that have multiple components, e.g., vectors or tuples in a Cartesian product.

[3]A standard measurable space is a measurable space (Ω, Σ) that is isomorphic to a measurable space $(\tilde{\Omega}, \tilde{\Sigma})$, where $\tilde{\Omega}$ is a Polish space (i.e., a complete separable metric space) and $\tilde{\Sigma} = \mathcal{B}(\tilde{\Omega})$ are the Borel subsets of $\tilde{\Omega}$ (i.e., the σ-algebra generated by the open sets in $\tilde{\Omega}$). In several proofs we will assume without loss of generality that the standard measurable space *is* a Polish space Ω with σ-algebra $\mathcal{B}(\Omega)$. Examples of standard measurable spaces are open and closed subsets of $\mathbb{R}^d$, and finite sets with the complete σ-algebra. See for example Cohn (2013) for more details.

Definition 2.1.2 *For an SCM $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$, we call the set of equations:*

$$\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{e}) \qquad \boldsymbol{x} \in \boldsymbol{\mathcal{X}}, \boldsymbol{e} \in \boldsymbol{\mathcal{E}}$$

the structural equations *of the structural causal model $\mathcal{M}$.*

Somebody familiar with structural causal models will note that we are still missing an important ingredient: random variables. In our approach, they come in as follows:

Definition 2.1.3 *A pair of random variables $(\boldsymbol{X}, \boldsymbol{E})$ is a* solution *of the SCM $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ if*

1. *$\boldsymbol{E}$ takes values in $\boldsymbol{\mathcal{E}}$,*
2. *$\boldsymbol{X}$ takes values in $\boldsymbol{\mathcal{X}}$,*
3. *$\mathbb{P}^{\boldsymbol{E}} = \mathbb{P}_{\boldsymbol{\mathcal{E}}}$, i.e., the distribution of $\boldsymbol{E}$ is equal to $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$,*
4. *the* structural equations

$$\boldsymbol{X} = \boldsymbol{f}(\boldsymbol{X}, \boldsymbol{E})$$

are satisfied almost surely.

For convenience, we will call a random variable $\boldsymbol{X}$ a solution of $\mathcal{M}$, if there exists a random variable $\boldsymbol{E}$ such that $(\boldsymbol{X}, \boldsymbol{E})$ forms a solution of $\mathcal{M}$.

Often, the endogenous random variables $\boldsymbol{X}$ can be observed, and the exogenous random variables $\boldsymbol{E}$ are latent. Latent exogenous variables are often referred to as "disturbance terms" or "noise variables". Whenever a solution $\boldsymbol{X}$ exists, we call the distribution $\mathbb{P}^{\boldsymbol{X}}$ the *observational distribution of $\mathcal{M}$ associated to $\boldsymbol{X}$*. Note that in general there may be several different observational distributions associated to an SCM due to the existence of different solutions. This is, as we will see later, a consequence of the allowance of the existence of cycles in SCMs.

The reason for stripping off the random variables from the definition of the structural causal model is that, as we will see in Section 2.4, it may happen that an SCM with a solution may have no solution anymore after performing an intervention on the SCM. Conversely, it may also happen that intervening on an SCM without any solution gives an SCM with a solution. By separating the SCMs from their solutions, interventions can always be defined.

In terms of the solutions of an SCM, this definition leads to an ambiguity, which can be illustrated by the following example:

Example 2.1.4 *Consider the SCM $\mathcal{M} = \langle \mathbf{1}, \mathbf{1}, \mathcal{X}, \mathcal{E}, f, \mathbb{P}_{\mathcal{E}} \rangle$, where we take $\mathcal{X} = \mathcal{E} = \{-1, 0, 1\}$ with $\mathbb{P}_{\mathcal{E}}(\{-1\}) = \mathbb{P}_{\mathcal{E}}(\{1\}) = \frac{1}{2}$ and causal mechanism $f(x, e) = e^2 + e - 1$. For brevity, we use here and throughout this paper the notation $\boldsymbol{n} := \{1, 2, \ldots, n\}$ for $n \in \mathbb{N}$. Let $\tilde{\mathcal{M}}$ be the SCM $\mathcal{M}$ but with a different causal mechanism $\tilde{f}(e) = e$. Then the structural equations for both SCMs have a different solution set, which only differs on the point where $e = 0$ which has probability equal to zero. Hence any random variable X is a solution for $\mathcal{M}$ if and only if it is for $\tilde{\mathcal{M}}$.*

This ambiguity follows from the fact that the structural equations are only satisfied almost surely. To accommodate for this it seems natural to not differentiate between causal mechanisms that have different solutions on at most a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set of exogenous variables and thus leads to an equivalence relation between the causal mechanisms.

Before we state the equivalence relation we will introduce the following notation: For a subset $\tilde{\mathcal{I}} \subseteq \mathcal{I}$ and $\tilde{\mathcal{J}} \subseteq \mathcal{J}$ we write $\boldsymbol{\mathcal{X}}_{\tilde{\mathcal{I}}} := \prod_{i \in \tilde{\mathcal{I}}} \mathcal{X}_i$ and $\boldsymbol{\mathcal{E}}_{\tilde{\mathcal{J}}} := \prod_{j \in \tilde{\mathcal{J}}} \mathcal{E}_j$. Moreover, for a subset $\mathcal{U} \subseteq \mathcal{I} \cup \mathcal{J}$, we use the convention that we write $\boldsymbol{\mathcal{X}}_{\mathcal{U}}$ and $\boldsymbol{\mathcal{E}}_{\mathcal{U}}$ instead of $\boldsymbol{\mathcal{X}}_{\mathcal{U} \cap \mathcal{I}}$ and $\boldsymbol{\mathcal{E}}_{\mathcal{U} \cap \mathcal{J}}$ respectively and we adopt a similar notation for the (random) variables in those spaces, that is, we write $\boldsymbol{x}_{\mathcal{U}}$ and $\boldsymbol{e}_{\mathcal{U}}$ instead of $\boldsymbol{x}_{\mathcal{U} \cap \mathcal{I}}$ and $\boldsymbol{e}_{\mathcal{U} \cap \mathcal{J}}$ respectively.

Definition 2.1.5 *Consider two mappings* $\boldsymbol{f} : \boldsymbol{\mathcal{X}}_{\mathcal{I}} \times \boldsymbol{\mathcal{E}}_{\mathcal{J}} \to \boldsymbol{\mathcal{X}}_{\mathcal{K}}$ *and* $\tilde{\boldsymbol{f}} : \boldsymbol{\mathcal{X}}_{\tilde{\mathcal{I}}} \times \boldsymbol{\mathcal{E}}_{\tilde{\mathcal{J}}} \to \boldsymbol{\mathcal{X}}_{\tilde{\mathcal{K}}}$ *together with a probability distribution* $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathcal{J} \cup \tilde{\mathcal{J}}}}$ *on* $\boldsymbol{\mathcal{E}}_{\mathcal{J} \cup \tilde{\mathcal{J}}}$. *The mappings* $\boldsymbol{f}$ *and* $\tilde{\boldsymbol{f}}$ *are* (componentwise) equivalent, *denoted by* $\boldsymbol{f} \equiv \tilde{\boldsymbol{f}}$, *if* $\mathcal{K} = \tilde{\mathcal{K}}$ *and for all* $k \in \mathcal{K}$, *for* $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathcal{J} \cup \tilde{\mathcal{J}}}}$*-almost every* $\boldsymbol{e}_{\mathcal{J} \cup \tilde{\mathcal{J}}}$ *for all* $\boldsymbol{x}_{\{k\} \cup \mathcal{I} \cup \tilde{\mathcal{I}}} \in \boldsymbol{\mathcal{X}}_{\{k\} \cup \mathcal{I} \cup \tilde{\mathcal{I}}}$:[4]

$$x_k = f_k(\boldsymbol{x}_{\mathcal{I}}, \boldsymbol{e}_{\mathcal{J}}) \iff x_k = \tilde{f}_k(\boldsymbol{x}_{\tilde{\mathcal{I}}}, \boldsymbol{e}_{\tilde{\mathcal{J}}})\,.$$

This gives rise to an equivalence relation for SCMs:

Definition 2.1.6 *Two SCMs* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *and* $\tilde{\mathcal{M}} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \tilde{\boldsymbol{f}}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *are* equivalent *(denoted by* $\mathcal{M} \equiv \tilde{\mathcal{M}}$*) if* $\boldsymbol{f} \equiv \tilde{\boldsymbol{f}}$.

This equivalence relation $\equiv$ on the set of all SCMs gives rise to the quotient set of equivalence classes of SCMs. Note that two equivalent SCMs can only differ by their causal mechanism. Importantly, equivalent SCMs have the same solutions.

In this paper we will prove several properties and define several operations and relations on the quotient set of equivalence classes of SCMs. A common approach for proving a certain property for an equivalence class of SCMs is that we start by proving that the property holds for a representative of the equivalence class, and then show that it holds for any other element of that equivalence class. Similarly, in order to define a certain operation (or relation) on the equivalence class of SCMs, we usually start by defining the operation on an SCM and then show that this operation preserves the equivalence relation. In both cases we say that the property or operation *descends* to the set of equivalences classes.

2.2 The (augmented) functional graph

The functional relations of an SCM are described by the causal mechanism of the SCM. The causal mechanism encodes an (augmented) functional graph that can be summarized by a directed (mixed) graph. In acyclic SCMs the functional graph describes the direct causal graph,

[4] Please note that in general the quantifier "for $\mathbb{P}$-almost every" does not commute with the quantifier "for all".

but in cyclic SCMs this may not be the case, as we will see in Section 2.8. Before we define the (augmented) functional graph in terms of directed (mixed) graphs, we first state some standard terminology on directed (mixed) graphs.

2.2.1 Graphical notation and terminology

A *directed graph* is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a set of nodes and $\mathcal{E}$ a set of directed edges, which is a subset $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ of ordered pairs of nodes. Each element $(i, j) \in \mathcal{E}$ can be represented by the directed edge $i \to j$. In particular, $(i, i) \in \mathcal{E}$ represents a *self-loop* $i \to i$.

A *directed mixed graph* is a triple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{B})$, where the pair $(\mathcal{V}, \mathcal{E})$ forms a directed graph and $\mathcal{B}$ is a set of bidirected edges, which is a subset $\mathcal{B} \subseteq (\mathcal{V} \times \mathcal{V}) \setminus \Delta$, where $\Delta := \{(i, i) : i \in \mathcal{V}\}$, of unordered (distinct) pairs of nodes. Each element $(i, j) \in \mathcal{B}$ can be represented by the bidirected edge $i \leftrightarrow j$. Note that a directed graph is a directed mixed graph where $\mathcal{B}$ is the empty set.

Consider a directed mixed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{B})$. Then, a directed mixed graph $\tilde{\mathcal{G}} = (\tilde{\mathcal{V}}, \tilde{\mathcal{E}}, \tilde{\mathcal{B}})$ is a *subgraph of* $\mathcal{G}$, if $\tilde{\mathcal{V}} \subseteq \mathcal{V}$, $\tilde{\mathcal{E}} \subseteq \mathcal{E}$ and $\tilde{\mathcal{B}} \subseteq \mathcal{B}$. For a subset $\mathcal{W} \subseteq \mathcal{V}$, we define the *induced subgraph of* $\mathcal{G}$ *on* $\mathcal{W}$ by $\mathcal{G}_{\mathcal{W}} := (\mathcal{W}, \tilde{\mathcal{E}}, \tilde{\mathcal{B}})$, where $\tilde{\mathcal{E}}$ and $\tilde{\mathcal{B}}$ are the set of directed and bidirected edges in $\mathcal{E}$ and $\mathcal{B}$ respectively that lie in $\mathcal{W} \times \mathcal{W}$.

A *walk* between i and j in a directed (mixed) graph $\mathcal{G}$ is a sequence of edges (directed or bidirected) $(\epsilon_1, \ldots, \epsilon_n)$ for which there exists a sequence of nodes $(i = i_0, i_1, \ldots, i_{n-1}, i_n = j)$ for some $n \geq 0$, such that $\epsilon_k \in \{i_{k-1} \to i_k, i_{k-1} \leftarrow i_k, i_{k-1} \leftrightarrow i_k\}$ for $k = 1, 2, \ldots, n$ (note that $n = 0$ corresponds with a walk consisting of a single node); if all nodes $i_0, \ldots, i_n$ are distinct, it is called a *path*. A path (walk) of the form $i \to \cdots \to j$, i.e., such that ϵ_k is $i_{k-1} \to i_k$ for all $k = 1, 2, \ldots, n$, is called a *directed path (walk)* from i to j. A *cycle* is a sequence of edges $(\epsilon_1, \ldots, \epsilon_{n+1})$ such that $(\epsilon_1, \ldots, \epsilon_n)$ is a directed path from i to j and ϵ_{n+1} is $j \to i$. In particular, a self-loop is a cycle. Note that a path cannot contain any cycles. A directed graph and a directed mixed graph are *acyclic*, if they contain no cycles, and are then referred to as a *directed acyclic graph (DAG)* and an *acyclic directed mixed graph (ADMG)* respectively.

For a directed (mixed) graph $\mathcal{G}$ and a node $i \in \mathcal{V}$ we define the set of *parents* of i by $\mathrm{pa}_{\mathcal{G}}(i) := \{j \in \mathcal{V} : j \to i \in \mathcal{E}\}$, the set of children of i by $\mathrm{ch}_{\mathcal{G}}(i) := \{j \in \mathcal{V} : i \to j \in \mathcal{E}\}$, the set of *ancestors* of i by

$$\mathrm{an}_{\mathcal{G}}(i) := \{j \in \mathcal{V} : \text{there is a directed path from } j \text{ to } i \text{ in } \mathcal{G}\}$$

and the set of *descendants* of i by

$$\mathrm{de}_{\mathcal{G}}(i) := \{j \in \mathcal{V} : \text{there is a directed path from } i \text{ to } j \text{ in } \mathcal{G}\} .$$

Note that we have $\{i\} \cup \mathrm{pa}_{\mathcal{G}}(i) \subseteq \mathrm{an}_{\mathcal{G}}(i)$ and $\{i\} \cup \mathrm{ch}_{\mathcal{G}}(i) \subseteq \mathrm{de}_{\mathcal{G}}(i)$. We can apply these definitions to subsets $\mathcal{U} \subseteq \mathcal{V}$ by taking the union of these sets, for example $\mathrm{pa}_{\mathcal{G}}(\mathcal{U}) := \cup_{i \in \mathcal{U}} \mathrm{pa}_{\mathcal{G}}(i)$.

We call a directed mixed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{B})$ *strongly connected* if every two distinct nodes $i, j \in \mathcal{V}$ are connected via a cycle, i.e. via two directed paths $i \to \cdots \to j$ and $j \to \cdots \to i$. The *strongly connected component of* $i \in \mathcal{V}$, denoted by $\mathrm{sc}_{\mathcal{G}}(i)$, is the biggest subset $\mathcal{S} \subseteq \mathcal{V}$ such that $i \in \mathcal{S}$ and the induced subgraph $\mathcal{G}_{\mathcal{S}}$ is strongly connected. In other words, $\mathrm{sc}_{\mathcal{G}}(i)$ is the biggest subset $\mathcal{S} \subseteq \mathcal{V}$ such that $i \in \mathcal{S}$ and every two distinct nodes $j, k \in \mathcal{S}$ are connected via a cycle, i.e.. via two directed paths $j \to \cdots \to k$ and $k \to \cdots \to j$.

In the rest of this paper we will omit the subscript $\mathcal{G}$ whenever it is clear which directed (mixed) graph $\mathcal{G}$ we are referring to.

2.2.2 The (augmented) functional graph

The functional relations of an SCM are given by:

Definition 2.2.1 *We say that for an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$*:*

(i) $i \in \mathcal{I}$ *is a* functional parent *of* $k \in \mathcal{I}$ *if and only if there does not exist a measurable function*[5] $\tilde{f}_k : \boldsymbol{\mathcal{X}}_{\setminus i} \times \boldsymbol{\mathcal{E}} \to \mathcal{X}_k$ *such that* $\tilde{f}_k \equiv f_k$ *(see Definition 2.1.5).*

(ii) $j \in \mathcal{J}$ *is a* functional parent *of* $k \in \mathcal{I}$ *if and only if there does not exist a measurable function* $\tilde{f}_k : \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}}_{\setminus j} \to \mathcal{X}_k$ *such that* $\tilde{f}_k \equiv f_k$.

By definition these functional relations are preserved under the equivalence relation $\equiv$ on SCMs and can be represented in a directed graph or a directed mixed graph:

Definition 2.2.2 *Given an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$. *We define:*

(i) the augmented functional graph $\mathcal{G}^a(\mathcal{M})$ *as the directed graph with nodes* $\mathcal{I} \cup \mathcal{J}$ *and directed edges* $i \to j$ *if and only if* i *is a functional parent of* j.

(ii) the functional graph $\mathcal{G}(\mathcal{M})$ *as the directed mixed graph with nodes* $\mathcal{I}$*, directed edges* $i \to j$ *if and only if* i *is a functional parent of* j *and bidirected edges* $i \leftrightarrow j$ *if and only if there exists a* $k \in \mathcal{J}$ *such that* k *is a functional parent of both* i *and* j.

In particular, the augmented functional graph contains no directed edges between exogenous variables, because they are not functionally related by the causal mechanism. These definitions map $\mathcal{M}$ to $\mathcal{G}^a(\mathcal{M})$ and $\mathcal{G}(\mathcal{M})$, we call these mappings the *augmented functional graph mapping* $\mathcal{G}^a$ and the *functional graph mapping* $\mathcal{G}$ respectively. By definition, the mappings $\mathcal{G}^a$ and $\mathcal{G}$ are invariant under the equivalence relation $\equiv$ on SCMs and hence the whole equivalence class of an SCM is mapped to a unique augmented functional graph and functional graph.

[5] For $\boldsymbol{\mathcal{X}} = \prod_{k \in \mathcal{I}} \mathcal{X}_k$ for some index set $\mathcal{I}$, we denote $\boldsymbol{\mathcal{X}}_{\setminus i} = \prod_{k \in \mathcal{I} \setminus \{i\}} \mathcal{X}_k$, and similarly for its elements.

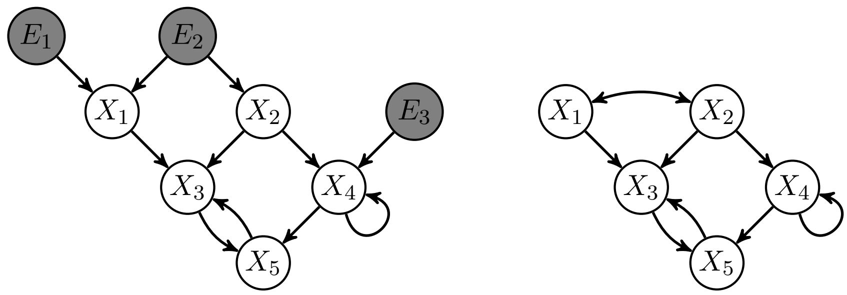

Figure 1: The augmented functional graph (left) and the functional graph (right) of the SCM $\mathcal{M}$ of Example 2.2.3.

Example 2.2.3 *Consider an SCM* $\mathcal{M} = \langle \mathbf{5}, \mathbf{3}, \mathbb{R}^5, \mathbb{R}^3, \boldsymbol{f}, \mathbb{P}_{\mathbb{R}^3} \rangle$ *with a causal mechanism given by:*

$$\begin{aligned} f_1(\boldsymbol{x}, \boldsymbol{e}) &= e_1 + e_2 \\ f_2(\boldsymbol{x}, \boldsymbol{e}) &= e_2 \\ f_3(\boldsymbol{x}, \boldsymbol{e}) &= x_1 \cdot x_2 + x_5 \\ f_4(\boldsymbol{x}, \boldsymbol{e}) &= x_4 + x_4^2 - e_3 \cdot x_2 \\ f_5(\boldsymbol{x}, \boldsymbol{e}) &= \sin(x_3)/x_4 \,, \end{aligned}$$

with $\mathbb{P}_{\mathbb{R}^3}$ *an arbitrary product probability distribution over* $\mathbb{R}^3$*. The augmented functional graph and the functional graph of* $\mathcal{M}$ *are depicted*[6] *in Figure 1.*

As is illustrated in this example, the augmented functional graph is a more verbose representation than the functional graph. Because of this, we will take in this paper the augmented functional graph as the standard graphical representation for SCMs. We will use the functional graph only in Section 2.9. For an SCM $\mathcal{M}$, we often write the sets $\text{pa}_{\mathcal{G}^a(\mathcal{M})}(\mathcal{U})$, $\text{ch}_{\mathcal{G}^a(\mathcal{M})}(\mathcal{U})$, $\text{an}_{\mathcal{G}^a(\mathcal{M})}(\mathcal{U})$, etc., for some subset $\mathcal{U} \subseteq \mathcal{I} \cup \mathcal{J}$, simply as $\text{pa}(\mathcal{U})$, $\text{ch}(\mathcal{U})$, $\text{an}(\mathcal{U})$, etc., whenever, besides the SCM $\mathcal{M}$, there is no explicit graph given.

Definition 2.2.4 *We call an SCM* $\mathcal{M}$ acyclic *if* $\mathcal{G}^a(\mathcal{M})$ *is a DAG, otherwise we call* $\mathcal{M}$ cyclic*.*

Equivalently, an SCM $\mathcal{M}$ is acyclic if $\mathcal{G}(\mathcal{M})$ is an ADMG, and otherwise it is cyclic.

Most of the existing literature considers only acyclic SCMs. In the structural equation model (SEM) literature, acyclic SCMs are referred to as *recursive* SEMs and cyclic SCMs as *non-recursive* SEMs. Particularly interesting classes of acyclic SCMs are the *Markovian* SCMs,

[6]For visualizing an (augmented) functional graph, we will stick to the common convention to use random variables, with the index set as subscript, instead of using the index set itself. With a slight abuse of notation, we will still use random variables notation in the (augmented) functional graph in the case that the SCM has no solution at all.

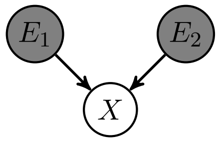

Figure 2: The augmented functional graph of the acyclic SCM $\mathcal{M}$ of Example 2.3.2.

which are acyclic SCMs for which each exogenous variable has at most one child, and the *semi-Markovian* SCMs, which are acyclic SCMs for which each exogenous variable has at most two childs. They satisfy several Markov properties (Pearl, 2009). In Section 2.9 we will discuss general Markov properties that holds for cyclic SCMs. Acyclic SCMs have a considerable technical advantage: they are always uniquely solvable, as we will see in Section 2.6.4. This property makes acyclic SCMs a convenient class to work with, which may explain the focus on acyclic SCMs in the literature.

2.3 The canonical representation

Definition 2.3.1 *An SCM $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ is called* canonical *if for all $i \in \mathcal{I}$ the components f_i are of the form $f_i : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(i)} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(i)} \to \mathcal{X}_i$.*

We already encountered a canonical SCM in Example 2.2.3. The next example illustrates that not all SCMs are canonical:

Example 2.3.2 *Consider the SCM $\mathcal{M} = \langle \mathbf{1}, \mathbf{2}, \mathbb{R}, \mathbb{R}^2, f, \mathbb{P}_{\mathbb{R}^2} \rangle$ with the causal mechanism $f(x, \boldsymbol{e}) = -x + e_1 + e_2$ and $\mathbb{P}_{\mathbb{R}^2}$ an arbitrary product probability distribution over $\mathbb{R}^2$. The augmented functional graph is depicted in Figure 2 and hence this SCM is not canonical.*

Proposition 2.3.3 *Let $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ be an SCM, then there exists an equivalent SCM $\tilde{\mathcal{M}} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \tilde{\boldsymbol{f}}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ such that each component f_i has an equivalent representation $\tilde{f}_i : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(i)} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(i)} \to \mathcal{X}_i$.*

Proof. For $i \in \mathcal{I}$, let $\{\ell_1, \ldots, \ell_n\}$ be the $n := |\mathcal{I} \setminus \mathrm{pa}(i)|$ different elements in $\mathcal{I} \setminus \mathrm{pa}(i)$ and let $\{k_1, \ldots, k_m\}$ be the $m := |\mathcal{J} \setminus \mathrm{pa}(i)|$ different elements in $\mathcal{J} \setminus \mathrm{pa}(i)$ (pick any arbitrary order). Then by Definition 2.2.1 there exists mappings $g_i^p : \boldsymbol{\mathcal{X}}_{\mathcal{I} \setminus \{\ell_1, \ldots, \ell_p\}} \times \boldsymbol{\mathcal{E}} \to \mathcal{X}_i$ and $h_i^q : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(i)} \times \boldsymbol{\mathcal{E}}_{\mathcal{J} \setminus \{k_1, \ldots, k_q\}} \to \mathcal{X}_i$ for $p = 1, \ldots, n$ and $q = 1, \ldots, m$ such that

$$f_i \equiv g_i^1 \equiv \cdots \equiv g_i^n \equiv h_i^1 \cdots \equiv h_i^m =: \tilde{f}_i \,.$$

□

That is, for every SCM there exists an equivalent canonical SCM.

Example 2.3.2 (Continued) *Let $\tilde{\mathcal{M}}$ be the SCM $\mathcal{M}$ but with the causal mechanism $\tilde{f}(x, \boldsymbol{e}) = \frac{1}{2}(e_1 + e_2)$. Then $\tilde{\mathcal{M}}$ is a canonical SCM which is equivalent to $\mathcal{M}$.*

For a causal mechanism $\boldsymbol{f} : \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{X}}$ and a subset $\mathcal{U} \subseteq \mathcal{I}$, we will write $\boldsymbol{f}_{\mathcal{U}} : \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{X}}_{\mathcal{U}}$ for the $\mathcal{U}$ components of $\boldsymbol{f}$. The canoincal representation is the most sparse representation that is compatible with the (augmented) functional graph, in the sense that for any $\mathcal{U} \subseteq \mathcal{I}$ one can construct the mapping $\tilde{\boldsymbol{f}}_{\mathcal{U}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{U})} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{U})} \to \boldsymbol{\mathcal{X}}_{\mathcal{U}}$ such that $\tilde{\boldsymbol{f}}_{\mathcal{U}} \equiv \boldsymbol{f}_{\mathcal{U}}$. Moreover, one has for each ancestral component $\boldsymbol{f}_{\mathrm{an}(\mathcal{U})}$ an equivalent representation $\hat{\boldsymbol{f}}_{\mathrm{an}(\mathcal{U})} : \boldsymbol{\mathcal{X}}_{\mathrm{an}(\mathcal{U})} \times \boldsymbol{\mathcal{E}}_{\mathrm{an}(\mathcal{U})} \to \boldsymbol{\mathcal{X}}_{\mathrm{an}(\mathcal{U})}$. For a mapping $\boldsymbol{f}_{\mathcal{U}} : \boldsymbol{\mathcal{X}}_{\tilde{\mathcal{I}}} \times \boldsymbol{\mathcal{E}}_{\tilde{\mathcal{J}}} \to \boldsymbol{\mathcal{X}}_{\mathcal{U}}$ for some sets $\tilde{\mathcal{I}}, \tilde{\mathcal{J}}$ and $\mathcal{U}$ and for some subset $\mathcal{V} \subseteq \mathcal{U}$ we will write $\boldsymbol{f}_{\mathcal{V}}$ instead of $(\boldsymbol{f}_{\mathcal{U}})_{\mathcal{V}}$ for the $\mathcal{V}$ components of $\boldsymbol{f}_{\mathcal{U}}$.

2.4 Interventions

To define the causal semantics of SCMs, we consider here an idealized class of interventions introduced by Pearl (2009). Other types of interventions, like probabilistic interventions, mechanism changes, activity interventions, fat-hand interventions, etc. are at least as interesting, but we will not consider those here.

Definition 2.4.1 *Given an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *and a subset* $I \subseteq \mathcal{I}$ *of endogenous variables and a value* $\boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$*, the* perfect intervention $\mathrm{do}(I, \boldsymbol{\xi}_I)$ *maps* $\mathcal{M}$ *to the intervened model* $\mathcal{M}_{\mathrm{do}(I,\boldsymbol{\xi}_I)} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \tilde{\boldsymbol{f}}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *where the intervened causal mechanism* $\tilde{\boldsymbol{f}}$ *is defined by:*

$$\tilde{f}_i(\boldsymbol{x}, \boldsymbol{e}) := \begin{cases} \xi_i & i \in I \\ f_i(\boldsymbol{x}, \boldsymbol{e}) & i \in \mathcal{I} \setminus I \,. \end{cases}$$

This operation $\mathrm{do}(I, \boldsymbol{\xi}_I)$ preserves the equivalence relation (see Definition 2.1.6) on the set of all SCMs and hence this mapping descends to the set of equivalence classes of SCMs.

We can define an operation $\mathrm{do}(I)$ that operates on directed graphs:

Definition 2.4.2 *Given a directed mixed graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{B})$ *and a subset* $I \subseteq \mathcal{V}$*, we define the intervened graph* $\mathrm{do}(I)(\mathcal{G}) = (\mathcal{V}, \tilde{\mathcal{E}}, \tilde{\mathcal{B}})$ *where* $\tilde{\mathcal{E}} := \mathcal{E} \setminus (\mathcal{V} \times I)$ *and* $\tilde{\mathcal{B}} := \mathcal{B} \setminus [(\mathcal{V} \times I) \cup (I \times \mathcal{V})]$.

It simply removes all incoming edges on the nodes in I. The two notions of intervention are compatible:

Proposition 2.4.3 *Given an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *and a subset* $I \subseteq \mathcal{I}$ *of endogenous variables and a value* $\boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$*, then* $\big(\mathcal{G}^a \circ \mathrm{do}(I, \boldsymbol{\xi}_I)\big)(\mathcal{M}) = \big(\mathrm{do}(I) \circ \mathcal{G}^a\big)(\mathcal{M})$ *and* $\big(\mathcal{G} \circ \mathrm{do}(I, \boldsymbol{\xi}_I)\big)(\mathcal{M}) = \big(\mathrm{do}(I) \circ \mathcal{G}\big)(\mathcal{M})$

Proof. The $\mathrm{do}(I, \boldsymbol{\xi}_I)$ operation on $\mathcal{M}$ completely removes the functional dependence on $\boldsymbol{x}$ and $\boldsymbol{e}$ from the f_j components for $j \in I$ and hence the corresponding directed edges. □

We have the following elementary properties:

Lemma 2.4.4 *For any SCM:*

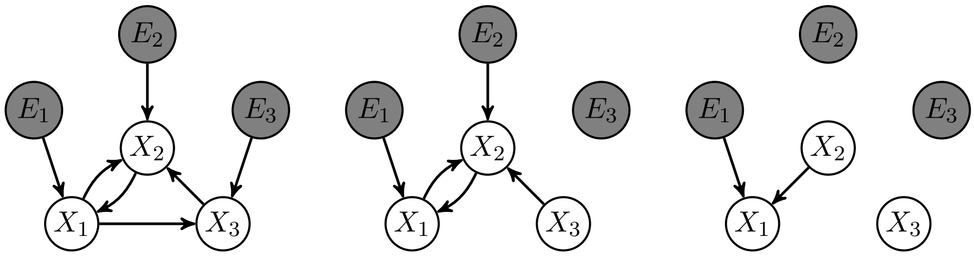

Figure 3: The augmented functional graph of the SCM of Example 2.4.5 (left), after the first intervention $\mathrm{do}(\{3\}, 1)$ (middle), and after the second intervention $\mathrm{do}(\{2\}, 1)$ (right).

1. *Perfect interventions on disjoint subsets of endogenous variables commute.*
2. *Acyclicity is preserved under perfect intervention.*

Proof. The first statement follows directly from the definitions. For the second statement, note that a perfect intervention can only remove functional parental relations, and therefore will never introduce a cycle. □

The reason why we separated the SCM from their solutions was that interventions can be defined independently of their solutions. It may for example happen that an SCM with a solution may have no solution anymore after performing a perfect intervention on the SCM. Or, it may happen that intervening on an SCM without any solution gives an SCM with a solution, as the following example illustrates.

Example 2.4.5 *Consider the SCM* $\mathcal{M} = \langle \mathbf{3}, \mathbf{3}, \mathbb{R}^3, \mathbb{R}^3, \boldsymbol{f}, \mathbb{P}_{\mathbb{R}^3} \rangle$ *with the following causal mechanism:*

$$
\begin{aligned}
f_1(\boldsymbol{x}, \boldsymbol{e}) &= x_2 + e_1 \\
f_2(\boldsymbol{x}, \boldsymbol{e}) &= x_1 + x_3 + e_2 \\
f_3(\boldsymbol{x}, \boldsymbol{e}) &= -x_1 + e_3
\end{aligned}
$$

and take for $\mathbb{P}_{\mathbb{R}^3}$ *the standard-normal measure on* $\mathbb{R}^3$*. It is easy to see that* $\mathcal{M}$ *has a solution. The augmented functional graph* $\mathcal{G}^a(\mathcal{M})$ *is depicted in Figure 3. The perfect intervention* $\mathrm{do}(\{3\}, 1)$ *gives the intervened causal mechanism*

$$
\begin{aligned}
\tilde{f}_1(\boldsymbol{x}, \boldsymbol{e}) &= x_2 + e_1 \\
\tilde{f}_2(\boldsymbol{x}, \boldsymbol{e}) &= x_1 + x_3 + e_2 \\
\tilde{f}_3(\boldsymbol{x}, \boldsymbol{e}) &= 1
\end{aligned}
$$

which does not have a solution anymore. The reverse is also possible: doing another perfect intervention $\mathrm{do}(\{2\}, 1)$ *on* $\mathcal{M}_{\mathrm{do}(\{3\},1)}$ *gives again an SCM with a solution.*

Remember that for each solution $\boldsymbol{X}$ of an SCM $\mathcal{M}$ we called the distribution $\mathbb{P}^{\boldsymbol{X}}$ an observational distribution of $\mathcal{M}$ associated to $\boldsymbol{X}$.

In the literature, the observational distribution is usually considered to be unique, as is for example the case for acyclic SCMs. However, as we will see in Section 2.6, an cyclic SCM may have different observational distributions. Whenever the intervened SCM $\mathcal{M}_{\mathrm{do}(I,\boldsymbol{\xi}_I)}$ has a solution $\boldsymbol{X}$ we call the distribution $\mathbb{P}^{\boldsymbol{X}}$ the *interventional distribution of* $\mathcal{M}$ *associated to* $\boldsymbol{X}$ *under the perfect intervention* $\mathrm{do}(I,\boldsymbol{\xi}_I)$.[7]

2.5 Counterfactuals

The causal semantics of an SCM are described by the interventions on the SCM. Similarly, one can describe the counterfactual semantics of an SCM by the interventions on the so-called *twin-SCM*. This twin-SCM was first introduced in the "twin network" method by Balke and Pearl (1994).

Definition 2.5.1 *Given an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$*, we define the* twin structural causal model (twin-SCM) *as the SCM*

$$\mathcal{M}^{\mathrm{twin}} := \langle \mathcal{I} \cup \mathcal{I}', \mathcal{J}, \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \tilde{\boldsymbol{f}}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle \,,$$

where $\mathcal{I}'$ *is a copy of* $\mathcal{I}$ *and the causal mechanism* $\tilde{\boldsymbol{f}} : \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{X}}$ *is the measurable function defined by* $\tilde{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{e}) = \big(\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{e}), \boldsymbol{f}(\boldsymbol{x}', \boldsymbol{e})\big)$.

By definition, the twin-operation on SCMs is preserves the equivalence relation $\equiv$ on SCMs.

A typical counterfactual query has the form "What is $p(\boldsymbol{X}_{L'} = \boldsymbol{x}_{L'} \,|\, \mathrm{do}(X_{I'} = \boldsymbol{\xi}_{I'}), \mathrm{do}(X_J = \boldsymbol{\eta}_J), \boldsymbol{X}_K = \boldsymbol{x}_K)$?", to be read as "Given that in the factual world, we performed a perfect intervention $\mathrm{do}(X_J = \boldsymbol{\eta}_J)$ and then observed $\boldsymbol{X}_K = \boldsymbol{x}_K$, what would be the probability of the observation $\boldsymbol{X}_{L'} = \boldsymbol{x}_{L'}$ in that counterfactual world in which we would instead have done the perfect intervention $\mathrm{do}(X_{I'} = \boldsymbol{\xi}_{I'})$?".

Consider the following example borrowed from Dawid (2002):

Example 2.5.2 *Consider the SCM* $\mathcal{M}_\rho = \langle \mathbf{2}, \mathbf{3}, \{0,1\} \times \mathbb{R}, \{0,1\} \times \mathbb{R}^2, \boldsymbol{f}, \mathbb{P}_{\{0,1\}\times\mathbb{R}^2} \rangle$ *with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}, \boldsymbol{e}) &= e_1 \\ f_2(\boldsymbol{x}, \boldsymbol{e}) &= e_{x_1+2} \end{aligned}$$

where $E_1 \sim \mathrm{Bernoulli}(1/2)$,

$$\begin{pmatrix} E_2 \\ E_3 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right),$$

and E_1 *is independent of* (E_2, E_3)*. In an epidemiological setting, this SCM could be used to model whether a patient was treated or not* (X_1) *and the corresponding outcome for that patient* (X_2)*.*

Suppose in the actual world we did not assign treatment $X_1 = 0$ *to a unit and the outcome was* $X_2 = c$*. Consider the counterfactual query "What would the outcome have been, if we had assigned*

[7] In the literature, one often finds the notation $p(\boldsymbol{x})$ and $p(\boldsymbol{x} \,|\, \mathrm{do}(\boldsymbol{X}_I = \boldsymbol{x}_I))$ for the densities of the observational and interventional distribution, respectively, in case these are uniquely defined by the SCM.

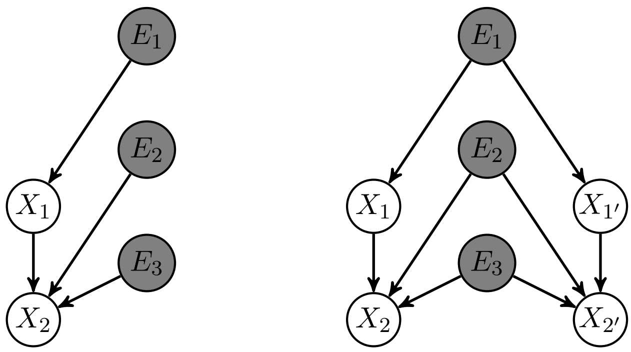

Figure 4: The augmented functional graph of the original SCM $\mathcal{M}_\rho$ (left) and the twin-SCM $\mathcal{M}_\rho^{\text{twin}}$ (right).

treatment to this unit?". We can answer this question by introducing a parallel counterfactual world that is modelled by the twin-SCM $\mathcal{M}_\rho^{\text{twin}}$*, as depicted in Figure 4. The counterfactual query then asks for* $p(X_{2'} \,|\, \text{do}(X_{1'} = 1), \text{do}(X_1 = 0), X_2 = c)$*. One can calculate that*

$$p\left(\begin{pmatrix} X_{2'} \\ X_2 \end{pmatrix} \,|\, \text{do}(X_{1'} = 1), \text{do}(X_1 = 0)\right) = \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

and hence $p(X_{2'} \,|\, \text{do}(X_{1'} = 1), \text{do}(X_1 = 0), X_2 = c) = \mathcal{N}(-\rho c, 1 - \rho^2)$*. Note that the counterfactual query depends on a quantity* ρ *that we cannot know from the observational density* $p(X_1, X_2)$ *and the interventional densities* $p(X_2 \,|\, \text{do}(X_1 = 0))$ *and* $p(X_2 \,|\, \text{do}(X_1 = 1))$ *(for example from randomized controlled trials).*

Whenever the intervened twin-SCM $\mathcal{M}^{\text{twin}}_{\text{do}(\tilde{I}, \boldsymbol{\xi}_{\tilde{I}})}$, where $\tilde{I} \subseteq \mathcal{I} \cup \mathcal{I}'$ and $\boldsymbol{\xi}_{\tilde{I}} \in \boldsymbol{\mathcal{X}}_{\tilde{I}}$ has a solution $(\boldsymbol{X}, \boldsymbol{X}')$ we call the distribution $\mathbb{P}^{(\boldsymbol{X}, \boldsymbol{X}')}$ the *counterfactual distribution of* $\mathcal{M}$ *associated to* $(\boldsymbol{X}, \boldsymbol{X}')$ *under the perfect intervention* $\text{do}(\tilde{I}, \boldsymbol{\xi}_{\tilde{I}})$.

2.6 Solvability

In this section we will describe the notion of solvability and show that it is a sufficient and necessary condition for the existence of a solution of an SCM. Moreover, we will show that there exists a sufficient and necessary condition for solvability. We will give several properties of solvability and show that all solutions of a uniquely solvable SCM induce the same observational distribution.

The notion of solvability with respect to a certain subset of the endogenous variables captures the existence of a solution of the structural equations with respect to this subset of variables given the input variables.

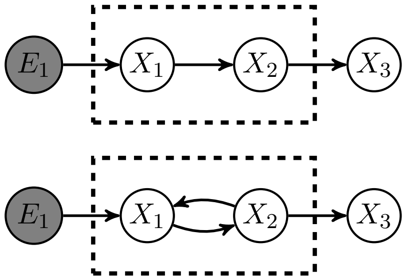

Figure 5: The augmented functional graph of the acyclic SCM from Example 2.6.1 that is solvable w.r.t. $\{1, 2\}$ (top) and of the cyclic SCM from Example 2.6.1 that can be either solvable, or not solvable, w.r.t. $\{1, 2\}$ (bottom).

Example 2.6.1 *A simple acyclic SCM is given by* $\mathcal{M} = \langle \mathbf{3}, \mathbf{1}, \mathbb{R}^3, \mathbb{R}, \boldsymbol{f}, \mathbb{P}_{\mathbb{R}} \rangle$*, where*

$$f_1(\boldsymbol{x}, e_1) = e_1^2$$
$$f_2(\boldsymbol{x}, e_1) = x_1$$
$$f_3(\boldsymbol{x}, e_1) = x_2 \,,$$

$\mathbb{P}_{\mathbb{R}}$ *is the standard-normal measure on* $\mathbb{R}$*. Consider the subset of endogenous variables* $\{1, 2\}$ *which is depicted by the box around the nodes in the augmented functional graph in Figure 5. For each input value* $e_1 \in \mathbb{R}$ *of this box, the structural equations for the variables* $\{1, 2\}$ *have a unique output for* x_1 *and* x_2*, which is given by the mapping* $\boldsymbol{g} : \mathbb{R} \to \mathbb{R}^2$ *defined by* $\boldsymbol{g}(e_1) := (e_1^2, e_1^2)$*. The existence of such a mapping means that* $\mathcal{M}$ *is solvable w.r.t.* $\{1, 2\}$*. Changing the causal mechanim* $\boldsymbol{f}$ *to*

$$f_1(\boldsymbol{x}, e_1) = x_2 - x_2^2 + e_1^2$$
$$f_2(\boldsymbol{x}, e_1) = x_1$$
$$f_3(\boldsymbol{x}, e_1) = x_2$$

gives an example of an cyclic SCM as depicted in Figure 5. Here there exists two mappings $\boldsymbol{g}, \tilde{\boldsymbol{g}} : \mathbb{R} \to \mathbb{R}^2$ *defined by* $\boldsymbol{g}(e_1) := (|e_1|, |e_1|)$ *and* $\tilde{\boldsymbol{g}}(e_1) := (-|e_1|, -|e_1|)$ *that make* $\mathcal{M}$ *solvable w.r.t.* $\{1, 2\}$*. Taking the causal mechanism* $\boldsymbol{f}$ *instead to be*

$$f_1(\boldsymbol{x}, e_1) = x_2 + e_1^2$$
$$f_2(\boldsymbol{x}, e_1) = x_1$$
$$f_3(\boldsymbol{x}, e_1) = x_2$$

gives an cyclic SCM that is not solvable w.r.t. $\{1, 2\}$*, since the structural equations for* $\{1, 2\}$ *do not have any solution for* x_1 *and* x_2 *for* $e_1 \neq \mathbb{R}$*.*

More precisely, we define solvability with respect to a certain subset as follows:

Definition 2.6.2 *We will call an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ solvable w.r.t. $\mathcal{O} \subseteq \mathcal{I}$ *if there exists a measurable mapping* $\boldsymbol{g}_{\mathcal{O}} : \boldsymbol{\mathcal{X}}_{\text{pa}(\mathcal{O}) \setminus \mathcal{O}} \times$

$\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \to \boldsymbol{\mathcal{X}}_{\mathcal{O}}$ *such that for* $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$*-almost every* $\boldsymbol{e}$ *for all* $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$*:*

$$\boldsymbol{x}_{\mathcal{O}} = \boldsymbol{g}_{\mathcal{O}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}) \quad \Longrightarrow \quad \boldsymbol{x}_{\mathcal{O}} = \boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}, \boldsymbol{e})\,.$$

We will call $\mathcal{M}$ solvable *if it is solvable w.r.t.* $\mathcal{I}$*.*

By definition, solvability w.r.t. a subset respects the equivalence relation $\equiv$ on SCMs.

Intuitively, the definition of solvability w.r.t. $\mathcal{O}$ means that there exists a specific assignment on the variables $\mathcal{O}$ given the input variables $\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}$. Note that $\mathcal{M}$ is solvable iff there exists a measurable mapping $\boldsymbol{g} : \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \to \boldsymbol{\mathcal{X}}$ such that for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$\boldsymbol{x} = \boldsymbol{g}(\boldsymbol{e}_{\mathrm{pa}(\mathcal{I})}) \quad \Longrightarrow \quad \boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{e})\,.$$

Next we will show that there exists a sufficient and necessary condition for the existence of a mapping that makes $\mathcal{M}$ solvable w.r.t. a certain subset. We consider here two cases: the case where $\mathcal{M}$ is solvable w.r.t. $\mathcal{I}$ (see Section 2.6.1) and where it is solvable w.r.t. a strict subset $\mathcal{O} \subset \mathcal{I}$ (see Section 2.6.2).

2.6.1 Solvability w.r.t. $\mathcal{I}$

We start by considering the case where $\mathcal{M}$ is solvable w.r.t. $\mathcal{I}$.

Lemma 2.6.3 *If* $\mathcal{M}$ *is solvable, then there exists a solution.*

Proof. Suppose the mapping $\boldsymbol{g} : \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \to \boldsymbol{\mathcal{X}}$ makes $\mathcal{M}$ solvable, then the measurable mappings $\boldsymbol{E} : \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{E}}$ and $\boldsymbol{X} : \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{X}}$, defined by $\boldsymbol{E}(\boldsymbol{e}) = \boldsymbol{e}$ and $\boldsymbol{X}(\boldsymbol{e}) = \boldsymbol{g}(\boldsymbol{e}_{\mathrm{pa}(\mathcal{I})})$ respectively, define a pair of random variables $(\boldsymbol{E}, \boldsymbol{X})$ such that $\boldsymbol{X} = \boldsymbol{f}(\boldsymbol{X}, \boldsymbol{E})$ holds a.s. and hence $\boldsymbol{X}$ is a solution. □

Before we prove the converse, in Proposition 2.6.8, we will give a sufficient and necessary condition in terms of the solution space $\boldsymbol{\mathcal{S}}(\mathcal{M})$ of $\mathcal{M}$, where:

$$\boldsymbol{\mathcal{S}}(\mathcal{M}) := \{(\boldsymbol{e}, \boldsymbol{x}) \in \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{X}} : \boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{e})\}\,.$$

Lemma 2.6.4 *This solution space* $\boldsymbol{\mathcal{S}}(\mathcal{M})$ *of an SCM* $\mathcal{M}$ *is a measurable set.*

Proof. Consider the measurable mapping $\boldsymbol{h} : \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{X}}$ defined by $\boldsymbol{h}(\boldsymbol{e}, \boldsymbol{x}) = (\boldsymbol{x}, \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{e}))$ and let $\Delta := \{(\boldsymbol{x}, \boldsymbol{x}) : \boldsymbol{x} \in \boldsymbol{\mathcal{X}}\}$. Then Δ is measurable since $\boldsymbol{\mathcal{X}}$ is Hausdorff and hence $\boldsymbol{h}^{-1}(\Delta) = \boldsymbol{\mathcal{S}}(\mathcal{M})$ is measurable. □

There is a close relation between the solution space $\boldsymbol{\mathcal{S}}(\mathcal{M})$ together with the probability distribution $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$ and the existence of a solution of an SCM $\mathcal{M}$, as the following example illustrates.

Example 2.6.5 *Consider an SCM* $\mathcal{M} = \langle \mathbf{1}, \mathbf{1}, \mathbb{R}, \mathbb{R}, f, \mathbb{P}_{\mathbb{R}} \rangle$ *with the following causal mechanism:*

$$f(x,e) = x \cdot (1 - \mathbf{1}_{\{0\}}(e)) + 1$$

where $\mathbf{1}_{\{0\}}$ *denotes the indicator function and we take for* $\mathbb{P}_{\mathbb{R}}$ *the standard-normal measure on* $\mathbb{R}$. *The solution space* $\boldsymbol{\mathcal{S}}(\mathcal{M})$ *consists only of one point* $(e,x) = (0,1)$. *If we project this point down on* $\mathcal{E}$, *then this point has a probability of zero and hence we cannot find random variables* (E,X) *such that* $X = f(X,E)$ *holds a.s.. However, if we change the standard-normal measure* $\mathbb{P}_{\mathbb{R}}$ *to the Dirac measure* δ_0, *then it has on* $e = 0$ *a probability mass equal to* 1, *and hence one can find such pair of random variables.*

More generally, we have the following lemma:

Lemma 2.6.6 *Given an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$. *Then* $\mathcal{M}$ *is solvable iff* $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M}))$ *is a* $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$*-null set, where* $\boldsymbol{pr}_{\boldsymbol{\mathcal{E}}} : \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{E}}$ *is the projection mapping on* $\boldsymbol{\mathcal{E}}$.

Proof. By Proposition 2.3.3 there exists an equivalent canonical SCM $\tilde{\mathcal{M}}$ with causal mechanism $\tilde{\boldsymbol{f}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{I})} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \to \boldsymbol{\mathcal{X}}$. By definition $\boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M}))$ and $\boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\tilde{\mathcal{M}}))$ differ by a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set and hence $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M}))$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set iff $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\tilde{\mathcal{M}}))$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set. Writing $\boldsymbol{\mathcal{S}}(\tilde{\mathcal{M}}) = \boldsymbol{\mathcal{E}}_{\setminus \mathrm{pa}(\mathcal{I})} \times \tilde{\boldsymbol{\mathcal{S}}}$, where

$$\tilde{\boldsymbol{\mathcal{S}}} := \left\{ (\boldsymbol{e}_{\mathrm{pa}(\mathcal{I})}, \boldsymbol{x}) \in \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \times \boldsymbol{\mathcal{X}} : \boldsymbol{x} = \tilde{\boldsymbol{f}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{I})}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{I})}) \right\},$$

gives that $\boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\tilde{\mathcal{M}})) = \boldsymbol{\mathcal{E}}_{\setminus \mathrm{pa}(\mathcal{I})} \times \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})}}(\tilde{\boldsymbol{\mathcal{S}}})$ and hence $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\tilde{\mathcal{M}})) = \boldsymbol{\mathcal{E}}_{\setminus \mathrm{pa}(\mathcal{I})} \times (\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})}}(\tilde{\boldsymbol{\mathcal{S}}}))$. Thus $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M}))$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set iff $\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})}}(\tilde{\boldsymbol{\mathcal{S}}})$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})}}$-null set.

Now suppose that $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}})$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set, then $\mathcal{M}$ is solvable by application of the measurable selection Theorem 6.0.2.

Conversely, let

$$\boldsymbol{\mathcal{T}} := \left\{ \boldsymbol{e}_{\mathrm{pa}(\mathcal{I})} \in \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \,:\, \boldsymbol{g}(\boldsymbol{e}_{\mathrm{pa}(\mathcal{I})}) = \tilde{\boldsymbol{f}}(\boldsymbol{g}_{\mathrm{pa}(\mathcal{I})}(\boldsymbol{e}_{\mathrm{pa}(\mathcal{I})}), \boldsymbol{e}_{\mathrm{pa}(\mathcal{I})}) \right\}.$$

By assumption $\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \setminus \boldsymbol{\mathcal{T}}$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})}}$-null set and since $\boldsymbol{\mathcal{T}} \subseteq \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})}}(\tilde{\boldsymbol{\mathcal{S}}})$ we have $\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})}}(\tilde{\boldsymbol{\mathcal{S}}}) \subseteq \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \setminus \boldsymbol{\mathcal{T}}$, hence $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M}))$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set. □

Note that the projection mapping does not map measurable sets to measurable sets perse and hence $\boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M}))$ may not be measurable at all.

Corollary 2.6.7 *The SCM* $\mathcal{M}$ *is solvable iff for* $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$*-almost every* $\boldsymbol{e}$ *the structural equations*

$$\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{e})$$

have a solution $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$.

Proof. This follows directly from Lemma 2.6.6. Note that

$$\boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M})) = \{\boldsymbol{e} \in \boldsymbol{\mathcal{E}} : \exists \boldsymbol{x} \in \boldsymbol{\mathcal{X}}(\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{e}))\} .$$

□

The condition that $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M}))$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set is a sufficient and a necessary condition for $\mathcal{M}$ to be solvable and this equivalence allows us to prove the following proposition:

Proposition 2.6.8 *For an SCM $\mathcal{M}$, there exists a solution iff $\mathcal{M}$ is solvable.*

Proof. One direction follows from Lemma 2.6.3. Now suppose we have a solution, then there exists a pair of random variables $(\boldsymbol{E}, \boldsymbol{X}) : \Omega \to \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{X}}$ such that $\boldsymbol{X} = \boldsymbol{f}(\boldsymbol{X}, \boldsymbol{E})$ a.s.. Let the measurable mapping $\boldsymbol{h} : \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{X}}$ be defined by $\boldsymbol{h}(\boldsymbol{e}, \boldsymbol{x}) = (\boldsymbol{x}, \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{e}))$ and consider the measurable set:

$$\begin{aligned}\Lambda :&= (\boldsymbol{h} \circ (\boldsymbol{E}, \boldsymbol{X}))^{-1}(\Delta) \\ &= \{\boldsymbol{\omega} \in \Omega : \boldsymbol{X}(\boldsymbol{\omega}) = \boldsymbol{f}(\boldsymbol{X}(\boldsymbol{\omega}), \boldsymbol{E}(\boldsymbol{\omega}))\} .\end{aligned}$$

Then $\boldsymbol{E}(\Lambda) \subseteq \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M}))$ and thus $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M})) \subseteq \boldsymbol{\mathcal{E}} \setminus \boldsymbol{E}(\Lambda)$, moreover since $\mathbb{P}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{E}(\Lambda)) = 1$ it follows that $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M}))$ lies in the $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{E}(\Lambda)$. The result follows by applying Lemma 2.6.6. □

This proposition implies that for a solution $\boldsymbol{X} : \Omega \to \boldsymbol{\mathcal{X}}$, there necessarily exists a random variable $\boldsymbol{E} : \Omega \to \boldsymbol{\mathcal{E}}$ and a mapping $\boldsymbol{g} : \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \to \boldsymbol{\mathcal{X}}$ such that $\boldsymbol{g}(\boldsymbol{E}_{\mathrm{pa}(\mathcal{I})})$ is a solution. However, it does not imply that there necessarily exists a random variable $\boldsymbol{E} : \Omega \to \boldsymbol{\mathcal{E}}$ and a mapping $\boldsymbol{g} : \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \to \boldsymbol{\mathcal{X}}$ such that $\boldsymbol{X} = \boldsymbol{g}(\boldsymbol{E}_{\mathrm{pa}(\mathcal{I})})$ holds a.s..

Example 2.6.9 *Consider an SCM $\mathcal{M} = \langle \mathbf{1}, \emptyset, \mathbb{R}, 1, f, \mathbb{P}_1 \rangle$ with the causal mechanism $f : \mathcal{X} \times \mathcal{E} \to \mathcal{X}$ defined by $f(x, e) = x - x^2 + 1$, where $\mathbb{P}_1$ is the trivial probability measure on a point. There exists only two mappings $g_\pm : \mathcal{E} \to \mathcal{X}$, defined by $g_\pm(e) = \pm 1$ that makes $\mathcal{M}$ solvable. Any random variable X such that $X \in \{\pm 1\}$ is a solution of $\mathcal{M}$, however it is only almost surely equal to $g_\pm(E)$ if $\mathbb{P}(X = \pm 1) = 1$.*

2.6.2 Solvability w.r.t. a subset

We have seen that for solvability there exists a sufficient and necessary condition in terms of the solution space of the SCM. However, in general there does not exist a similar necessary condition for solvability w.r.t. a strict subset $\mathcal{O} \subset \mathcal{I}$. This is due to the fact that in general there does not exist a measurable section of a measurable set (see chapter 18 of Kechris (1995)). Consider the measurable set:

$$\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M}) = \{(\boldsymbol{e}, \boldsymbol{x}) \in \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{X}} \,:\, \boldsymbol{x}_{\mathcal{O}} = \boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}, \boldsymbol{e}))\} .$$

Then in general there does not exist a measurable mapping $\boldsymbol{g}_{\mathcal{O}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \to \boldsymbol{\mathcal{X}}_{\mathcal{O}}$ such that for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x}_{\setminus \mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\setminus \mathcal{O}}$ we have $(\boldsymbol{e}, \boldsymbol{x}_{\setminus \mathcal{O}}, \boldsymbol{g}_{\mathcal{O}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})})) \in \boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M})$. However, this mapping exists under the following sufficient condition:

Proposition 2.6.10 *Given an SCM $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ and a subset $\mathcal{O} \subseteq \mathcal{I}$. If for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x}_{\setminus \mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\setminus \mathcal{O}}$ the fiber $\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M})_{(\boldsymbol{e}, \boldsymbol{x}_{\setminus \mathcal{O}})}$ over $(\boldsymbol{e}, \boldsymbol{x}_{\setminus \mathcal{O}})$ given by*

$$\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M})_{(\boldsymbol{e}, \boldsymbol{x}_{\setminus \mathcal{O}})} := \{\boldsymbol{x}_{\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathcal{O}} : \boldsymbol{x}_{\mathcal{O}} = \boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}, \boldsymbol{e})\}$$

is non-empty and σ-compact, then $\mathcal{M}$ is solvable w.r.t. $\mathcal{O}$.

Proof. Consider the equivalent canonical SCM $\tilde{\mathcal{M}}$ with the canonical causal mechanism $\tilde{\boldsymbol{f}}$ as given in Proposition 2.3.3. Then $\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\tilde{\mathcal{M}}) = \boldsymbol{\mathcal{E}}_{\setminus \mathrm{pa}(\mathcal{O})} \times \tilde{\boldsymbol{\mathcal{S}}} \times \boldsymbol{\mathcal{X}}_{\setminus(\mathcal{O} \cup \mathrm{pa}(\mathcal{O}))}$, where

$$\begin{aligned}\tilde{\boldsymbol{\mathcal{S}}} := \{(\boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}, \boldsymbol{x}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}}, \boldsymbol{x}_{\mathcal{O}}) \in \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \times \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}} \times \boldsymbol{\mathcal{X}}_{\mathcal{O}} :\\ \boldsymbol{x}_{\mathcal{O}} = \tilde{\boldsymbol{f}}_{\mathcal{O}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O})}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})})\}\end{aligned}$$

is a measurable set. Because $\boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M}))$ and $\boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\tilde{\mathcal{M}}))$ differ by a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set, we have in particular that for $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})}}$-almost every $\boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}$ for all $\boldsymbol{x}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}}$ the fiber

$$\tilde{\boldsymbol{\mathcal{S}}}_{(\boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}, \boldsymbol{x}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}})} = \{\boldsymbol{x}_{\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathcal{O}} : \boldsymbol{x}_{\mathcal{O}} = \tilde{\boldsymbol{f}}_{\mathcal{O}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O})}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})})\}$$

is non-empty and σ-compact. In other words, we have that $\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \setminus \boldsymbol{\mathcal{K}}_{\sigma}$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})}}$-null set, where

$$\begin{aligned}\boldsymbol{\mathcal{K}}_{\sigma} := \{&\boldsymbol{e}_{\mathrm{pa}(\mathcal{O})} \in \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \, : \, \forall \boldsymbol{x}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}} \\ &(\tilde{\boldsymbol{\mathcal{S}}}_{(\boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}, \boldsymbol{x}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}})} \text{ is non-empty and } \sigma\text{-compact})\} \,,\end{aligned}$$

The result follows now directly from the second measurable selection Theorem 6.0.3. □

For many purposes, this condition of σ-compactness suffices since it contains for example all countable discrete spaces, any interval of the real line, and moreover all the Euclidean spaces. For larger fibers $\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M})_{(\boldsymbol{e}, \boldsymbol{x}_{\setminus \mathcal{O}})}$ we refer the reader to (Kechris, 1995).

Note that for the case $\mathcal{O} = \mathcal{I}$ the condition in the proof of Proposition 2.6.10 that $\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \setminus \boldsymbol{\mathcal{K}}_{\sigma}$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})}}$-null set implies in particular that $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}}(\mathcal{M}))$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set (see also Lemma 2.6.6), however the converse does not hold in general.

Example 2.6.11 *Consider an SCM $\mathcal{M} = \langle \mathbf{2}, \mathbf{2}, \mathbb{C}^2, \mathbb{R}^2, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}, \boldsymbol{e}) &= x_1 - x_1^n + x_2 + e_1 \\ f_2(\boldsymbol{x}, \boldsymbol{e}) &= e_2 \end{aligned}$$

where $n \in \mathbb{N}$ and $\mathbb{P}_{\boldsymbol{\mathcal{E}}} = \mathbb{P}^{\boldsymbol{E}}$ with $\boldsymbol{E} \sim \mathcal{N}(\mathbf{0}, \mathbb{I})$. For $n \geq 1$, this SCM is solvable w.r.t. the subset $\{1\}$, since for each $(x_2, e_1) \in \boldsymbol{\mathcal{X}}_2 \times \boldsymbol{\mathcal{E}}_1$ the fiber $\boldsymbol{\mathcal{S}}_{\{1\}}(\mathcal{M})_{(x_2, e_1)}$ consists of at most n elements. This gives n mappings $g_1 : \boldsymbol{\mathcal{X}}_2 \times \boldsymbol{\mathcal{E}}_1 \to \boldsymbol{\mathcal{X}}_1$, which each map the value (x_2, e_1) to one of the n distinct n^{th} roots of $x_2 + e_1$.

2.6.3 Several solvability properties

In general, solvability w.r.t. $\mathcal{O} \subseteq \mathcal{I}$ does not imply solvability w.r.t. a superset $\mathcal{V} \supset \mathcal{O}$ nor w.r.t. a subset $\mathcal{W} \subset \mathcal{O}$ as can be seen in the following example:

Example 2.6.12 *Consider the SCM $\mathcal{M} = \langle \mathbf{3}, \emptyset, \mathbb{R}^3, 1, \boldsymbol{f}, \mathbb{P}_1 \rangle$ with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}) &= x_1 \cdot (1 - \mathbf{1}_{\{1\}}(x_2)) + 1 \\ f_2(\boldsymbol{x}) &= x_2 \\ f_3(\boldsymbol{x}) &= x_3 \cdot (1 - \mathbf{1}_{\{-1\}}(x_2)) + 1 \,. \end{aligned}$$

This SCM is solvable w.r.t. $\{1,2\}$, $\{2,3\}$ and $\{2\}$, however it is not solvable w.r.t. $\{1\}$, $\{3\}$ and $\{1,2,3\}$.

However, solvability w.r.t. $\mathcal{O} \subseteq \mathcal{I}$ does imply solvability w.r.t. ancestral subsets $\mathcal{V} \subseteq \mathcal{O}$:

Proposition 2.6.13 *Given an SCM $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ that is solvable w.r.t. $\mathcal{O} \subseteq \mathcal{I}$. Then $\mathcal{M}$ is solvable w.r.t. $\widetilde{\mathrm{an}}_{\mathcal{G}(\mathcal{M})_{\mathcal{O}}}(\mathcal{V})$ for every $\mathcal{V} \subseteq \mathcal{O}$, where $\widetilde{\mathrm{an}}_{\mathcal{G}(\mathcal{M})_{\mathcal{O}}}(\mathcal{V})$ are the ancestors of $\mathcal{V}$ according to the induced subgraph $\mathcal{G}(\mathcal{M})_{\mathcal{O}}$ of the augmented functional graph $\mathcal{G}^a(\mathcal{M})$ on $\mathcal{O}$.*

Proof. Solvability of $\mathcal{M}$ w.r.t. $\mathcal{O}$ implies that there exists a mapping $\boldsymbol{g}_{\mathcal{O}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \to \boldsymbol{\mathcal{X}}_{\mathcal{O}}$ such that for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$\boldsymbol{x}_{\mathcal{O}} = \boldsymbol{g}_{\mathcal{O}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}) \quad \Longrightarrow \quad \boldsymbol{x}_{\mathcal{O}} = \boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}, \boldsymbol{e}) \,.$$

Let $\mathcal{P} := \widetilde{\mathrm{an}}_{\mathcal{G}(\mathcal{M})_{\mathcal{O}}}(\mathcal{V})$ for some $\mathcal{V} \subseteq \mathcal{O}$, then by choosing an equivalent causal mechanism $\tilde{\boldsymbol{f}}_{\mathcal{P}}$ and $\tilde{\boldsymbol{f}}_{\mathcal{O} \setminus \mathcal{P}}$ that only depend on their parents we have that for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$\begin{cases} \boldsymbol{x}_{\mathcal{P}} &= \boldsymbol{g}_{\mathcal{P}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}) \\ \boldsymbol{x}_{\mathcal{O} \setminus \mathcal{P}} &= \boldsymbol{g}_{\mathcal{O} \setminus \mathcal{P}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}) \end{cases} \quad \Longrightarrow \quad \begin{cases} \boldsymbol{x}_{\mathcal{P}} &= \tilde{\boldsymbol{f}}_{\mathcal{P}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{P})}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{P})}) \\ \boldsymbol{x}_{\mathcal{O} \setminus \mathcal{P}} &= \tilde{\boldsymbol{f}}_{\mathcal{O} \setminus \mathcal{P}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O} \setminus \mathcal{P})}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O} \setminus \mathcal{P})}) \,. \end{cases}$$

Since for the endogenous variables $\mathrm{pa}(\mathcal{P}) \subseteq \mathcal{P} \cup (\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O})$, we have that in particular for $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})}}$-almost every $\boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}$ for all $\boldsymbol{x}_{\mathcal{P} \cup (\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O})} \in \boldsymbol{\mathcal{X}}_{\mathcal{P} \cup (\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O})}$:

$$\boldsymbol{x}_{\mathcal{P}} = \boldsymbol{g}_{\mathcal{P}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}) \quad \Longrightarrow \quad \boldsymbol{x}_{\mathcal{P}} = \tilde{\boldsymbol{f}}_{\mathcal{P}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{P})}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{P})}) \,.$$

This implies that the mapping $\boldsymbol{g}_{\mathcal{P}}$ cannot depend on elements different from $\mathrm{pa}(\mathcal{P})$, because if it does, it leads to a contradiction. Moreover, it follows from the definition of $\mathcal{P}$ that $(\mathrm{pa}(\mathcal{O}) \setminus \mathcal{O}) \cap \mathrm{pa}(\mathcal{P}) = \mathrm{pa}(\mathcal{P}) \setminus \mathcal{P}$. Hence, we have for $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{P})}}$-almost every $\boldsymbol{e}_{\mathrm{pa}(\mathcal{P})}$ for all $\boldsymbol{x}_{\mathcal{P} \cup \mathrm{pa}(\mathcal{P})} \in \boldsymbol{\mathcal{X}}_{\mathcal{P} \cup \mathrm{pa}(\mathcal{P})}$:

$$\boldsymbol{x}_{\mathcal{P}} = \boldsymbol{g}_{\mathcal{P}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{P}) \setminus \mathcal{P}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{P})}) \quad \Longrightarrow \quad \boldsymbol{x}_{\mathcal{P}} = \tilde{\boldsymbol{f}}_{\mathcal{P}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{P})}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{P})}) \,.$$

which is equivalent to the statement that $\mathcal{M}$ is solvable w.r.t. $\widetilde{\text{an}}_{\mathcal{G}(\mathcal{M})_{\mathcal{O}}}(\mathcal{V})$.
□

In general, solvability is not preserved under union and intersection. Example 2.6.12 gives an example where solvability is not preserved under union. Even for the union of disjoint subsets, solvability is not preserved (see Example 2.7.2 with $\alpha = \beta = 1$). The next example illustrates that solvability is in general not preserved under intersection.

Example 2.6.14 *Consider the SCM* $\mathcal{M} = \langle \mathbf{3}, \emptyset, \mathbb{R}^3, 1, \boldsymbol{f}, \mathbb{P}_1 \rangle$ *with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}) &= 0 \\ f_2(\boldsymbol{x}) &= x_2 \cdot (1 - \mathbf{1}_{\{0\}}(x_1 \cdot x_3)) + 1 \\ f_3(\boldsymbol{x}) &= 0 \,. \end{aligned}$$

Then $\mathcal{M}$ *is solvable w.r.t.* $\{1, 2\}$ *and* $\{2, 3\}$*, however it is not solvable w.r.t. their intersection.*

2.6.4 Unique Solvability

Definition 2.6.15 *We will call an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ uniquely solvable w.r.t. $\mathcal{O} \subseteq \mathcal{I}$, *if there exists a measurable mapping* $\boldsymbol{g}_{\mathcal{O}} : \boldsymbol{\mathcal{X}}_{\text{pa}(\mathcal{O}) \setminus \mathcal{O}} \times \boldsymbol{\mathcal{E}}_{\text{pa}(\mathcal{O})} \to \boldsymbol{\mathcal{X}}_{\mathcal{O}}$ *such that for* $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$*-almost every* $\boldsymbol{e}$ *for all* $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$*:*

$$\boldsymbol{x}_{\mathcal{O}} = \boldsymbol{g}_{\mathcal{O}}(\boldsymbol{x}_{\text{pa}(\mathcal{O}) \setminus \mathcal{O}}, \boldsymbol{e}_{\text{pa}(\mathcal{O})}) \quad \Longleftrightarrow \quad \boldsymbol{x}_{\mathcal{O}} = \boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}, \boldsymbol{e}) \,.$$

We will call $\mathcal{M}$ uniquely solvable *if it is uniquely solvable w.r.t.* $\mathcal{I}$*.*

Note that if $\mathcal{M} \equiv \tilde{\mathcal{M}}$ and $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{O}$, then $\tilde{\mathcal{M}}$ is as well, and the same mapping $\boldsymbol{g}_{\mathcal{O}}$ makes both $\mathcal{M}$ and $\tilde{\mathcal{M}}$ uniquely solvable w.r.t. $\mathcal{O}$.

In Section 2.6.1 we gave a sufficient and necessary condition for solvability w.r.t. $\mathcal{I}$ in terms of the solution space. We saw in Section 2.6.2 that in general there does not exist a similar necessary condition for solvability w.r.t. a strict subset $\mathcal{O} \subset \mathcal{I}$. The next proposition shows that under the additional uniqueness assumption there does exist a sufficient and necessary condition for unique solvability w.r.t. any subset.

Proposition 2.6.16 *Given an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *and a subset* $\mathcal{O} \subseteq \mathcal{I}$*. Then for* $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$*-almost every* $\boldsymbol{e}$ *for all* $\boldsymbol{x}_{\setminus \mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\setminus \mathcal{O}}$*, the fiber* $\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M})_{(\boldsymbol{e}, \boldsymbol{x}_{\setminus \mathcal{O}})}$ *over* $(\boldsymbol{e}, \boldsymbol{x}_{\setminus \mathcal{O}})$ *given by*

$$\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M})_{(\boldsymbol{e}, \boldsymbol{x}_{\setminus \mathcal{O}})} := \{ \boldsymbol{x}_{\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathcal{O}} : \boldsymbol{x}_{\mathcal{O}} = \boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}, \boldsymbol{e})) \} \,.$$

is a singleton iff $\mathcal{M}$ *is uniquely solvable w.r.t.* $\mathcal{O}$*.*

Proof. Suppose the mapping $\boldsymbol{g}_{\mathcal{O}}$ makes $\mathcal{M}$ uniquely solvable w.r.t. $\mathcal{O}$. Let $\tilde{\mathcal{M}}$ be the equivalent canonical SCM with the canonical causal mechanism $\tilde{\boldsymbol{f}}$ as given in Proposition 2.3.3. Then, for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x}_{\setminus\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\setminus\mathcal{O}}$ we have that

$$\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\tilde{\mathcal{M}})_{(\boldsymbol{e},\boldsymbol{x}_{\setminus\mathcal{O}})} = \{\boldsymbol{x}_{\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathcal{O}} : \boldsymbol{x}_{\mathcal{O}} = \boldsymbol{g}_{\mathcal{O}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})})\}$$

is a singleton. Moreover, for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x}_{\setminus\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\setminus\mathcal{O}}$

$$\boldsymbol{x}_{\mathcal{O}} \in \boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M})_{(\boldsymbol{e},\boldsymbol{x}_{\setminus\mathcal{O}})} \quad\iff\quad \boldsymbol{x}_{\mathcal{O}} \in \boldsymbol{\mathcal{S}}_{\mathcal{O}}(\tilde{\mathcal{M}})_{(\boldsymbol{e},\boldsymbol{x}_{\setminus\mathcal{O}})}\,,$$

hence $\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M})_{(\boldsymbol{e},\boldsymbol{x}_{\setminus\mathcal{O}})}$ is a singleton. Conversely, by Proposition 2.6.10 there exists a mapping $\boldsymbol{g}_{\mathcal{O}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \to \boldsymbol{\mathcal{X}}_{\mathcal{O}}$ that makes $\mathcal{M}$ solvable w.r.t. $\mathcal{O}$. Suppose we have two such mappings $\boldsymbol{g}_{\mathcal{O}}$ and $\tilde{\boldsymbol{g}}_{\mathcal{O}}$, then for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$, for all $\boldsymbol{x}_{\setminus\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\setminus\mathcal{O}}$ we have that both $\boldsymbol{g}_{\mathcal{O}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})})$ and $\tilde{\boldsymbol{g}}_{\mathcal{O}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})})$ lie in the fiber $\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M})_{(\boldsymbol{e},\boldsymbol{x}_{\setminus\mathcal{O}})}$. Since $\boldsymbol{\mathcal{S}}_{\mathcal{O}}(\mathcal{M})_{(\boldsymbol{e},\boldsymbol{x}_{\setminus\mathcal{O}})}$ is a singleton, the mappings $\boldsymbol{g}_{\mathcal{O}}$ and $\tilde{\boldsymbol{g}}_{\mathcal{O}}$ have to be equivalent. Hence $\mathcal{M}$ has to be uniquely solvable w.r.t. $\mathcal{O}$. □

Analogous to Corollary 2.6.7 we have:

Corollary 2.6.17 *The SCM $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ is uniquely solvable w.r.t. $\mathcal{O} \subseteq \mathcal{I}$ iff for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x}_{\setminus\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\setminus\mathcal{O}}$ the structural equations*

$$\boldsymbol{x}_{\mathcal{O}} = \boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}, \boldsymbol{e})$$

have a unique solution $\boldsymbol{x}_{\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathcal{O}}$.

Proof. This result follows directly from Proposition 2.6.16. □

Unique solvability w.r.t. $\mathcal{O}$ implies in particular solvability w.r.t. $\mathcal{O}$. Hence unique solvability in turn implies the existence of a solution. In fact, if an SCM $\mathcal{M}$ is uniquely solvable, then the solution space $\boldsymbol{\mathcal{S}}(\mathcal{M})$ can be $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-uniquely parametrized by the mapping $\boldsymbol{g} : \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \to \boldsymbol{\mathcal{X}}$ and we will show in Lemma 2.6.19 that this always leads to a unique observational distribution. This is already known for acyclic SCMs (Pearl, 2009), and follows also from the fact that:

Proposition 2.6.18 *Every acyclic SCM $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ is uniquely solvable w.r.t. any subset $\mathcal{O} \subseteq \mathcal{I}$.*

Proof. For the equivalent canonical SCM $\tilde{\mathcal{M}}$ consider the structural equations for the variables $\mathcal{O}$. Then the acyclicity of the augmented functional graph implies the existence of a topological ordening on the nodes $\mathcal{O}$. Following this topological ordering, we can substitute the components of the causal mechanism into each other. This gives the mapping $\boldsymbol{g}_{\mathcal{O}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}} \times \boldsymbol{\mathcal{E}}_{\mathrm{an}(\mathcal{O})} \to \boldsymbol{\mathcal{X}}_{\mathcal{O}}$ which makes $\tilde{\mathcal{M}}$ and hence $\mathcal{M}$ uniquely solvable w.r.t. $\mathcal{O}$. Note that this mapping $\boldsymbol{g}_{\mathcal{O}}$ is independent of the choice of the topological ordering. □

Lemma 2.6.19 *Consider an SCM $\mathcal{M}$ that is uniquely solvable, then there exists a solution, and all solutions have the same observational distribution.*

Proof. The existence of a solution follows directly from Proposition 2.6.8. Let the random variable $\boldsymbol{E}$ be such that $\mathbb{P}^{\boldsymbol{E}} = \mathbb{P}_{\boldsymbol{\mathcal{E}}}$, then $\boldsymbol{g}(\boldsymbol{E}_{\text{pa}(\mathcal{I})})$ is a solution of $\mathcal{M}$, where $\boldsymbol{g}$ is the mapping that makes $\mathcal{M}$ uniquely solvable. For any other solution $\boldsymbol{X}$ of $\mathcal{M}$ we have by unique solvability that $\boldsymbol{X} = \boldsymbol{g}(\boldsymbol{E}_{\text{pa}(\mathcal{I})})$ a.s.. That is, for any solution $\boldsymbol{X}$ we have that the observational distribution $\mathbb{P}^{\boldsymbol{X}}$ is the push-forward of $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\text{pa}(\mathcal{I})}}$ under $\boldsymbol{g}$. □

This property that all solutions of uniquely solvable SCMs induce the same observational distribution is why they form a convenient class of SCMs. However, we would like to stress that this does not mean that all solutions of a uniquely solvable SCM $\mathcal{M}$ are almost surely equal to each other, as the following example illustrates:

Example 2.6.20 *Consider the uniquely solvable SCM $\mathcal{M} = \langle \mathbf{1}, \mathbf{1}, \mathcal{X}, \mathcal{E}, f, \mathbb{P}_{\mathcal{E}} \rangle$, where we take $\mathcal{X} = \mathcal{E} = \{-1, 1\}$ with the discrete σ-algebra $2^{\mathcal{E}}$, the causal mechanism $f(x, e) = e$ and $\mathbb{P}_{\mathcal{E}}(e = 1) = \frac{1}{2}$. If we take as background probability space $(\Omega, \mathcal{F}, \mathbb{P}) = (\mathcal{E}, 2^{\mathcal{E}}, \mathbb{P}_{\mathcal{E}})$, then the pair of random variables $(E, X) = (Id_\Omega, Id_\Omega)$ and $(\tilde{E}, \tilde{X}) = (-Id_\Omega, -Id_\Omega)$ define both a solution of $\mathcal{M}$ for which $\mathbb{P}^X = \mathbb{P}^{\tilde{X}}$. However, X and $\tilde{X}$ are not almost surely equal.*

Some of the solvability properties discussed in Section 2.6.3 also hold for unique solvability. We first have that in general unique solvability w.r.t. $\mathcal{O} \subseteq \mathcal{I}$ does not imply unique solvability w.r.t. a superset nor a subset of $\mathcal{O}$ (see Example 2.6.12). Next, we have that in general unique solvability is not preserved under union and intersection (see Examples 2.6.12 and 2.6.14). However, we do not have a similar property to the one stated in Proposition 2.6.13, if we restrict it to the case of unique solvability. That is, in general, unique solvability w.r.t. a subset $\mathcal{O} \subseteq \mathcal{I}$ does not imply unique solvability w.r.t. an ancestral subset $\mathcal{V} \subseteq \mathcal{O}$, as the following example illustrates:

Example 2.6.21 *Consider the SCM $\mathcal{M} = \langle \mathbf{3}, \mathbf{1}, \mathbb{R}^3, \mathbb{R}, \boldsymbol{f}, \mathbb{P}_{\mathbb{R}} \rangle$ with the causal mechanism:*

$$
\begin{aligned}
f_1(\boldsymbol{x}, e) &= x_1 \cdot (1 - \mathbf{1}_{\{0\}}(x_2 - x_3)) + 1 \\
f_2(\boldsymbol{x}, e) &= x_2 \\
f_3(\boldsymbol{x}, e) &= e\,,
\end{aligned}
$$

where $\mathbb{P}_{\mathbb{R}}$ the standard-normal measure on $\mathbb{R}$. This SCM is uniquely solvable w.r.t. the set $\{1, 2\}$, and thus solvable w.r.t. this set. However, although it is solvable w.r.t. the ancestral subset $\{2\}$ of $\{1, 2\}$ it is not uniquely solvable w.r.t. this subset.

Linear SCMs Linear SCMs form a special class that has seen much attention in the literature (see e.g. Bollen, 1989).

Definition 2.6.22 *We call $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \mathbb{R}^{\mathcal{I}}, \mathbb{R}^{\mathcal{J}}, \boldsymbol{f}, \mathbb{P}_{\mathbb{R}^{\mathcal{J}}} \rangle$* linear*, if each causal mechanism is a linear combination of its endogenous and exoge-*

nous variables:

$$f_i(\boldsymbol{x}, \boldsymbol{e}) = \sum_{j \in \mathcal{I}} B_{ij} x_j + \sum_{k \in \mathcal{J}} \Gamma_{ik} e_k \,,$$

where $i \in \mathcal{I}$, and $B \in \mathbb{R}^{\mathcal{I} \times \mathcal{I}}$ and $\Gamma \in \mathbb{R}^{\mathcal{I} \times \mathcal{J}}$ are matrices, and where $\mathbb{P}_{\mathbb{R}^{\mathcal{J}}}$ can be any product probability measure[8] *on $\mathbb{R}^{\mathcal{J}}$.*

For a subset $\mathcal{O} \subseteq \mathcal{I}$ we will also use the shorthand vector-notation:

$$\boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}, \boldsymbol{e}) = B_{\mathcal{O}\mathcal{I}} \boldsymbol{x} + \Gamma_{\mathcal{O}\mathcal{J}} \boldsymbol{e} \,.$$

In the literature, the coefficient B_{ij} is often referred to as the "causal effect" of x_j to x_i. See Section 2.8 for a discussion about the direct causal interpretation of cyclic SCMs.

We have the following result:

Lemma 2.6.23 *Given a linear SCM $\mathcal{M}$, a subset $\mathcal{O} \subseteq \mathcal{I}$ and let $\mathcal{L} := \mathcal{I} \setminus \mathcal{O}$. Then $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{O}$ if and only if the matrix $\mathbb{I}_{\mathcal{O}\mathcal{O}} - B_{\mathcal{O}\mathcal{O}}$ is invertible. Moreover, if $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{O}$, then the mapping $\boldsymbol{g}_{\mathcal{O}} : \mathbb{R}^{\mathcal{L}} \times \mathbb{R}^{\mathcal{J}} \to \mathbb{R}^{\mathcal{O}}$ given by*

$$\boldsymbol{g}_{\mathcal{O}}(\boldsymbol{x}_{\mathcal{L}}, \boldsymbol{e}) = (\mathbb{I}_{\mathcal{O}\mathcal{O}} - B_{\mathcal{O}\mathcal{O}})^{-1} B_{\mathcal{O}\mathcal{L}} \boldsymbol{x}_{\mathcal{L}} + (\mathbb{I}_{\mathcal{O}\mathcal{O}} - B_{\mathcal{O}\mathcal{O}})^{-1} \Gamma_{\mathcal{O}\mathcal{J}} \boldsymbol{e} \,,$$

makes $\mathcal{M}$ uniquely solvable w.r.t. $\mathcal{O}$.

Proof. $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{O}$ iff for $\mathbb{P}_{\mathbb{R}^{\mathcal{J}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x}_{\mathcal{L}} \in \boldsymbol{\mathcal{X}}_{\mathcal{L}}$:

$$\begin{aligned}
& \boldsymbol{x}_{\mathcal{O}} = \boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}, \boldsymbol{e}) \\
& \iff \boldsymbol{x}_{\mathcal{O}} = B_{\mathcal{O}\mathcal{O}} \boldsymbol{x}_{\mathcal{O}} + B_{\mathcal{O}\mathcal{L}} \boldsymbol{x}_{\mathcal{L}} + \Gamma_{\mathcal{O}\mathcal{J}} \boldsymbol{e} \\
& \iff (\mathbb{I}_{\mathcal{O}\mathcal{O}} - B_{\mathcal{O}\mathcal{O}}) \boldsymbol{x}_{\mathcal{O}} = B_{\mathcal{O}\mathcal{L}} \boldsymbol{x}_{\mathcal{L}} + \Gamma_{\mathcal{O}\mathcal{J}} \boldsymbol{e}
\end{aligned}$$

has a unique solution $\boldsymbol{x}_{\mathcal{O}} \in \mathcal{X}_{\mathcal{O}}$. Hence $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{O}$ iff $\mathbb{I}_{\mathcal{O}\mathcal{O}} - B_{\mathcal{O}\mathcal{O}}$ is invertible. □

2.6.5 The obstruction to a self-loop

The obstruction to the existence of a self-loop at an endogenous variable is unique solvability w.r.t. that variable.

Proposition 2.6.24 *The SCM $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ is uniquely solvable w.r.t. $\{i\}$ for $i \in \mathcal{I}$ iff $\mathcal{G}^a(\mathcal{M})$ (or $\mathcal{G}(\mathcal{M})$) has no self-loop at $i \in \mathcal{I}$.*

Proof. By unique solvability w.r.t. $\{i\}$ we have for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$x_i = g_i(\boldsymbol{x}_{\mathrm{pa}(i) \setminus \{i\}}, \boldsymbol{e}_{\mathrm{pa}(i)}) \qquad \iff \qquad x_i = f_i(\boldsymbol{x}, \boldsymbol{e}) \,,$$

from which the result follows. □

The obstruction to the existence of any self-loop, can be described by the following notion:

[8] Note that we do not assume that the probability measure $\mathbb{P}_{\mathbb{R}^{\mathcal{J}}}$ is Gaussian.

Definition 2.6.25 *We call an SCM $\mathcal{M}$* structurally uniquely solvable,[9] *if for all $i \in \mathcal{I}$ the SCM $\mathcal{M}$ is uniquely solvable w.r.t. $\{i\}$.*

That is, a structurally uniquely solvable SCM $\mathcal{M}$ does not have any self-loops. In particular, every acyclic SCM is structurally uniquely solvable.

2.6.6 Interventions

We saw already in Example 2.4.5 that (unique) solvability is not preserved under perfect intervention. Moreover, a uniquely solvable SCM can lead to a not uniquely solvable SCM after intervention, which either has no solution at all or has multiple solutions.

Example 2.6.26 *Consider the SCM $\mathcal{M} = \langle \mathbf{2}, \emptyset, \mathbb{R}^2, \boldsymbol{f}, \mathbb{P}_1 \rangle$ with the following causal mechanism:*

$$f_1(\boldsymbol{x}) = x_1 + x_1^2 - x_2 + 1$$
$$f_2(\boldsymbol{x}) = x_2(1 - \mathbf{1}_{\{0\}}(x_1)) + 1\,.$$

This SCM is uniquely solvable, however doing a perfect intervention $\mathrm{do}(\{1\}, \xi)$ *for some $\xi \neq 0$, leads to an intervened model $\mathcal{M}_{\mathrm{do}(\{1\},\xi)}$ without any solution. Doing instead the perfect intervention* $\mathrm{do}(\{2\}, \xi)$ *for some $\xi > 1$ leads also to a not uniquely solvable SCM $\mathcal{M}_{\mathrm{do}(\{2\},\xi)}$. However, this intervened model $\mathcal{M}_{\mathrm{do}(\{2\},\xi)}$ has two solutions $(X_1, X_2) = (\pm\sqrt{\xi - 1}, \xi)$.*

We see that unique solvability is not preserved under intervention in general, however it is known to be preserved for the class of acyclic SCMs, which follows directly from the fact that acyclicity is preserved under intervention (see Lemma 2.4.4). The fact that acyclic SCMs are closed under perfect intervention makes acyclic SCMs a convenient class of SCMs. By closer inspection, we saw in Proposition 2.6.18 that acyclic SCMs are in particular uniquely solvable w.r.t. every subset and it is this weaker property that is preserved under perfect intervention, as is shown in the next proposition.

Proposition 2.6.27 *Consider an SCM $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ that is (uniquely) solvable w.r.t. every subset. Let $I \subseteq \mathcal{I}$ and $\boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$, then $\mathcal{M}_{\mathrm{do}(I,\boldsymbol{\xi}_I)}$ is (uniquely) solvable w.r.t. every subset.*

Proof. Let $\mathcal{O} \subseteq \mathcal{I}$ and define $\mathcal{O}_1 := \mathcal{O} \setminus I$ and $\mathcal{O}_2 := \mathcal{O} \cap I$, that is $\mathcal{O} = \mathcal{O}_1 \cup \mathcal{O}_2$. Consider the mapping $\boldsymbol{g}_{\mathcal{O}_1} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O}_1)\setminus\mathcal{O}_1} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O}_1)} \to \boldsymbol{\mathcal{X}}_{\mathcal{O}_1}$ that makes $\mathcal{M}$ (uniquely) solvable w.r.t. $\mathcal{O}_1$ by assumption. Then the mapping $\tilde{\boldsymbol{g}}_{\mathcal{O}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \to \boldsymbol{\mathcal{X}}_{\mathcal{O}}$ defined by

$$\tilde{\boldsymbol{g}}_{\mathcal{O}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}) := \boldsymbol{g}_{\mathcal{O}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O}_1)\setminus\mathcal{O}}, \boldsymbol{\xi}_{\mathrm{pa}(\mathcal{O}_1)\cap\mathcal{O}_2}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O}_1)})$$
$$\tilde{\boldsymbol{g}}_{\mathcal{O}_2}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}) := \boldsymbol{\xi}_{\mathcal{O}_2}$$

makes $\mathcal{M}_{\mathrm{do}(I,\boldsymbol{\xi}_I)}$ (uniquely) solvable w.r.t. $\mathcal{O}$. □

[9] Note structural unique solvability is called structural solvability in Mooij et al. (2013) via Propoposition 2.6.28.

This proposition tells us that the class of SCMs that are (uniquely) solvable w.r.t. every subset is closed under perfect intervention. This class extends the interventionally robust class of acyclic SCMs to certain cyclic SCMs.

More generally, one can also intuitively think of the (unique) solvability w.r.t. a certain subset $\mathcal{O} \subseteq \mathcal{I}$ as a (unique) solvability property of all intervened models $\mathcal{M}_{\mathrm{do}(\mathcal{I}\setminus\mathcal{O},\boldsymbol{\xi}_{\mathcal{I}\setminus\mathcal{O}})}$ for $\boldsymbol{\xi}_{\mathcal{I}\setminus\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathcal{I}\setminus\mathcal{O}}$.

Proposition 2.6.28 *Let $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ be an SCM and let $\mathcal{O} \subseteq \mathcal{I}$. If $\mathcal{M}$ is (uniquely) solvable w.r.t. $\mathcal{O}$, then for every $\boldsymbol{\xi}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}}$ the intervened SCM $\mathcal{M}_{\mathrm{do}(\mathrm{pa}(\mathcal{O})\setminus\mathcal{O},\boldsymbol{\xi}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}})}$ is (uniquely) solvable w.r.t.* $\mathrm{pa}(\mathcal{O}) \cup \mathcal{O}$*. Moreover, for every $\boldsymbol{\xi}_{\mathcal{I}\setminus\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathcal{I}\setminus\mathcal{O}}$ the intervened SCM $\mathcal{M}_{\mathrm{do}(\mathcal{I}\setminus\mathcal{O},\boldsymbol{\xi}_{\mathcal{I}\setminus\mathcal{O}})}$ is (uniquely) solvable.*

Proof. Let $\boldsymbol{g}_{\mathcal{O}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \to \boldsymbol{\mathcal{X}}_{\mathcal{O}}$ be the mapping that makes $\mathcal{M}$ (uniquely) solvable w.r.t. $\mathcal{O}$. Then the mapping $\tilde{\boldsymbol{g}}_{\mathrm{pa}(\mathcal{O})\cup\mathcal{O}} : \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \to \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O})\cup\mathcal{O}}$ defined by $\tilde{\boldsymbol{g}}_{\mathrm{pa}(\mathcal{O})\cup\mathcal{O}}(\boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}) := (\boldsymbol{\xi}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}}, \boldsymbol{g}_{\mathcal{O}}(\boldsymbol{\xi}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}))$ makes $\mathcal{M}_{\mathrm{do}(\mathrm{pa}(\mathcal{O})\setminus\mathcal{O},\boldsymbol{\xi}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}})}$ (uniquely) solvable w.r.t. $\mathrm{pa}(\mathcal{O}) \cup \mathcal{O}$. Moreover, the mapping $\hat{\boldsymbol{g}} : \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{I})} \to \boldsymbol{\mathcal{X}}$ defined by $\hat{\boldsymbol{g}}(\boldsymbol{e}_{\mathrm{pa}(\mathcal{I})}) := (\boldsymbol{\xi}_{\mathcal{I}\setminus\mathcal{O}}, \boldsymbol{g}_{\mathcal{O}}(\boldsymbol{\xi}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{O})}))$ makes $\mathcal{M}_{\mathrm{do}(\mathcal{I}\setminus\mathcal{O},\boldsymbol{\xi}_{\mathcal{I}\setminus\mathcal{O}})}$ (uniquely) solvable. □

2.7 Equivalences

In Section 2 we already encountered an equivalence relation on the class of SCMs (see Definition 2.1.6). The (augmented) functional graph is preserved under this equivalence relation as well as their solutions and their observational, interventional and counterfactual distributions. In this section we will give several coarser equivalence relations on the class of SCMs.

2.7.1 Observational equivalence

Observational equivalence is the property that two or more SCMs are indistinguishable on the basis of their observational distributions.

Definition 2.7.1 *Two SCMs $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are* observationally equivalent with respect to $\mathcal{O} \subseteq \mathcal{I} \cap \tilde{\mathcal{I}}$*, denoted by $\mathcal{M} \equiv_{obs(\mathcal{O})} \tilde{\mathcal{M}}$, if $\boldsymbol{\mathcal{X}}_{\mathcal{O}} = \tilde{\boldsymbol{\mathcal{X}}}_{\mathcal{O}}$ and for all solutions $\boldsymbol{X}$ there exists a solution $\tilde{\boldsymbol{X}}$ such that $\mathbb{P}^{\boldsymbol{X}_{\mathcal{O}}} = \mathbb{P}^{\tilde{\boldsymbol{X}}_{\mathcal{O}}}$ and for all solutions $\tilde{\boldsymbol{X}}$ there exists a solution $\boldsymbol{X}$ such that $\mathbb{P}^{\boldsymbol{X}_{\mathcal{O}}} = \mathbb{P}^{\tilde{\boldsymbol{X}}_{\mathcal{O}}}$. They are simply called* observationally equivalent *if they are observationally equivalent with respect to $\mathcal{I} = \tilde{\mathcal{I}}$.*

Consider the following well-known example:

Example 2.7.2 *Consider the uniquely solvable SCM $\mathcal{M} = \langle \mathbf{2}, \mathbf{2}, \mathbb{R}^2, \mathbb{R}^2, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ with causal mechanism*

$$
\begin{aligned}
f_1(\boldsymbol{x}, \boldsymbol{e}) &= \alpha x_2 + e_1 \\
f_2(\boldsymbol{x}, \boldsymbol{e}) &= \beta x_1 + e_2 \,,
\end{aligned}
$$

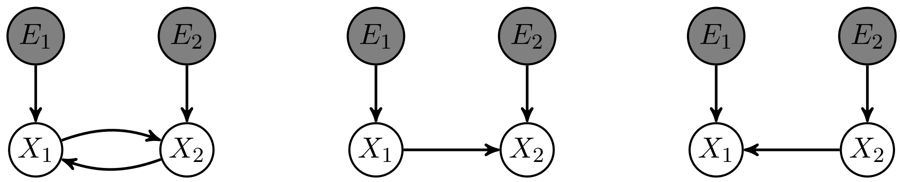

Figure 6: The augmented functional graph of the SCMs $\mathcal{M}$ (left), $\tilde{\mathcal{M}}$ (middle) and $\bar{\mathcal{M}}$ (right) of Example 2.7.2. They are observationally equivalent, but not interventionally equivalent.

where $\alpha, \beta \neq 1$ *and* $\mathbb{P}_{\mathcal{E}} = \mathbb{P}^{\boldsymbol{E}}$ *with*

$$\begin{aligned} &E_1 \sim \mathcal{N}(\mu_1, \sigma_1^2) \\ &E_2 \sim \mathcal{N}(\mu_2, \sigma_2^2) \\ &E_1 \perp\!\!\!\perp E_2 \,. \end{aligned}$$

The augmented functional graph is depicted in Figure 6. Consider also the uniquely solvable SCM $\tilde{\mathcal{M}} = \langle \mathbf{2}, \mathbf{2}, \mathbb{R}^2, \mathbb{R}^2, \tilde{\boldsymbol{f}}, \mathbb{P}_{\tilde{\mathcal{E}}} \rangle$ *with causal mechanism*

$$\begin{aligned} f_1(\boldsymbol{x}, \tilde{\boldsymbol{e}}) &= \tilde{e}_1 \\ f_2(\boldsymbol{x}, \tilde{\boldsymbol{e}}) &= \gamma x_2 + \tilde{e}_2 , \end{aligned}$$

where

$$\gamma = \frac{\beta\sigma_1^2 + \alpha\sigma_2^2}{\sigma_1^2 + \alpha^2\sigma_2^2} \,,$$

and $\mathbb{P}_{\mathcal{E}} = \mathbb{P}^{\boldsymbol{E}}$ *with*

$$\begin{aligned} &\tilde{E}_1 \sim \mathcal{N}(\tilde{\mu}_1, \tilde{\sigma}_1^2) \\ &\tilde{E}_2 \sim \mathcal{N}(\tilde{\mu}_2, \tilde{\sigma}_2^2) \\ &\tilde{E}_1 \perp\!\!\!\perp \tilde{E}_2 \,, \end{aligned}$$

where

$$\begin{aligned} \tilde{\mu}_1 &= c[\mu_1 + \alpha\mu_2], & \tilde{\sigma}_1^2 &= c^2[\sigma_1^2 + \alpha^2\sigma_2^2], \\ \tilde{\mu}_2 &= c[(\beta - \gamma)\mu_1 + (1 - \alpha\gamma)\mu_2], & \tilde{\sigma}_2^2 &= c^2[(\beta - \gamma)^2\sigma_1^2 + (1 - \alpha\gamma)^2\sigma_2^2] \end{aligned}$$

with

$$c = (1 - \alpha\beta)^{-1} \,.$$

Then $\tilde{\mathcal{M}}$ *and* $\mathcal{M}$ *are observationally equivalent. Similarly, one can define an SCM* $\bar{\mathcal{M}}$ *with augmented functional graph as depicted in Figure 6 that is observationally equivalent to both* $\tilde{\mathcal{M}}$ *and* $\mathcal{M}$*.*

Although the SCMs of this example are observationally equivalent, they are not interventionally equivalent, as we will see in the next section.

Proposition 2.7.3 *If two SCMs* $\mathcal{M}$ *and* $\tilde{\mathcal{M}}$ *are observationally equivalent w.r.t.* $\mathcal{O}$*, then they are observationally equivalent w.r.t. every subset* $\mathcal{V} \subset \mathcal{O}$*.*

Proof. Let $\mathcal{V} \subset \mathcal{O}$ and assume without loss of generality that for all solutions $\boldsymbol{X}$ there exists a solution $\tilde{\boldsymbol{X}}$ such that $\mathbb{P}^{\boldsymbol{X}_{\mathcal{O}}} = \mathbb{P}^{\tilde{\boldsymbol{X}}_{\mathcal{O}}}$, then in particular $\mathbb{P}^{\boldsymbol{X}_{\mathcal{V}}} = \mathbb{P}^{\tilde{\boldsymbol{X}}_{\mathcal{V}}}$. □

Proposition 2.7.4 *Let $\mathcal{M}$ and $\tilde{\mathcal{M}}$ be observationally equivalent w.r.t. $\mathcal{O} \subseteq \mathcal{I} \cap \tilde{\mathcal{I}}$. If $\mathcal{M}$ is solvable, then $\tilde{\mathcal{M}}$ is solvable.*

Proof. This follows directly from Proposition 2.6.8. □

If two SCMs $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are observationally equivalent, then their associated augmented functional graphs $\mathcal{G}^a(\mathcal{M})$ and $\mathcal{G}^a(\tilde{\mathcal{M}})$ are not necessarily equal to each other, as we already saw in Example 2.7.2.

2.7.2 Interventional equivalence

We consider two SCMs to be interventionally equivalent if they induce the same interventional distributions under any perfect intervention:

Definition 2.7.5 *Two SCMs $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are* interventionally equivalent with respect to $\mathcal{O} \subseteq \mathcal{I} \cap \tilde{\mathcal{I}}$, *denoted by $\mathcal{M} \equiv_{int(\mathcal{O})} \tilde{\mathcal{M}}$, if for any $I \subseteq \mathcal{O}$ and any value $\boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$ their intervened models $\mathcal{M}_{\mathrm{do}(I,\boldsymbol{\xi}_I)}$ and $\tilde{\mathcal{M}}_{\mathrm{do}(I,\boldsymbol{\xi}_I)}$ are observationally equivalent with respect to $\mathcal{O}$. They are called* interventionally equivalent *if they are interventionally equivalent with resepct to $\mathcal{I} = \tilde{\mathcal{I}}$.*

Example 2.7.2 shows clearly the difference between observational and interventional equivalence (see Figure 6). The following example shows that more complicated behavior is also possible:

Example 2.7.6 *Consider the SCM $\mathcal{M}$ from Example 2.6.26 and consider the SCM $\tilde{\mathcal{M}} = \langle \mathbf{2}, \emptyset, \mathbb{R}^2, \boldsymbol{f}, \mathbb{P}_1 \rangle$ with the following causal mechanism:*

$$\tilde{f}_1(\boldsymbol{x}) = x_1(1 - \mathbf{1}_{\{1\}}(x_2)) - x_2 + 1$$
$$\tilde{f}_2(\boldsymbol{x}) = x_2(1 - \mathbf{1}_{\{0\}}(x_1)) + 1.$$

Both SCMs are uniquely solvable and they are observationally equivalent. For every perfect intervention $\mathrm{do}(\{1\}, \xi)$ *with $\xi \in \mathbb{R}$ their intervened models are observationally equivalent, and hence $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are interventionally equivalent w.r.t. $\{1\}$. Doing a perfect intervention* $\mathrm{do}(\{2\}, 1)$ *on both models also yields observationally equivalent SCMs. However, doing a perfect intervention* $\mathrm{do}(\{2\}, \xi)$ *for some $\xi \geq 1$ leads for both to a not uniquely solvable SCM, for which $\mathcal{M}_{\mathrm{do}(\{2\},\xi)}$ is solvable and $\tilde{\mathcal{M}}_{\mathrm{do}(\{2\},\xi)}$ is non-solvable. Hence $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are not interventionally equivalent.*

On the other hand:

Proposition 2.7.7 *If two SCMs $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are interventionally equivalent w.r.t. $\mathcal{O} \subseteq \mathcal{I} \cap \tilde{\mathcal{I}}$, then they are interventionally equivalent w.r.t. every subset $\mathcal{V} \subset \mathcal{O}$.*

Proof. This follows directly from Proposition 2.7.3. □

Interventional equivalence implies in particular observational equivalence, as the empty perfect intervention ($I = \emptyset$) is a special case of a perfect intervention. We saw already in Examples 2.7.2 and 2.7.6 that interventional equivalence is a strictly stronger notion than observational equivalence. Even for this stronger notion of interventional equivalence, we have that if two SCMs $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are interventionally equivalent, then their associated augmented functional graphs $\mathcal{G}^a(\mathcal{M})$ and $\mathcal{G}^a(\tilde{\mathcal{M}})$ are not necessarily equal to each other, as is shown in the following example:

Example 2.7.8 *Consider the SCM* $\mathcal{M} = \langle \mathbf{2}, \mathbf{2}, \{-1,1\}^2, \{-1,1\}^2, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *with the causal mechanism*

$$\begin{aligned} f_1(\boldsymbol{x}, \boldsymbol{e}) &= e_1 \\ f_2(\boldsymbol{x}, \boldsymbol{e}) &= x_1 e_2 \end{aligned}$$

where $\mathbb{P}_{\boldsymbol{\mathcal{E}}} = \mathbb{P}^{\boldsymbol{E}}$ *with* $E_1, E_2 \sim \mathcal{U}(\{-1,1\})$ *and* $E_1 \perp\!\!\!\perp E_2$*. In addition, consider the SCM* $\tilde{\mathcal{M}}$ *that differs only by the causal mechanism*

$$\begin{aligned} \tilde{f}_1(\boldsymbol{x}, \boldsymbol{e}) &= e_1 \\ \tilde{f}_2(\boldsymbol{x}, \boldsymbol{e}) &= e_2 \,. \end{aligned}$$

Then $\mathcal{M}$ *and* $\tilde{\mathcal{M}}$ *are interventionally equivalent although* $\mathcal{G}^a(\mathcal{M})$ *is not equal to* $\mathcal{G}^a(\tilde{\mathcal{M}})$*. However, as we will see in Proposition 2.8.5, if* $\mathcal{M}$ *and* $\tilde{\mathcal{M}}$ *have no self-loops, their direct causal graphs are identical.*

2.7.3 Counterfactual equivalence

We consider two SCMs to be counterfactually equivalent if their twin-SCMs induce the same counterfactual distributions under any perfect intervention:

Definition 2.7.9 *Two SCMs* $\mathcal{M}$ *and* $\tilde{\mathcal{M}}$ *are* counterfactually equivalent with respect to $\mathcal{O} \subseteq \mathcal{I} \cap \tilde{\mathcal{I}}$*, denoted by* $\mathcal{M} \equiv_{cf(\mathcal{O})} \tilde{\mathcal{M}}$*, if* $\mathcal{M}^{\text{twin}}$ *and* $\tilde{\mathcal{M}}^{\text{twin}}$ *are interventionally equivalent with respect to* $\mathcal{O} \cup \mathcal{O}'$*, where* $\mathcal{O}'$ *corresponds to the copy of* $\mathcal{O}$ *in* $\mathcal{I}'$*. They are called* counterfactually equivalent *if they are counterfactually equivalent with respect to* $\mathcal{I} = \tilde{\mathcal{I}}$*.*

Proposition 2.7.10 *If two SCMs* $\mathcal{M}$ *and* $\tilde{\mathcal{M}}$ *are counterfactually equivalent w.r.t.* $\mathcal{O} \subseteq \mathcal{I} \cap \tilde{\mathcal{I}}$*, then they are counterfactually equivalent w.r.t. every subset* $\mathcal{V} \subset \mathcal{O}$*.*

Proof. This follows directly from Proposition 2.7.7. □

Lemma 2.7.11 *An SCM* $\mathcal{M}$ *is observationally equivalent w.r.t.* $\mathcal{O} \subseteq \mathcal{I}$ *to* $\mathcal{M}^{\text{twin}}$*.*

Proof. Let $\boldsymbol{X}$ be a solution of $\mathcal{M}$, then $(\boldsymbol{X}, \boldsymbol{X})$ is a solution of $\mathcal{M}^{\text{twin}}$. Conversely, let $(\boldsymbol{X}, \boldsymbol{X}')$ be a solution of $\mathcal{M}^{\text{twin}}$, then $\boldsymbol{X}$ is a solution of $\mathcal{M}$. □

Proposition 2.7.12 *If two SCMs $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are counterfactually equivalent w.r.t. $\mathcal{O} \subseteq \mathcal{I} \cap \tilde{\mathcal{I}}$, then $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are interventionally equivalent w.r.t. $\mathcal{O}$.*

Proof. Let $\mathcal{M}$ and $\tilde{\mathcal{M}}$ be counterfactually equivalent w.r.t. $\mathcal{O}$, then $\mathcal{M}^{\text{twin}}$ and $\tilde{\mathcal{M}}^{\text{twin}}$ are interventionally equivalent w.r.t. $\mathcal{O} \cup \mathcal{O}'$. In particular $\mathcal{M}^{\text{twin}}_{\text{do}(I \cup I', \boldsymbol{\xi}_{I \cup I'})}$ and $\tilde{\mathcal{M}}^{\text{twin}}_{\text{do}(I \cup I', \boldsymbol{\xi}_{I \cup I'})}$ for I' a copy of I and $\boldsymbol{\xi}_{I'} = \boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$ are observationally equivalent w.r.t. $\mathcal{O}$, where we used here Proposition 2.7.3. Since $\mathcal{M}^{\text{twin}}_{\text{do}(I \cup I', \boldsymbol{\xi}_{I \cup I'})} = (\mathcal{M}_{\text{do}(I, \boldsymbol{\xi}_I)})^{\text{twin}}$ and $\tilde{\mathcal{M}}^{\text{twin}}_{\text{do}(I \cup I', \boldsymbol{\xi}_{I \cup I'})} = (\tilde{\mathcal{M}}_{\text{do}(I, \boldsymbol{\xi}_I)})^{\text{twin}}$ we have by Lemma 2.7.11 that $\mathcal{M}_{\text{do}(I, \boldsymbol{\xi}_I)}$ and $\tilde{\mathcal{M}}_{\text{do}(I, \boldsymbol{\xi}_I)}$ are observationally equivalent w.r.t. $\mathcal{O}$. □

However, the converse is in general not true:

Example 2.7.13 *Consider the same SCMs as in Example 2.7.8. We saw in that example that they are interventionally equivalent. However, they are not counterfactually equivalent, as $\mathcal{M}^{\text{twin}}_{\text{do}(\{1',1\},(1,-1))}$ is not observationally equivalent to $\tilde{\mathcal{M}}^{\text{twin}}_{\text{do}(\{1',1\},(1,-1))}$. To see this, consider the counterfactual query $p(X_{2'} = 1 | \text{do}(X_{1'} = 1), \text{do}(X_1 = -1), X_2 = 1)$. Both SCMs will give a different answer and hence $\mathcal{M}$ and $\tilde{\mathcal{M}}$ cannot be counterfactually equivalent.*

Even interventional equivalent SCMs with the same causal mechanism may not be counterfactually equivalent. Take for example the SCMs $\mathcal{M}_\rho$ and $\mathcal{M}_{\rho'}$ with $\rho \neq \rho'$ in Example 2.5.2, then they are interventionally but not counterfactually equivalent.

If two SCMs $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are counterfactually equivalent, then their associated augmented functional graphs $\mathcal{G}^a(\mathcal{M})$ and $\mathcal{G}^a(\tilde{\mathcal{M}})$ are not necessarily equal to each other, as is shown in the following example:

Example 2.7.14 *Consider the SCM $\mathcal{M} = \langle \mathbf{1}, \mathbf{1}, \mathbb{R}, \mathbb{R}, f, \mathbb{P}_{\mathbb{R}} \rangle$ with causal mechanism $f(e) = e$ and where $\mathbb{P}_{\mathbb{R}} = \mathbb{P}^E$ with $E \sim \mathcal{N}(0,2)$ and consider the SCM $\tilde{\mathcal{M}} = \langle \mathbf{1}, \mathbf{2}, \mathbb{R}, \mathbb{R}^2, \tilde{f}, \mathbb{P}_{\mathbb{R}^2} \rangle$ with causal mechanism $\tilde{f}(\boldsymbol{e}) = e_1 + e_2$ and where $\mathbb{P}_{\mathbb{R}^2} = \mathbb{P}^{\boldsymbol{E}}$ with $E_1, E_2 \sim \mathcal{N}(0,1)$ and $E_1 \perp\!\!\!\perp E_2$. Then $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are counterfactually equivalent but their augmented functional graphs $\mathcal{G}^a(\mathcal{M})$ and $\mathcal{G}^a(\tilde{\mathcal{M}})$ differ.*

2.7.4 Relations between equivalences

The definitions of observational, interventional and counterfactual equivalence define equivalence relations on the set of all SCMs. Note that two SCMs are observationally, interventionally and counterfactually equivalent if they are so with respect to the set of endogenous variables. This means that between those SCMs the index set of exogenous variables, the space of exogenous variables, the exogenous probability distribution and the causal mechanism may all differ. The set of observational, interventional and counterfactual equivalence classes of SCMs are related in the following way:

$$\begin{aligned}
&\mathcal{M} \text{ and } \tilde{\mathcal{M}} \text{ are equivalent} \\
\Longrightarrow\ &\mathcal{M} \text{ and } \tilde{\mathcal{M}} \text{ are counterfactually equivalent} \\
\Longrightarrow\ &\mathcal{M} \text{ and } \tilde{\mathcal{M}} \text{ are interventionally equivalent} \\
\Longrightarrow\ &\mathcal{M} \text{ and } \tilde{\mathcal{M}} \text{ are observationally equivalent.}
\end{aligned}$$

Example 2.7.15 *Consider the SCM $\mathcal{M}$ and $\tilde{\mathcal{M}}$ from Example 2.7.8 (or from Example 2.7.13). In addition, consider the SCM $\hat{\mathcal{M}}$ that differs only by its causal mechanism in the following way:*

$$\begin{aligned}
f_1(\boldsymbol{x}, \boldsymbol{e}) &= e_1, & \tilde{f}_1(\boldsymbol{x}, \boldsymbol{e}) &= e_1, & \hat{f}_1(\boldsymbol{x}, \boldsymbol{e}) &= e_2 \\
f_2(\boldsymbol{x}, \boldsymbol{e}) &= x_1 e_2, & \tilde{f}_2(\boldsymbol{x}, \boldsymbol{e}) &= e_2, & \hat{f}_2(\boldsymbol{x}, \boldsymbol{e}) &= -e_1
\end{aligned}$$

where $\mathbb{P}_{\boldsymbol{\mathcal{E}}} = \mathbb{P}^{\boldsymbol{E}}$ with $E_1, E_2 \sim \mathcal{U}(\{-1, 1\})$ and $E_1 \perp\!\!\!\perp E_2$. Then $\mathcal{M}$, $\tilde{\mathcal{M}}$ and $\hat{\mathcal{M}}$ are all interventionally equivalent and thus in particular they are all observationally equivalent. Moreover, the only two SCMs that are counterfactually equivalent are the SCMs $\tilde{\mathcal{M}}$ and $\hat{\mathcal{M}}$. Nevertheless, the associated augmented functional graphs of $\mathcal{M}$, $\tilde{\mathcal{M}}$ and $\hat{\mathcal{M}}$ all differ.

In general, counterfactually, interventionally and observationally equivalent SCMs may not have the same associated augmented functional graphs, as we have shown in Examples 2.7.2, 2.7.8, 2.7.14 and 2.7.15.

2.8 The direct causal graph

For acyclic SCMs it is known that the induced subgraph of the augmented functional graph on the endogenous variables has a direct causal interpretation. For cyclic SCMs such an interpretation does not hold in general. In the acyclic setting, an endogenous variable x_i is a direct cause of x_j with respect to the endogenous variables $\mathcal{I}$, if two different interventions on x_i will lead to a change in (the probability distribution of) x_j when all the other variables in $\mathcal{I}$ besides x_i and x_j are held fixed at some value by intervention. For cyclic SCMs there is the obstruction that the possible effect x_j may have a self-loop, and hence, by Proposition 2.6.24, is not uniquely solvable w.r.t. that variable. In this section we will investigate how the existence of self-loops can obstruct a direct causal interpration of the induced subgraph of augmented functional graph on the endogenous variables and show that a structurally uniquely solvable SCM leads to a well-defined direct causal graph for cyclic SCMs.

2.8.1 Self-loops obstruct the causal semantics

Consider an SCM $\mathcal{M}$ and let x_i and x_j be different endogenous variables. Suppose that x_j has a self-loop, i.e. $\mathcal{M}$ is not uniquely solvable w.r.t. $\{j\}$, then $\mathcal{M}$ is either one of the following two cases, it is either

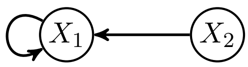

Figure 7: The augmented functional graph $\mathcal{G}^a(\mathcal{M})$ of the SCM $\mathcal{M}$ which is not solvable w.r.t. $\{1\}$ (see Example 2.8.1).

solvable w.r.t. $\{j\}$ or it is not solvable w.r.t. $\{j\}$. Performing an intervention $\operatorname{do}(\setminus j, \boldsymbol{\xi}_{\setminus j})$, for $\boldsymbol{\xi}_{\setminus j} \in \boldsymbol{\mathcal{X}}_{\setminus j}$, on all other variables yields an intervened model $\mathcal{M}_{\operatorname{do}(\setminus j, \boldsymbol{\xi}_{\setminus j})}$ that has a self-loop at the endogenous variable j and that is either solvable or not solvable, depending on if $\mathcal{M}$ is either solvable w.r.t. $\{j\}$ or not solvable w.r.t. $\{j\}$ respectively (see Proposition 2.6.28). Hence, the intervened model has either several solutions (with different induced probability distributrions) or no solution. We illustrate the direct causal interpretation of these two cases by the next two examples and show that in both cases performing two different interventions on x_i could obstruct of having a direct effect on x_j. We first consider the case where the SCM is not solvable w.r.t. $\{j\}$.

Example 2.8.1 *Consider the SCM* $\mathcal{M} = \langle \mathbf{2}, \emptyset, \mathbb{R}^2, 1\boldsymbol{f}, \mathbb{P}_1 \rangle$ *with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}) &= x_1 \cdot (1 - \mathbf{1}_{\{0\}}(x_2)) + 1 \\ f_2(\boldsymbol{x}) &= 0\,. \end{aligned}$$

This SCM is not solvable w.r.t. $\{1\}$. *Performing a perfect intervention on* x_2 *by setting it to a value different from* 0 *leads to an SCM that is not solvable and hence to an SCM without any solution. Although the augmented functional graph in Figure 7 depicts a directed edge from* x_2 *to* x_1, *this edge cannot be detected from a distributional change on* x_1 *due to a perfect intervention on* x_2, *and hence cannot be interpreted as a direct causal relation.*

The next example illustrates the case where the SCM is solvable w.r.t. $\{j\}$.

Example 2.8.2 *Consider the SCM* $\mathcal{M} = \langle \mathbf{2}, \emptyset, \mathbb{R}^2, 1, \boldsymbol{f}, \mathbb{P}_1 \rangle$ *with causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}) &= 0 \\ f_2(\boldsymbol{x}) &= x_1\,. \end{aligned}$$

Consider also the SCM $\tilde{\mathcal{M}}$ *with causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}) &= 0 \\ f_2(\boldsymbol{x}) &= x_2\,. \end{aligned}$$

Note both $\mathcal{M}$ *and* $\tilde{\mathcal{M}}$ *are solvable w.r.t.* $\{2\}$, *but only* $\mathcal{M}$ *is uniquely solvable w.r.t.* $\{2\}$. *The SCM* $\tilde{\mathcal{M}}$ *has for every interventions* $\operatorname{do}(I, \boldsymbol{\xi}_I)$, *with* $I \subseteq \mathbf{2}$ *and* $\boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$, *all the solutions of* $\mathcal{M}_{\operatorname{do}(I, \boldsymbol{\xi}_I)}$. *Performing for example the perfect intervention* $\operatorname{do}(\{1\}, -1)$ *and* $\operatorname{do}(\{1\}, 1)$ *on the SCM* $\mathcal{M}$ *gives rise to different distributions of* x_2. *We might see*

Figure 8: The augmented functional graph of the SCM $\mathcal{M}$ (left) and $\tilde{\mathcal{M}}$ (right) of Example 2.8.2.

the same distributional change in x_2 after performing these interventions on the SCM $\tilde{\mathcal{M}}$, if nature picks corresponding solutions in the intervened models of $\tilde{\mathcal{M}}$. That solutions with different interventional distributions exist is due to the fact that the SCM $\tilde{\mathcal{M}}$ is solvable w.r.t. $\{2\}$, but not uniquely solvable w.r.t. $\{2\}$. This means that the causal relations of the SCM $\tilde{\mathcal{M}}$ are underdetermined, since the model supports more possible causal relations than are described by the (augmented) functional graph (see Figure 8 on the right).

The fact that in these examples we can for an endogenous variable j under intervention on $\setminus j$, pick no solution or solutions with different distributions, is due to the property that the original SCM has a self-loop at the endogenous variable j and is either not solvable w.r.t. $\{j\}$ or solvable w.r.t. $\{j\}$ respectively (see Proposition 2.6.28 and 2.6.8). Both solvability and non-solvability w.r.t. an endogenous variables with a self-loop obstructs a direct causal interpretation of the (augmented) functional graph of the SCM. We conclude that every self-loop in an SCM obstructs a direct causal interpretation.

2.8.2 The direct causal graph

For SCMs that are structurally uniquely solvable (see Definition 2.6.25), and thus have no self-loops (see Proposition 2.6.24), we can define the notion of "direct cause" formally:

Definition 2.8.3 *Let $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ be an SCM that is structurally uniquely solvable. Then we say that $i \in \mathcal{I}$ is a* direct cause of $j \in \mathcal{I} \setminus \{i\}$ *if and only if there exist two different $\boldsymbol{\xi}_{\setminus\{j\}}$ and $\tilde{\boldsymbol{\xi}}_{\setminus\{j\}}$ in $\boldsymbol{\mathcal{X}}_{\setminus\{j\}}$ such that $\boldsymbol{\xi}_{\setminus\{i,j\}} = \tilde{\boldsymbol{\xi}}_{\setminus\{i,j\}}$, and $\mathcal{M}_{\mathrm{do}(\setminus\{j\},\boldsymbol{\xi}_{\setminus\{j\}})}$ and $\mathcal{M}_{\mathrm{do}(\setminus\{j\},\tilde{\boldsymbol{\xi}}_{\setminus\{j\}})}$ are not observationally equivalent w.r.t. $\{j\}$.*

Note that this definition of direct cause is invariant under the equivalence relation $\equiv$ on SCMs.

This leads to the definition of the "direct causal graph":

Definition 2.8.4 *Let $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ be an SCM that is structurally uniquely solvable. We define the* direct causal graph *as the directed graph $\mathcal{G}^{dc}(\mathcal{M})$ with nodes $\mathcal{I}$ and directed edges $i \to j$ if and only if i is a direct cause of j.*

The direct causal graph of such an SCM represents the direct causal relations that are described by the SCM. Note that latent confounders are not part of the direct causal graph, since the direct causal graph describes only the direct causal relations among the endogenous variables and not among the endogenous and exogenous variables together. See Section 3.4 for a discussion about latent confounders.

Interventionally equivalent SCMs describe the same set of interventional distributions and hence we have the following result:

Proposition 2.8.5 *Let $\mathcal{M}$ and $\tilde{\mathcal{M}}$ be interventionally equivalent. If both $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are structurally uniquely solvable, then $\mathcal{G}^{dc}(\mathcal{M}) = \mathcal{G}^{dc}(\tilde{\mathcal{M}})$.*

Proof. Suppose i is a direct cause of j in $\mathcal{M}$, then for any two different $\boldsymbol{\xi}_{\setminus j}$ and $\tilde{\boldsymbol{\xi}}_{\setminus j}$ in $\boldsymbol{\mathcal{X}}_{\setminus j}$ such that $\boldsymbol{\xi}_{\setminus \{i,j\}} = \tilde{\boldsymbol{\xi}}_{\setminus \{i,j\}}$ we have that $\mathcal{M}_{\mathrm{do}(\setminus j, \boldsymbol{\xi}_{\setminus j})}$ and $\mathcal{M}_{\mathrm{do}(\setminus j, \tilde{\boldsymbol{\xi}}_{\setminus j})}$ are not observationally equivalent. From the interventional equivalence of $\mathcal{M}$ and $\tilde{\mathcal{M}}$ it follows that $\tilde{\mathcal{M}}_{\mathrm{do}(\setminus j, \boldsymbol{\xi}_{\setminus j})}$ and $\tilde{\mathcal{M}}_{\mathrm{do}(\setminus j, \tilde{\boldsymbol{\xi}}_{\setminus j})}$ are also not observationally equivalent and hence i is a direct cause of j in $\tilde{\mathcal{M}}$. By symmetry it follows that $\mathcal{G}^{dc}(\mathcal{M}) = \mathcal{G}^{dc}(\tilde{\mathcal{M}})$. □

The direct causal graph is a subgraph of the (augmented) functional graph.

Proposition 2.8.6 *Let $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ be an SCM that is structurally uniquely solvable. Then $\mathcal{G}^{dc}(\mathcal{M}) \subseteq \mathcal{G}(\mathcal{M})$ and $\mathcal{G}^{dc}(\mathcal{M}) \subseteq \mathcal{G}^{a}(\mathcal{M})$.*

Proof. We need to prove that $i \in \mathrm{pa}(j)$ if i is a direct cause of j. Suppose $i \notin \mathrm{pa}(j)$, that is there exists a measurable mapping $\tilde{f}_j : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(j)\setminus\{i\}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(j)} \to \boldsymbol{\mathcal{X}}_j$ such that $\tilde{f}_j \equiv f_j$. Pick any different $\boldsymbol{\xi}_{\setminus j}$ and $\tilde{\boldsymbol{\xi}}_{\setminus j}$ in $\boldsymbol{\mathcal{X}}_{\setminus j}$ such that $\boldsymbol{\xi}_{\setminus \{i,j\}} = \tilde{\boldsymbol{\xi}}_{\setminus \{i,j\}}$. Then for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost all $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ the structural equation for the j^{th} variable of the intervened model $\mathcal{M}_{\mathrm{do}(\setminus j, \boldsymbol{\xi}_{\setminus j})}$ (or similar for $\mathcal{M}_{\mathrm{do}(\setminus j, \tilde{\boldsymbol{\xi}}_{\setminus j})}$) holds iff we have that

$$x_j = \tilde{f}_j(\boldsymbol{\xi}_{\mathrm{pa}(j)\setminus\{i,j\}}, x_j, \boldsymbol{e}_{\mathrm{pa}(j)})$$

holds. Without loss of generality, a solution X_j of $\mathcal{M}_{\mathrm{do}(\setminus j, \boldsymbol{\xi}_{\setminus j})}$ is also a solution of $\mathcal{M}_{\mathrm{do}(\setminus j, \tilde{\boldsymbol{\xi}}_{\setminus j})}$. Hence i cannot be a direct cause of j w.r.t. $\mathcal{I}$. □

The following example shows that $\mathcal{G}^{a}(\mathcal{M})_{\mathcal{I}} \subseteq \mathcal{G}^{dc}(\mathcal{M})$ does not hold in general:

Example 2.8.7 *Consider the SCM $\mathcal{M}$ of Example 2.7.8. There, x_1 is a parent of x_2, but x_1 is not a direct cause of x_2, since*

$$\mathbb{P}^{X_2}_{\mathcal{M}_{\mathrm{do}(\{1\},-1)}} = \mathbb{P}^{\tilde{X}_2}_{\mathcal{M}_{\mathrm{do}(\{1\},1)}}$$

for all solutions X_2 and $\tilde{X}_2$ of $\mathcal{M}_{\mathrm{do}(\{1\},-1)}$ and $\mathcal{M}_{\mathrm{do}(\{1\},1)}$ respectively. This is a consequence of the symmetric distribution of E_2. Changing the distribution of E_2 to any non-symmetric distribution will lead to a direct causal effect of x_1 on x_2.

2.9 Markov properties

In this section we will give a short overview of some Markov properties for SCMs, including cyclic SCMs, as derived in Forré and Mooij (2017). We show that under certain assumptions the general directed global Markov property, which associates a set of conditional independence relations to the functional graph of the SCM, holds for its observational distribution. This Markov property is defined in terms of σ-separation, which is a general extension of d-separation. In particular, σ-separation implies d-separation, and for the acyclic case, σ-separation is equivalent to d-separation. We show that under the assumption of unique solvability w.r.t. each strongly connected component, the observational distribution satisfies the general directed global Markov property relative to the functional graph. Moreover, we mention several special cases for which the stronger directed global Markov property holds. These assumptions are in general not preserved under intervention. We show that under the stronger assumption of unique solvability w.r.t. every subset, every interventional distribution satisfies the general directed global Markov property relative to the functional graph of the intervened model. For an extensive study of the different Markov properties that can be associated to SCMs we refer the reader to the work of Forré and Mooij (2017).

Definition 2.9.1 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{B})$ be a directed mixed graph and let $\mathcal{A}, \mathcal{B}, \mathcal{S} \subseteq \mathcal{V}$ be subsets of nodes. A path*[10] $(i = i_0, i_1, \dots, i_{n-1}, i_n = j)$ *in $\mathcal{G}$ is called* $\mathcal{S}$-σ-blocked or σ-blocked by $\mathcal{S}$ *if:*

1. *$i \in \mathcal{S}$ or $j \in \mathcal{S}$, or*
2. *there exists a node $i_k \notin \mathrm{an}_{\mathcal{G}}(\mathcal{S})$ with two adjacent edges that form a collider at i_k, i.e.*

$$i_{k-1} \to i_k \leftarrow i_{k+1}\,, \quad i_{k-1} \leftrightarrow i_k \leftarrow i_{k+1}\,,$$
$$i_{k-1} \to i_k \leftrightarrow i_{k+1}\,, \quad i_{k-1} \leftrightarrow i_k \leftrightarrow i_{k+1}\,,$$

   or

3. *there exists a node $i_k \in \mathcal{S}$ with two adjacent edges that form a non-collider at i_k, i.e.*

$$i_{k-1} \to i_k \to i_{k+1}\,, \quad i_{k-1} \leftrightarrow i_k \to i_{k+1}\,, \quad i_{k-1} \leftarrow i_k \to i_{k+1}\,,$$
$$i_{k-1} \leftarrow i_k \leftarrow i_{k+1}\,, \quad i_{k-1} \leftarrow i_k \leftrightarrow i_{k+1}\,,$$

   where at least one child (i_{k-1} or i_{k+1}) of i_k is not in $\mathrm{sc}_{\mathcal{G}}(i_k)$.

We call a path $\mathcal{S}$-σ-open *or* $\mathcal{S}$-σ-active *if it is not $\mathcal{S}$-σ-blocked. We say that $\mathcal{A}$ is* σ-separated from $\mathcal{B}$ given $\mathcal{S}$ *if every path $(i, \dots, j)$ in $\mathcal{G}$ with $i \in \mathcal{A}$ and $j \in \mathcal{B}$ is σ-blocked by $\mathcal{S}$, and write:*

$$\mathcal{A} \overset{\sigma}{\underset{\mathcal{G}}{\perp\!\!\!\perp}} \mathcal{B}\,.$$

[10] Alternatively, one may look at walks.

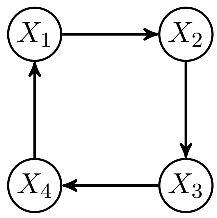

Figure 9: A directed graph $\mathcal{G}$ for which d-separation does not imply σ-separation (see Example 2.9.2).

In other words, a path in $\mathcal{G}$ is $\mathcal{S}$-σ-blocked iff it has an endnode in $\mathcal{S}$, or it contains a collider not in $\text{an}_{\mathcal{G}}(\mathcal{S})$, or it contains a non-collider in $\mathcal{S}$ that points to a node in a different strongly connected component. It is shown in Forré and Mooij (2017) that σ-separation implies d-separation. The other way around does not hold in general, as can be seen in the following example:

Example 2.9.2 *Consider the directed graph $\mathcal{G}$ as depicted in Figure 9. Here X_1 is d-separated from X_3 given $\{X_2, X_4\}$, but X_1 is not σ-separated from X_3 given $\{X_2, X_4\}$.*

In particular, σ-Separation is equivalent to d-separation for DAGs and ADMGs.

Definition 2.9.3 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{B})$ be a directed mixed graph and $\mathbb{P}_{\mathcal{V}}$ a probability distribution on $\boldsymbol{\mathcal{X}}_{\mathcal{V}} = \prod_{i\in\mathcal{V}} \mathcal{X}_i$. The probability distribution $\mathbb{P}_{\mathcal{V}}$ obeys the* general directed global Markov property *relative to $\mathcal{G}$, if for all subsets $\mathcal{A}, \mathcal{B}, \mathcal{S} \subseteq \mathcal{V}$ we have:*

$$\mathcal{A} \underset{\mathcal{G}}{\overset{\sigma}{\perp\!\!\!\perp}} \mathcal{B}|\mathcal{S} \quad \Longrightarrow \quad \mathcal{A} \underset{\mathbb{P}_{\mathcal{V}}}{\perp\!\!\!\perp} \mathcal{B}|\mathcal{S}$$

Here we have used the notation $\mathcal{A} \perp\!\!\!\perp_{\mathbb{P}_{\mathcal{V}}} \mathcal{B}|\mathcal{S}$ as short for $(X_i)_{i\in\mathcal{A}}$ and $(X_i)_{i\in\mathcal{B}}$ are conditionally independent given $(X_i)_{i\in\mathcal{S}}$ under $\mathbb{P}_{\mathcal{V}}$, where we take for $X_i : \boldsymbol{\mathcal{X}}_{\mathcal{V}} \to \mathcal{X}_i$ the canonical projections as random variables.

Theorem 2.9.4 *(Forré and Mooij, 2017) Consider an SCM $\mathcal{M}$ that is uniquely solvable w.r.t. each strongly connected component of $\mathcal{G}(\mathcal{M})$, then its observational distribution $\mathbb{P}^{\boldsymbol{X}}$ exists and is unique, and it obeys the general directed global Markov property relative to $\mathcal{G}(\mathcal{M})$.*

Proof. An SCM $\mathcal{M}$ that is uniquely solvable w.r.t. each strongly connected component is uniquely solvable and hence, by Lemma 2.6.19, all its solutions have the same observational distribution. The last statement follows from Theorem 3.8.2, 3.8.11, Lemma 3.7.7 and Remark 3.7.2 in Forré and Mooij (2017). □

The fact that σ-separation implies d-separation means that the directed global Markov property implies the general directed global Markov property. The other way around does not hold in general. Under stronger conditions we do have the following results:

Theorem 2.9.5 *(Forré and Mooij, 2017) Consider an SCM $\mathcal{M}$ that is uniquely solvable w.r.t. each strongly connected component of $\mathcal{G}(\mathcal{M})$ and moreover satisfies at least one of the following three conditions:*

1. *$\mathcal{M}$ is linear and $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$ has a density w.r.t. the Lebesgue measure;*
2. *$\mathcal{M}$ has discrete-valued endogenous variables;*
3. *$\mathcal{M}$ is acyclic;*

then its observational distribution $\mathbb{P}^{\boldsymbol{X}}$ obeys the directed global Markov property relative to $\mathcal{G}(\mathcal{M})$.

Proof. The linear case is proven by Theorem 3.8.17, the discrete case is proven by Theorem 3.8.2, 3.8.12, Remark 3.7.2, Theorem 3.6.6 and 3.5.2 and the acyclic case is proven by Remark 3.3.4 in Forré and Mooij (2017). □

For the linear case, one can relax the condition of unique solvability w.r.t. each strongly connected component to the condition of unique solvability w.r.t. $\mathcal{I}$. For the discrete case one can relax the condition of unique solvability w.r.t. each strongly connected component to the condition of unique solvability w.r.t. each ancestral subgraph of $\mathcal{G}(\mathcal{M})$ (Forré and Mooij, 2017).

The results in Theorems 2.9.4 and 2.9.5 are not preserved under intervention, because intervening on a strongly connected component could split it into several strongly connected components. The stronger condition that the SCM is uniquely solvable w.r.t. each strongly connected induced subgraph is preserved under intervention.

Definition 2.9.6 *A* loop *in a directed mixed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{B})$ is a subset $\mathcal{O} \subseteq \mathcal{V}$ that is strongly connected in the induced subgraph $\mathcal{G}_{\mathcal{O}}$ of $\mathcal{G}$ on $\mathcal{O}$.*

A cycle is a directed path $i \to \cdots \to j$ plus an edge $j \to i$, a loop can have many more edges between the nodes. For an SCM $\mathcal{M}$ we denote the set of loops of the functional graph $\mathcal{G}(\mathcal{M})$ by $L(\mathcal{M})$.

Proposition 2.9.7 *An SCM $\mathcal{M}$ is uniquely solvable w.r.t. every subset in $L(\mathcal{M})$ iff it is uniquely solvable w.r.t. every subset of $\mathcal{I}$.*

Proof. Suppose $\mathcal{M}$ is uniquely solvable w.r.t. every subset in $L(\mathcal{M})$ and consider a subset $\mathcal{O} \subseteq \mathcal{I}$. Consider the induced subgraph $\mathcal{G}^a(\mathcal{M})_{\mathcal{O}}$ of $\mathcal{G}^a(\mathcal{M})$ on the nodes $\mathcal{O}$. Then every strongly connected component of $\mathcal{G}^a(\mathcal{M})_{\mathcal{O}}$ is an element of $L(\mathcal{M})$. Let $\mathcal{C}$ be a strongly connected component in $\mathcal{G}^a(\mathcal{M})_{\mathcal{O}}$, and $\boldsymbol{g}_{\mathcal{C}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{C})\setminus\mathcal{C}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{C})} \to \boldsymbol{\mathcal{X}}_{\mathcal{C}}$ the mapping that makes $\mathcal{M}$ uniquely solvable w.r.t. $\mathcal{C}$. Since $\mathcal{G}^a(\mathcal{M})_{\mathcal{O}}$ partitions into strongly connected components, we can recursively (by following a topological ordening) insert these mappings into each other to obtain a mapping $\boldsymbol{g}_{\mathcal{O}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{O})\setminus\mathcal{O}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{O})} \to \boldsymbol{\mathcal{X}}_{\mathcal{O}}$ that makes $\mathcal{M}$ uniquely solvable w.r.t. $\mathcal{O}$. □

Note that we already proved in Proposition 2.6.27 that the property that an SCM is uniquely solvable w.r.t. every subset is preserved under intervention.

Corollary 2.9.8 *Consider an SCM $\mathcal{M}$ that is uniquely solvable w.r.t. every subset of $\mathcal{I}$. Then, for every $I \subseteq \mathcal{I}$ and $\boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$ the interventional distribution $\mathbb{P}^{\boldsymbol{X}}$ of $\mathcal{M}_{\mathrm{do}(I,\boldsymbol{\xi}_I)}$ obeys the general directed global Markov property relative to $\mathcal{G}(\mathcal{M}_{\mathrm{do}(I,\boldsymbol{\xi}_I)})$.*

Proof. This follows from Proposition 2.9.7 (or Proposition 2.6.27) and Theorem 2.9.4. □

Under this assumption that the SCM is uniquely solvable w.r.t. every subset, we know that the direct causal graph is defined under all interventions. Moreover, this assumption is, as we will see in the next section, also closed under marginalization and hence the direct causal graph is also defined under any marginalization. Therefore, such SCMs form a particularly convenient subclass of SCMs that include cycles.

3 Marginalizations

In this section we will show how we can marginalize an SCM over a subset $\mathcal{L} \subset \mathcal{I}$ of endogenous variables to another SCM on the margin $\mathcal{I} \setminus \mathcal{L}$ such that it preserves the causal semantics on this margin. In particular, this could be seen as marginalizing all observational and interventional distributions associated to the SCM down to their corresponding distributions on the margin.

Intuitively, the idea is that we would like to treat the subsystem $\mathcal{L}$ as a "black box", and only describe how the rest of the system interacts with it. Thereby we completely remove the representation of the internals of the subsystem, preserving only its essential input-output characteristics. We will show that if part of the causal mechanism that describes the causal relations in the subsystem $\mathcal{L}$ (i.e., the restriction of the causal mechanism to $\mathcal{L}$) satisfies the unique solvability w.r.t. $\mathcal{L}$ condition (intuitively: it gives a unique output for any possible input), then we can effectively remove this subsystem of endogenous variables from the model, treating it as a black box. An important property of this marginalization operation is that it preserves the causal semantics, meaning that the interventional distributions induced by the SCM are on the margin identical to those induced by its marginalization. Moreover, we show that this marginalization operation preserves the counterfactual semantics.

Similarly to the marginalization of an SCM we will define the marginalization of a directed graph, which is called the latent projection. We show that in general the marginalization of an SCM does not obey the latent projection of its associated augmented functional graph, i.e. that the augmented functional graph of the marginal SCM is not always a subgraph of the latent projection of the augmented functional graph of the original SCM. However, as we will see, this does hold given a sufficient condition. This leads to the result that SCMs that are uniquely solvable w.r.t. every subset are closed under intervention and marginalization, for which the acyclic SCMs form a particular subclass.

Next, we will define the direct causal graph w.r.t. a context set, which leads to the notion of indirect cause. We will show that in general the marginalization of an SCM does not obey the latent projection of its corresponding direct causal graph. In the end, we will discuss latent confounders, i.e. latent common causes.

3.1 Marginal structural causal models

Before we show how we can marginalize an SCM over a subset of endogenous variables, we will first like to mention that in general it is not always possible to find an SCM on the margin that preserves the causal semantics, i.e. that they are interventionally equivalent on the margin, as the following example ilustrates.

Example 3.1.1 *Consider the SCM $\mathcal{M} = \langle \mathbf{3}, \emptyset, \mathbb{R}^3, 1, \boldsymbol{f}, \mathbb{P}_1 \rangle$ with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}) &= x_1 + x_2 + x_3 \\ f_2(\boldsymbol{x}) &= x_2 \\ f_3(\boldsymbol{x}) &= 0\,. \end{aligned}$$

Then there exists no SCM $\tilde{\mathcal{M}}$ on the variables $\{2,3\}$ that is interventionally equivalent to $\mathcal{M}$ w.r.t. $\{2,3\}$. To see this, suppose there exists such an $\tilde{\mathcal{M}}$, then for any $\xi_2, \xi_3 \in \boldsymbol{\mathcal{X}}_{\{2,3\}}$ such that $\xi_2 + \xi_3 \neq 0$ the intervened model $\tilde{\mathcal{M}}_{\mathrm{do}(\{2,3\},(\xi_2,\xi_3))}$ has a solution but $\mathcal{M}_{\mathrm{do}(\{2,3\},(\xi_2,\xi_3))}$ does not.

More generally, for an SCM $\mathcal{M}$ that is not solvable w.r.t. a subset $\mathcal{L} \subset \mathcal{I}$ one can never find an SCM $\tilde{\mathcal{M}}$ on the endogenous variables $\mathcal{I} \setminus \mathcal{L}$ that is interventionally equivalent w.r.t. $\mathcal{I} \setminus \mathcal{L}$.

One may wonder, if instead for an SCM $\mathcal{M}$ that is solvable w.r.t. the subset $\mathcal{L}$, one can find an SCM on the margin $\mathcal{I} \setminus \mathcal{L}$ such that it is interventionally equivalent w.r.t. $\mathcal{I} \setminus \mathcal{L}$. The next example shows that this is in general not possible either:

Example 3.1.2 *Consider the SCM $\mathcal{M} = \langle \mathbf{3}, \mathbf{1}, \mathbb{R}^3, \mathbb{R}, \boldsymbol{f}, \mathbb{P}_{\mathbb{R}} \rangle$ with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}) &= e_1 \\ f_2(\boldsymbol{x}) &= x_2 - x_2^2 + x_1^2 \\ f_3(\boldsymbol{x}) &= x_2\,. \end{aligned}$$

and $\mathbb{P}_{\mathbb{R}}$ the standard-normal measure on $\mathbb{R}$. This SCM is solvable w.r.t. $\{2\}$. Then for an intervention $\mathrm{do}(\{1\},\xi_1)$, with $\xi_1 \in \mathbb{R}$, on $\mathcal{M}$, the solution on x_3 is either equal to $-|\xi_1|$ or to $|\xi_1|$. There exist no SCM $\tilde{\mathcal{M}}$ on $\{1,3\}$ such that $\mathcal{M}_{\mathrm{do}(\{1\},\xi)}$ and $\tilde{\mathcal{M}}_{\mathrm{do}(\{1\},\xi)}$ are observationally equivalent w.r.t. $\{1,3\}$, since the causal mechanism of $\tilde{f}_3$ of $\tilde{\mathcal{M}}_{\mathrm{do}(\{1\},\xi)}$ could never map x_1 to more than one value.

As we will show next, the stronger condition of unique solvability w.r.t. a subset, is a sufficient condition for the existence of an SCM on the margin that preserves the causal semantics. Consider the following example:

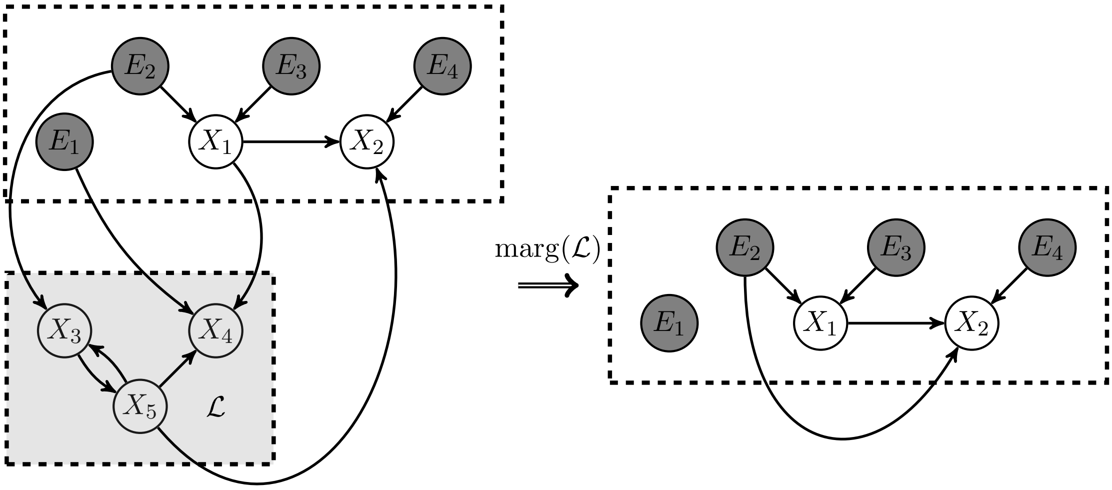

Figure 10: The augmented functional graphs of the SCM $\mathcal{M}$ (left) and $\tilde{\mathcal{M}}$ (right) of Example 3.1.3, where the SCM $\tilde{\mathcal{M}}$ defines a marginalization of $\mathcal{M}$ w.r.t. $\mathcal{L}$.

Example 3.1.3 *Consider the SCM* $\mathcal{M} = \langle \mathbf{5}, \mathbf{4}, \mathbb{R}^5, \mathbb{R}^4, \boldsymbol{f}, \mathbb{P}_{\mathbb{R}^4} \rangle$ *with the following causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}, \boldsymbol{e}) &= e_2 + e_3 \\ f_2(\boldsymbol{x}, \boldsymbol{e}) &= x_1 + x_5 + e_4 \\ f_3(\boldsymbol{x}, \boldsymbol{e}) &= x_5 + e_2 \\ f_4(\boldsymbol{x}, \boldsymbol{e}) &= x_1 + x_5 + e_1 \\ f_5(\boldsymbol{x}, \boldsymbol{e}) &= \tfrac{1}{2} x_3 \end{aligned}$$

and take for $\mathbb{P}_{\mathbb{R}^4}$ *the standard-normal measure on* $\mathbb{R}^4$*. This SCM, depicted in Figure 10, is uniquely solvable w.r.t.* $\mathcal{L} := \{3, 4, 5\}$ *by the mapping* $\boldsymbol{g}_{\{3,4,5\}} : \mathcal{X}_1 \times \boldsymbol{\mathcal{E}}_{\{1,2\}} \to \boldsymbol{\mathcal{X}}_{\{3,4,5\}}$ *given by:*

$$\begin{aligned} g_3(\boldsymbol{x}, \boldsymbol{e}) &= 2e_2 \\ g_4(\boldsymbol{x}, \boldsymbol{e}) &= x_1 + e_1 + e_2 \\ g_5(\boldsymbol{x}, \boldsymbol{e}) &= e_2 \end{aligned}$$

The structural equations for the variables $\mathcal{L}$ *can be seen as a subsystem, that is for every input* $(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L})\setminus\mathcal{L}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}) \in \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{L})\setminus\mathcal{L}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{L})}$ *these mappings give rise to a unique output* $\boldsymbol{x}_{\mathcal{L}} \in \boldsymbol{\mathcal{X}}_{\mathcal{L}}$*. Substituting these mappings into the causal mechanism of the variables* $\{1, 2\}$ *gives a "marginal" causal mechanism:*

$$\begin{aligned} \tilde{f}_1(\boldsymbol{x}, \boldsymbol{e}) &:= e_2 + e_3 \\ \tilde{f}_2(\boldsymbol{x}, \boldsymbol{e}) &:= x_1 + e_2 + e_4 \,. \end{aligned}$$

These mappings define an SCM $\tilde{\mathcal{M}} := \langle \mathbf{2}, \mathbf{4}, \mathbb{R}^2, \mathbb{R}^4, \tilde{\boldsymbol{f}}, \mathbb{P}_{\mathbb{R}^4} \rangle$ *on the margin. This constructed SCM* $\tilde{\mathcal{M}}$ *is interventionally equivalent w.r.t.* $\mathcal{L}$*, which can be checked manually or by applying Theorem 3.1.10.*

In general, for an SCM $\mathcal{M}$ and a given subset $\mathcal{L} \subset \mathcal{I}$ of endogenous variables and its complement $\mathcal{O} := \mathcal{I} \setminus \mathcal{L}$, the structural equations $\boldsymbol{x}_\mathcal{L} = \boldsymbol{f}_\mathcal{L}(\boldsymbol{x}_\mathcal{L}, \boldsymbol{x}_\mathcal{O}, \boldsymbol{e})$ define a "subsystem". In Example 3.1.3 this is depicted as a gray box in Figure 10. If $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}$ by some mapping $\boldsymbol{g}_\mathcal{L} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{L})\setminus\mathcal{L}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{L})} \to \boldsymbol{\mathcal{X}}_\mathcal{L}$, then for each input $(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L})\setminus\mathcal{L}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}) \in \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{L})\setminus\mathcal{L}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{L})}$ of the subsystem, there exists a $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{L})}}$-unique[11] output $\boldsymbol{x}_\mathcal{L} \in \boldsymbol{\mathcal{X}}_\mathcal{L}$. We can effectively remove this subystem of endogenous variables from the model by substitution. This will lead to a marginal SCM that is interventionally equivalent to the original SCM w.r.t. the margin, which we will prove in Theorem 3.1.10.

Definition 3.1.4 *Consider an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *and a subset* $\mathcal{L} \subset \mathcal{I}$*, and let* $\mathcal{O} = \mathcal{I} \setminus \mathcal{L}$*. If* $\mathcal{M}$ *is uniquely solvable w.r.t.* $\mathcal{L}$*, then we define a* marginalization of $\mathcal{M}$ w.r.t. $\mathcal{L}$*, denoted by* $\mathrm{marg}(\mathcal{L})$*, as the SCM* $\mathcal{M}_{\mathrm{marg}(\mathcal{L})} := \langle \mathcal{O}, \mathcal{J}, \boldsymbol{\mathcal{X}}_\mathcal{O}, \boldsymbol{\mathcal{E}}, \tilde{\boldsymbol{f}}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *with "marginal" causal mechanism* $\tilde{\boldsymbol{f}} : \boldsymbol{\mathcal{X}}_\mathcal{O} \times \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{X}}_\mathcal{O}$ *defined by*

$$\tilde{\boldsymbol{f}}(\boldsymbol{x}_\mathcal{O}, \boldsymbol{e}) := \boldsymbol{f}_\mathcal{O}(\boldsymbol{g}_\mathcal{L}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L})\setminus\mathcal{L}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}), \boldsymbol{x}_\mathcal{O}, \boldsymbol{e}) \, .$$

where $\boldsymbol{g}_\mathcal{L} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{L})\setminus\mathcal{L}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{L})} \to \boldsymbol{\mathcal{L}}$ *is a mapping that makes* $\mathcal{M}$ *uniquely solvable w.r.t.* $\mathcal{L}$*.*

Note that for a specific $\mathcal{L} \subset \mathcal{I}$ there may exist more than one marginalization $\mathrm{marg}(\mathcal{L})(\mathcal{M})$, depending on the choice of the mapping $\boldsymbol{g}_\mathcal{L}$ that makes $\mathcal{M}$ uniquely solvable w.r.t. $\mathcal{L}$. However, all marginalizations map $\mathcal{M}$ to a representative of the same equivalence class of SCMs. Moreover, marginalizing two equivalent SCMs over $\mathcal{L}$ yields two equivalent marginal SCMs. Thus, the relation $\mathrm{marg}(\mathcal{L})$ between SCMs induces a mapping between the equivalences classes of SCMs.

With this definition at hand, we recover the known result that we can always construct a marginal SCM over a subset of endogenous variables from an acyclic SCM by merely substitution (see Proposition 2.6.18). Moreover, it extends to cyclic SCMs, namely to those that are uniquely solvable w.r.t. a certain subset, as we already saw in Example 3.1.3. We will prove in Theorem 3.1.10 the important property that for those SCMs, the SCM and its corresponding marginal SCM are interventionally equivalent w.r.t. the margin. We would like to stress that for an SCM $\mathcal{M}$, unique solvability w.r.t. a certain subset $\mathcal{L} \subseteq \mathcal{I}$, is a sufficient, but not a necessary condition, for the existence of an SCM $\tilde{\mathcal{M}}$ on the margin $\mathcal{I} \setminus \mathcal{L}$ such that $\mathcal{M}$ and $\tilde{\mathcal{M}}$ are interventionally equivalent w.r.t. $\mathcal{I} \setminus \mathcal{L}$. This is illustrated by the following example:

Example 3.1.5 *Consider the SCM* $\mathcal{M} = \langle \mathbf{4}, \mathbf{1}, \mathbb{R}^4, \mathbb{R}, \boldsymbol{f}, \mathbb{P}_\mathbb{R} \rangle$ *with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}, e) &= e \\ f_2(\boldsymbol{x}, e) &= x_1 \\ f_3(\boldsymbol{x}, e) &= x_2 \\ f_4(\boldsymbol{x}, e) &= x_4 \, . \end{aligned}$$

[11] A mapping is $\mathbb{P}$-unique if it is unique up to $\mathbb{P}$-null set.

and $\mathbb{P}_{\mathbb{R}}$ the standard-normal measure on $\mathbb{R}$. This SCM is solvable w.r.t. $\mathcal{L} = \{2, 4\}$, but not uniquely solvable w.r.t. $\mathcal{L}$, and hence we cannot apply Definition 3.1.4 to $\mathcal{L}$. However, the SCM $\tilde{\mathcal{M}}$ on the endogenous variables $\{1, 3\}$ with the causal mechanism $\tilde{\boldsymbol{f}}$ given by:

$$\begin{aligned}\tilde{f}_1(\boldsymbol{x}, e) &= e\\ \tilde{f}_3(\boldsymbol{x}, e) &= x_1 \,.\end{aligned}$$

is interventionally equivalent w.r.t. $\{1, 3\}$, which can be checked manually.

Hence, in certain cases it is possible to relax the uniqueness condition. Classifying all these cases is beyond the scope of this paper.

The definition of marginalization is well-defined, in the sense that, if we can marginalize over two disjoint subsets after each other, then we can also marginalize over the union of those subsets at once.

Lemma 3.1.6 *Given an SCM $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ and two disjoint subsets $\mathcal{L}_1, \mathcal{L}_2 \subset \mathcal{I}$. Assume $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}_1$ and $\mathcal{M}_{\mathrm{marg}(\mathcal{L}_1)}$ is uniquely solvable w.r.t. $\mathcal{L}_2$, then $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}_1 \cup \mathcal{L}_2$. Moreover* $\mathrm{marg}(\mathcal{L}_2) \circ \mathrm{marg}(\mathcal{L}_1)(\mathcal{M}) \equiv \mathrm{marg}(\mathcal{L}_1 \cup \mathcal{L}_2)(\mathcal{M})$.

Proof. From unique solvability of $\mathcal{M}$ w.r.t. $\mathcal{L}_1$ it follows that there exists a mapping $\boldsymbol{g}_{\mathcal{L}_1}$ such that for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$\boldsymbol{x}_{\mathcal{L}_1} = \boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}) \quad \Longleftrightarrow \quad \boldsymbol{x}_{\mathcal{L}_1} = \boldsymbol{f}_{\mathcal{L}_1}(\boldsymbol{x}, \boldsymbol{e}) \,.$$

If $\mathcal{M}_{\mathrm{marg}(\mathcal{L}_1)}$ is uniquely solvable w.r.t. $\mathcal{L}_2$, then there exists a mapping $\boldsymbol{g}_{\mathcal{L}_2}$ such that for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x}_{\mathcal{I}\setminus\mathcal{L}_1} \in \boldsymbol{\mathcal{X}}_{\mathcal{I}\setminus\mathcal{L}_1}$:

$$\begin{aligned}&\boldsymbol{x}_{\mathcal{L}_2} = \boldsymbol{g}_{\mathcal{L}_2}(\boldsymbol{x}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)\setminus\mathcal{L}_2}, \boldsymbol{e}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)})\\ \Longleftrightarrow\ &\boldsymbol{x}_{\mathcal{L}_2} = \boldsymbol{f}_{\mathcal{L}_2}(\boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}), \boldsymbol{x}_{\mathcal{I}\setminus\mathcal{L}_1}, \boldsymbol{e}) \,,\end{aligned}$$

where $\widehat{\mathrm{pa}}(\mathcal{L}_2)$ denotes the parents of $\mathcal{L}_2$ w.r.t. the graph $\mathcal{G}^a(\mathcal{M}_{\mathrm{marg}(\mathcal{L}_1)})$. Note that for the endogenous parents $\widehat{\mathrm{pa}}(\mathcal{L}_2)\setminus\mathcal{L}_2 \subseteq \mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2)\setminus(\mathcal{L}_1 \cup \mathcal{L}_2)$ and for the exogenous parents $\widehat{\mathrm{pa}}(\mathcal{L}_2) \subseteq \mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2)$. Take the Ansatz for the mapping $\tilde{\boldsymbol{g}}_{\mathcal{L}_1\cup\mathcal{L}_2} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{L}_1\cup\mathcal{L}_2)\setminus(\mathcal{L}_1\cup\mathcal{L}_2)} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{L}_1\cup\mathcal{L}_2)} \to \boldsymbol{\mathcal{X}}_{\mathcal{L}_1\cup\mathcal{L}_2}$:

$$\begin{aligned}&(\tilde{\boldsymbol{g}}_{\mathcal{L}_1}, \tilde{\boldsymbol{g}}_{\mathcal{L}_2})(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1\cup\mathcal{L}_2)\setminus(\mathcal{L}_1\cup\mathcal{L}_2)}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1\cup\mathcal{L}_2)}) :=\\ &(\boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{g}_{\mathrm{pa}(\mathcal{L}_1)\cap\mathcal{L}_2}(\boldsymbol{x}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)\setminus\mathcal{L}_2}, \boldsymbol{e}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)}), \boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus(\mathcal{L}_1\cup\mathcal{L}_2)}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}), \boldsymbol{g}_{\mathcal{L}_2}(\boldsymbol{x}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)\setminus\mathcal{L}_2}, \boldsymbol{e}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)})) \,.\end{aligned}$$

Then for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$
\begin{aligned}
&\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{f}_{\mathcal{L}_1}(\boldsymbol{x}, \boldsymbol{e}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{f}_{\mathcal{L}_2}(\boldsymbol{x}, \boldsymbol{e}) \end{cases} \\
\iff &\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{f}_{\mathcal{L}_2}(\boldsymbol{x}, \boldsymbol{e}) \end{cases} \\
\iff &\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{f}_{\mathcal{L}_2}(\boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}), \boldsymbol{x}_{\mathcal{I}\setminus\mathcal{L}_1}, \boldsymbol{e}) \end{cases} \\
\iff &\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{g}_{\mathcal{L}_2}(\boldsymbol{x}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)\setminus\mathcal{L}_2}, \boldsymbol{e}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)}) \end{cases} \\
\iff &\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{g}_{\mathrm{pa}(\mathcal{L}_1)\cap\mathcal{L}_2}(\boldsymbol{x}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)\setminus\mathcal{L}_2}, \boldsymbol{e}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)}), \boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus(\mathcal{L}_1\cup\mathcal{L}_2)}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{g}_{\mathcal{L}_2}(\boldsymbol{x}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)\setminus\mathcal{L}_2}, \boldsymbol{e}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)}) \end{cases} \\
\iff &\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \tilde{\boldsymbol{g}}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1\cup\mathcal{L}_2)\setminus(\mathcal{L}_1\cup\mathcal{L}_2)}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1\cup\mathcal{L}_2)}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \tilde{\boldsymbol{g}}_{\mathcal{L}_2}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1\cup\mathcal{L}_2)\setminus(\mathcal{L}_1\cup\mathcal{L}_2)}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1\cup\mathcal{L}_2)}) , \end{cases}
\end{aligned}
$$

where in the first equivalence we used solvability w.r.t. $\mathcal{L}_1$ of $\mathcal{M}$, in the second we used substitution, in the third we used solvability w.r.t. $\mathcal{L}_2$ of $\mathcal{M}_{\mathrm{marg}(\mathcal{L}_1)}$, in the fourth we used again substitution and in the last equivalence we used the Ansatz. From this we conclude that $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}_1 \cup \mathcal{L}_2$. Hence, by definition it follows that $\mathrm{marg}(\mathcal{L}_2) \circ \mathrm{marg}(\mathcal{L}_1)(\mathcal{M}) \equiv \mathrm{marg}(\mathcal{L}_1 \cup \mathcal{L}_2)(\mathcal{M})$. □

Note that in the previous lemma $\mathcal{L}_1$ and $\mathcal{L}_2$ have to be disjoint, since marginalizing first over $\mathcal{L}_1$ gives a maginal SCM $\mathcal{M}_{\mathrm{marg}(\mathcal{L}_1)}$ with endogenous variables $\mathcal{I} \setminus \mathcal{L}_1$.

This leads to the following commutativity result for marginalization:

Proposition 3.1.7 *Given an SCM $\mathcal{M}$ and disjoint subsets $\mathcal{L}_1, \mathcal{L}_2 \subset \mathcal{I}$. If $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}_1$ and $\mathcal{L}_2$, $\mathcal{M}_{\mathrm{marg}(\mathcal{L}_1)}$ is uniquely solvable w.r.t. $\mathcal{L}_2$ and $\mathcal{M}_{\mathrm{marg}(\mathcal{L}_2)}$ is uniquely solvable w.r.t. $\mathcal{L}_1$, then marginalizing subsequently over $\mathcal{L}_1$ and $\mathcal{L}_2$ commutes, i.e.* $\mathrm{marg}(\mathcal{L}_2) \circ \mathrm{marg}(\mathcal{L}_1)(\mathcal{M}) = \mathrm{marg}(\mathcal{L}_1) \circ \mathrm{marg}(\mathcal{L}_2)(\mathcal{M}) = \mathrm{marg}(\mathcal{L}_1 \cup \mathcal{L}_2)(\mathcal{M})$.

Proof. This follows from Lemma 3.1.6. □

In general, marginalization is not always defined for all subsets as can be seen from Example 2.6.21. There we cannot marginalize $\mathcal{M}$ over the variable x_2, but we can marginalize it over the variables x_1 and x_2 together. The fact that we cannot marginalize over the single variable x_2 over there is due to the existence of a self-loop at x_2. It follows from Proposition 2.6.24 that we can only marginalize over a single variable if that variable has no self-loop. Note that we may introduce new self-loops if we marginalize over a subset of variables, as can be seen, for example, from the last two SCMs in Example 2.6.1. These SCMs have no self-loops, however marginalizing over x_2 yield a marginal SCM with a self-loop at the variable x_1.

Before we show the important property that marginalization preserves the causal semantics, we first show, as an intermediate result, that the marginalization operation does not change the solution space of the observed variables:

Lemma 3.1.8 *Given an SCM $\mathcal{M}$ and a subset $\mathcal{L} \subset \mathcal{I}$ such that $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}$. Then $\mathcal{M}$ and* $\mathrm{marg}(\mathcal{L})(\mathcal{M})$ *are observationally equivalent w.r.t. $\mathcal{I} \setminus \mathcal{L}$.*

Proof. Let $(\boldsymbol{X}, \boldsymbol{E})$ be random variables such that $\mathbb{P}^{\boldsymbol{E}} = \mathbb{P}_{\boldsymbol{\mathcal{E}}}$ and $\boldsymbol{X} = \boldsymbol{f}(\boldsymbol{X}, \boldsymbol{E})$ holds a.s.. By unique solvability w.r.t. $\mathcal{L}$ it follows that for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$\begin{cases} \boldsymbol{x}_{\mathcal{L}} &= \boldsymbol{f}_{\mathcal{L}}(\boldsymbol{x}, \boldsymbol{e}) \\ \boldsymbol{x}_{\mathcal{O}} &= \boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}, \boldsymbol{e}) \end{cases}$$

$$\iff \begin{cases} \boldsymbol{x}_{\mathcal{L}} &= \boldsymbol{g}_{\mathcal{L}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}) \setminus \mathcal{L}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}) \\ \boldsymbol{x}_{\mathcal{O}} &= \boldsymbol{f}_{\mathcal{O}}(\boldsymbol{g}_{\mathcal{L}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}) \setminus \mathcal{L}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}), \boldsymbol{x}_{\mathcal{O}}, \boldsymbol{e}) \end{cases}$$

$$\iff \begin{cases} \boldsymbol{x}_{\mathcal{L}} &= \boldsymbol{g}_{\mathcal{L}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}) \setminus \mathcal{L}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}) \\ \boldsymbol{x}_{\mathcal{O}} &= \tilde{\boldsymbol{f}}_{\mathcal{O}}(\boldsymbol{x}_{\mathcal{O}}, \boldsymbol{e}), \end{cases}$$

where $\mathcal{O} = \mathcal{I} \setminus \mathcal{L}$. Thus $\boldsymbol{X}_{\mathcal{O}} = \tilde{\boldsymbol{f}}(\boldsymbol{X}_{\mathcal{O}}, \boldsymbol{E})$ holds a.s.. Conversely, suppose $(\boldsymbol{X}_{\mathcal{O}}, \boldsymbol{E})$ are random variables such that $\mathbb{P}^{\boldsymbol{E}} = \mathbb{P}_{\boldsymbol{\mathcal{E}}}$ and $\boldsymbol{X}_{\mathcal{O}} = \tilde{\boldsymbol{f}}(\boldsymbol{X}_{\mathcal{O}}, \boldsymbol{E})$ holds a.s., where $\tilde{\boldsymbol{f}}$ is the marginal causal mechanism of $\mathrm{marg}(\mathcal{L})(\mathcal{M})$. Let $\boldsymbol{g}_{\mathcal{L}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{L}) \setminus \mathcal{L}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{L})} \to \boldsymbol{\mathcal{X}}_{\mathcal{L}}$ be a mapping that makes $\mathcal{M}$ uniquely solvable w.r.t. $\mathcal{L}$, then the random variable $\boldsymbol{X}_{\mathcal{L}} := \boldsymbol{g}_{\mathcal{L}}(\boldsymbol{X}_{\mathrm{pa}(\mathcal{L}) \setminus \mathcal{L}}, \boldsymbol{E}_{\mathrm{pa}(\mathcal{L})})$ makes $\boldsymbol{X} := (\boldsymbol{X}_{\mathcal{O}}, \boldsymbol{X}_{\mathcal{L}})$ a solution of $\mathcal{M}$. □

Next we show that the interventional distributions of a marginal SCM are identical to the marginal interventional distributions induced by the original SCM. A simple proof of this result proceeds by showing that the operations of intervening and marginalizing commute.

Lemma 3.1.9 *Given an SCM $\mathcal{M}$, a subset $\mathcal{L} \subset \mathcal{I}$ such that $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}$, a subset $I \subseteq \mathcal{I} \setminus \mathcal{L}$ and a value $\boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$. Marginalization* $\mathrm{marg}(\mathcal{L})$ *commutes with perfect intervention* $\mathrm{do}(I, \boldsymbol{\xi}_I)$*, i.e.* $(\mathrm{marg}(\mathcal{L}) \circ \mathrm{do}(I, \boldsymbol{\xi}_I))(\mathcal{M}) \equiv (\mathrm{do}(I, \boldsymbol{\xi}) \circ \mathrm{marg}(\mathcal{L}))(\mathcal{M})$.

Proof. This follows straightforwardly from the definitions of perfect intervention and marginalization and the fact that if $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}$, then $\mathcal{M}_{\mathrm{do}(I, \boldsymbol{\xi}_I)}$ is also uniquely solvable w.r.t. $\mathcal{L}$, since the structural equations for $\mathcal{L}$ are the same for $\mathcal{M}$ and $\mathcal{M}_{\mathrm{do}(I, \boldsymbol{\xi}_I)}$. □

With Lemmas 3.1.8 and 3.1.9 at hand we can prove the main theorem:

Theorem 3.1.10 *Given an SCM $\mathcal{M}$ and a subset $\mathcal{L} \subset \mathcal{I}$ such that $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}$. Then $\mathcal{M}$ and* $\mathrm{marg}(\mathcal{L})(\mathcal{M})$ *are interventionally equivalent w.r.t. $\mathcal{I} \setminus \mathcal{L}$.*

Proof. From Lemma 3.1.9 we know that for a subset $I \subseteq \mathcal{I} \setminus \mathcal{L}$ and a value $\boldsymbol{\xi}_I \in \boldsymbol{\mathcal{X}}_I$, $(\text{marg}(\mathcal{L}) \circ \text{do}(I, \boldsymbol{\xi}_I))(\mathcal{M})$ exists. By Lemma 3.1.8 we know that $\text{do}(I, \boldsymbol{\xi}_I)(\mathcal{M})$ and $(\text{marg}(\mathcal{L}) \circ \text{do}(I, \boldsymbol{\xi}_I))(\mathcal{M})$ are observationally equivalent w.r.t. $\mathcal{O}$ and hence by applying again Lemma 3.1.9, $\text{do}(I, \boldsymbol{\xi}_I)(\mathcal{M})$ and $(\text{do}(I, \boldsymbol{\xi}) \circ \text{marg}(\mathcal{L}))(\mathcal{M})$ are observationally equivalent w.r.t. $\mathcal{O}$. This implies that $\mathcal{M}$ and $\text{marg}(\mathcal{L})(\mathcal{M})$ are interventionally equivalent. □

This shows that our definition of a marginal SCM is indeed an SCM that is interventionally equivalent w.r.t. the margin. As we saw in Example 2.7.13 it is generally not true that interventional equivalence implies counterfactual equivalence. However, for our definition of a marginal SCM we have, in addition, the following interesting corollary:

Corollary 3.1.11 *Given an SCM $\mathcal{M}$ and a subset $\mathcal{L} \subset \mathcal{I}$ such that $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}$. Then $\mathcal{M}$ and* $\text{marg}(\mathcal{L})(\mathcal{M})$ *are counterfactually equivalent w.r.t. $\mathcal{I} \setminus \mathcal{L}$.*

Proof. We need to show that $\mathcal{M}^{\text{twin}}$ and $(\mathcal{M}_{\text{marg}(\mathcal{L})})^{\text{twin}}$ are interventionally equivalent w.r.t. $(\mathcal{I} \cup \mathcal{I}') \setminus (\mathcal{L} \cup \mathcal{L}')$, where $\mathcal{L}'$ is a copy of $\mathcal{L}$ in $\mathcal{I}'$. If $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}$, then $\mathcal{M}^{\text{twin}}$ is uniquely solvable w.r.t. $\mathcal{L} \cup \mathcal{L}'$. Moreover, $(\mathcal{M}_{\text{marg}(\mathcal{L})})^{\text{twin}}$ is equivalent to $(\mathcal{M}^{\text{twin}})_{\text{marg}(\mathcal{L} \cup \mathcal{L}')}$. From Theorem 3.1.10 it follows that $(\mathcal{M}^{\text{twin}})_{\text{marg}(\mathcal{L} \cup \mathcal{L}')}$ and $\mathcal{M}^{\text{twin}}$ are interventionally equivalent w.r.t. $(\mathcal{I} \cup \mathcal{I}') \setminus (\mathcal{L} \cup \mathcal{L}')$, from which the result follows. □

3.2 Latent projections

We define a marginalization operation for directed graphs, which we will call the "latent projection", as follows:

Definition 3.2.1 *Given a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and a subset of nodes $\mathcal{L} \subseteq \mathcal{V}$, we define the* latent projection of $\mathcal{G}$ w.r.t. $\mathcal{L}$ *as the graph* $\text{marg}(\mathcal{L})(\mathcal{G}) := (\mathcal{V} \setminus \mathcal{L}, \tilde{\mathcal{E}})$ *where*

$$\tilde{\mathcal{E}} := \{i \to j : i, j \in \mathcal{V} \setminus \mathcal{L}, i \to \ell_1 \to \cdots \to \ell_n \to j \in \mathcal{G} \text{ for } n \geq 0,\ \ell_1, \ldots, \ell_n \in \mathcal{L}\}$$

The name "latent projection" is inspired from a similar construction on mixed graphs in Verma (1993). However, the latent projection defined there does not provide a mapping between SCMs, but only a mapping between mixed graphs that is shown to preserve conditional independence properties (see also Tian, 2002). Here, we provide a sufficient condition for the marginalization of an SCM to obey the latent projection, i.e. that the augmented functional graph of the marginal SCM is a subgraph of the latent projection of the augmented functional graph of the original SCM.

In Example 3.1.3 we already saw an example of a marginalization that obeys the latent projection. However, not all marginalizations give rise to a latent projection, as is illustrated in the following example:

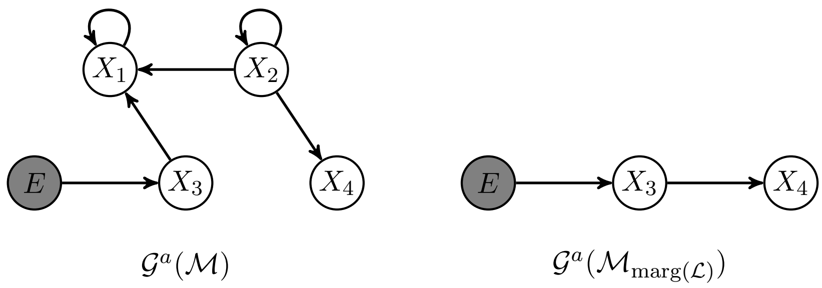

Figure 11: The augmented functional graph of the SCM $\mathcal{M}$ (left) and of $\mathcal{M}_{\mathrm{marg}(\mathcal{L})}$ (right) as in Example 3.2.2.

Example 3.2.2 *Consider the SCM* $\mathcal{M} = \langle \mathbf{4}, \mathbf{1}, \mathbb{R}^4, \mathbb{R}, \boldsymbol{f}, \mathbb{P}_{\mathbb{R}} \rangle$ *with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}, e) &= x_1 \cdot (1 - \mathbf{1}_{\{0\}}(x_2 - x_3)) + 1 \\ f_2(\boldsymbol{x}, e) &= x_2 \\ f_3(\boldsymbol{x}, e) &= e \\ f_4(\boldsymbol{x}, e) &= x_2 \,, \end{aligned}$$

where $\mathbb{P}_{\mathbb{R}}$ *the standard-normal measure on* $\mathbb{R}$*. Although* $\mathcal{M}$ *and its marginalization* $\mathcal{M}_{\mathrm{marg}(\mathcal{L})}$ *with* $\mathcal{L} = \{1, 2\}$ *are interventionally equivalent w.r.t.* $\mathcal{L}$*, the augmented functional graph* $\mathcal{G}^a(\mathcal{M}_{\mathrm{marg}(\mathcal{L})})$ *is not a subgraph of the latent projection of* $\mathcal{G}^a(\mathcal{M})$ *w.r.t.* $\mathcal{L}$*, as can be verified from the augmented functional graphs depicted in Figure 11.*

The next lemma shows that under certain unique solvability conditions one can restrict the components of the mappings that makes the SCM uniquely solvable w.r.t. a certain subset $\mathcal{L}$ to the parents of its ancestral components within the induced subgraph of the augmented functional graph on $\mathcal{L}$.

Lemma 3.2.3 *Given an SCM* $\mathcal{M}$ *and a subset* $\mathcal{L} \subset \mathcal{I}$ *such that* $\mathcal{M}$ *is*

(i) uniquely solvable w.r.t. $\mathcal{L}$*, and*

(ii) uniquely solvable w.r.t. $\widetilde{\mathrm{an}}_{\mathcal{G}^a(\mathcal{M})_{\mathcal{L}}}(\ell)$ *for every* $\ell \in \mathcal{L}$*, where* $\widetilde{\mathrm{an}}_{\mathcal{G}^a(\mathcal{M})_{\mathcal{L}}}(\ell)$ *are the ancestors of* ℓ *according to the induced subgraph* $\mathcal{G}^a(\mathcal{M})_{\mathcal{L}}$ *of the augmented functional graph* $\mathcal{G}^a(\mathcal{M})$ *on* $\mathcal{L}$*.*

Let $\boldsymbol{g}_{\mathcal{L}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{L}) \setminus \mathcal{L}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{L})} \to \boldsymbol{\mathcal{X}}_{\mathcal{L}}$ *be the mapping that makes* $\mathcal{M}$ *uniquely solvable w.r.t.* $\mathcal{L}$*, then for every subset* $\mathcal{K} \subseteq \mathcal{L}$ *the components* $\boldsymbol{g}_{\mathcal{K}}$ *of* $\boldsymbol{g}_{\mathcal{L}}$ *is a mapping of the form* $\boldsymbol{g}_{\mathcal{K}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{A}_{\mathcal{K}}) \setminus \mathcal{A}_{\mathcal{K}}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{A}_{\mathcal{K}})} \to \boldsymbol{\mathcal{X}}_{\mathcal{K}}$*, where* $\mathcal{A}_{\mathcal{K}} := \widetilde{\mathrm{an}}_{\mathcal{G}^a(\mathcal{M})_{\mathcal{L}}}(\mathcal{K})$ *are the ancestors of* $\mathcal{K}$ *according to the induced subgraph* $\mathcal{G}^a(\mathcal{M})_{\mathcal{L}}$ *of the augmented functional graph* $\mathcal{G}^a(\mathcal{M})$ *on* $\mathcal{L}$*.*

Proof. The mapping $\boldsymbol{g}_{\mathcal{L}}$ satisfies for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$\boldsymbol{x}_{\mathcal{L}} = \boldsymbol{g}_{\mathcal{L}}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}) \setminus \mathcal{L}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}) \quad \Longrightarrow \quad \boldsymbol{x}_{\mathcal{L}} = \boldsymbol{f}_{\mathcal{L}}(\boldsymbol{x}, \boldsymbol{e}) \,.$$

Define $\mathcal{A}_\ell := \widetilde{\mathrm{an}}_{\mathcal{G}^a(\mathcal{M})_{\mathcal{L}}}(\ell)$ for some $\ell \in \mathcal{L}$ and let $\tilde{\boldsymbol{g}}_{\mathcal{A}_\ell}$ be the mapping that makes $\mathcal{M}$ uniquely solvable w.r.t. $\mathcal{A}_\ell$. Then, for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$\begin{cases} \boldsymbol{x}_{\mathcal{A}_\ell} &= \boldsymbol{g}_{\mathcal{A}_\ell}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L})\setminus\mathcal{L}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}) \\ \boldsymbol{x}_{\mathcal{L}\setminus\mathcal{A}_\ell} &= \boldsymbol{g}_{\mathcal{L}\setminus\mathcal{A}_\ell}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L})\setminus\mathcal{L}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}) \end{cases} \implies \begin{cases} \boldsymbol{x}_{\mathcal{A}_\ell} &= \tilde{\boldsymbol{g}}_{\mathcal{A}_\ell}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{A}_\ell)\setminus\mathcal{A}_\ell}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{A}_\ell)}) \\ \boldsymbol{x}_{\mathcal{L}\setminus\mathcal{A}_\ell} &= \boldsymbol{f}_{\mathcal{L}\setminus\mathcal{A}_\ell}(\boldsymbol{x}, \boldsymbol{e}) \end{cases}$$

Since for the endogenous variables $\mathrm{pa}(\mathcal{A}_\ell) \setminus \mathcal{A}_\ell \subseteq \mathrm{pa}(\mathcal{L}) \setminus \mathcal{L}$, we have that in particular for $\mathbb{P}_{\boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{L})}}$-almost every $\boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}$ for all $\boldsymbol{x}_{\mathcal{A}_\ell \cup (\mathrm{pa}(\mathcal{L})\setminus\mathcal{L})} \in \boldsymbol{\mathcal{X}}_{\mathcal{A}_\ell \cup (\mathrm{pa}(\mathcal{L})\setminus\mathcal{L})}$:

$$\boldsymbol{x}_{\mathcal{A}_\ell} = \boldsymbol{g}_{\mathcal{A}_\ell}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L})\setminus\mathcal{L}}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L})}) \implies \boldsymbol{x}_{\mathcal{A}_\ell} = \tilde{\boldsymbol{g}}_{\mathcal{A}_\ell}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{A}_\ell)\setminus\mathcal{A}_\ell}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{A}_\ell)}) \,.$$

This implies that that the mapping $\boldsymbol{g}_{\mathcal{A}_\ell}$ cannot depend on variables different from $\boldsymbol{x}_{\mathrm{pa}(\mathcal{A}_\ell)\setminus\mathcal{A}_\ell}$ and $\boldsymbol{e}_{\mathrm{pa}(\mathcal{A}_\ell)}$, because if it does, it leads to a contradiction. In particular, the component g_ℓ can only depend on the endogenous variables $\mathrm{pa}(\mathcal{A}_\ell) \setminus \mathcal{A}_\ell$ and exogenous variables $\mathrm{pa}(\mathcal{A}_\ell)$. We conclude that for any subset $\mathcal{K} \subseteq \mathcal{I}$, the components $\boldsymbol{g}_{\mathcal{K}}$ is a mapping of the form $\boldsymbol{g}_{\mathcal{K}} : \boldsymbol{\mathcal{X}}_{\mathrm{pa}(\mathcal{A}_{\mathcal{K}})\setminus\mathcal{A}_{\mathcal{K}}} \times \boldsymbol{\mathcal{E}}_{\mathrm{pa}(\mathcal{A}_{\mathcal{K}})} \to \boldsymbol{\mathcal{X}}_{\mathcal{K}}$. □

Note that in general condition (ii) in Lemma 3.2.3 does not follow from condition (i) (see Section 2.6.4), this holds only when one drops the uniqueness condition, as was proven in Proposition 2.6.13.

With Lemma 3.2.3 at hand we can prove that the marginalization of an SCM under additional conditions does obey the latent projection.

Proposition 3.2.4 *Given an SCM $\mathcal{M}$ and a subset $\mathcal{L} \subset \mathcal{I}$ such that $\mathcal{M}$ is*

(i) uniquely solvable w.r.t. $\mathcal{L}$, and

(ii) uniquely solvable w.r.t. $\widetilde{\mathrm{an}}_{\mathcal{G}^a(\mathcal{M})_{\mathcal{L}}}(\ell)$ for every $\ell \in \mathcal{L}$, where $\widetilde{\mathrm{an}}_{\mathcal{G}^a(\mathcal{M})_{\mathcal{L}}}(\ell)$ are the ancestors of ℓ according to the induced subgraph $\mathcal{G}^a(\mathcal{M})_{\mathcal{L}}$ of the augmented functional graph $\mathcal{G}^a(\mathcal{M})$ on $\mathcal{L}$,

then $\big(\mathcal{G}^a \circ \mathrm{marg}(\mathcal{L})\big)(\mathcal{M}) \subseteq \big(\mathrm{marg}(\mathcal{L}) \circ \mathcal{G}^a\big)(\mathcal{M})$.

Proof. Let $\boldsymbol{g}_{\mathcal{L}}$ be the mapping that makes $\mathcal{M}$ uniquely solvable w.r.t. $\mathcal{L}$, then by Lemma 3.2.3 the component g_ℓ for every $\ell \in \mathcal{L}$ can only depend on the endogenous variables $\mathrm{pa}(\mathcal{A}_\ell) \setminus \mathcal{A}_\ell$ and exogenous variables $\mathrm{pa}(\mathcal{A}_\ell)$. Hence, every component $\tilde{f}_j$ of the marginal causal mechanism $\tilde{\boldsymbol{f}}$ of $\mathrm{marg}(\mathcal{L})(\mathcal{M})$ for $j \in \mathcal{I} \setminus \mathcal{L}$ depends on no other variables than those $i \in \mathcal{I} \setminus \mathcal{L}$ such that there exists a path $i \to \ell_1 \to \cdots \to \ell_n \to j \in \mathcal{G}^a(\mathcal{M})$ for $n \geq 0$ and $\ell_1, \ldots, \ell_n \in \mathcal{L}$. Therefore, the augmented functional graph $\mathcal{G}^a\big(\mathrm{marg}(\mathcal{L})(\mathcal{M})\big)$ is a subgraph of the latent projection $\mathrm{marg}(\mathcal{L})\big(\mathcal{G}^a(\mathcal{M})\big)$. □

The following example illustrates why the augmented functional graph of a marginalized SCM can be a strict subgraph of the corresponding latent projection:

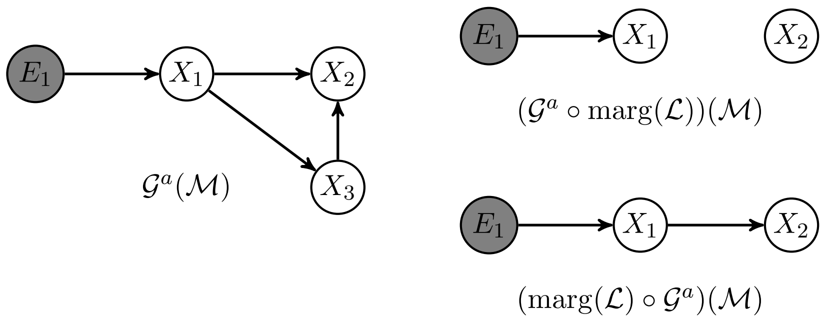

Figure 12: Example of a marginalization of $\mathcal{M}$ w.r.t. $\mathcal{L} = \{3\}$ for which the augmented functional graph is a strict subgraph of the latent projection of $\mathcal{G}^a(\mathcal{M})$ w.r.t. $\mathcal{L}$, as described in Example 3.2.5.

Example 3.2.5 *Consider the SCM given by* $\mathcal{M} = \langle \mathbf{3}, \mathbf{1}, \mathbb{R}^3, \mathbb{R}, \boldsymbol{f}, \mathbb{P}_\mathbb{R} \rangle$, *where*

$$\begin{aligned} f_1(\boldsymbol{x}, \boldsymbol{e}) &= e_1 \\ f_2(\boldsymbol{x}, \boldsymbol{e}) &= x_1 - x_3 \\ f_3(\boldsymbol{x}, \boldsymbol{e}) &= x_1 \,, \end{aligned}$$

and take for $\mathbb{P}_\mathbb{R}$ *the standard-normal measure on* $\mathbb{R}$*. Marginalizing over* $\{3\}$ *gives us the marginal causal mechanism*

$$\begin{aligned} \tilde{f}_1(\boldsymbol{x}, \boldsymbol{e}) &= e_1 \\ \tilde{f}_2(\boldsymbol{x}, \boldsymbol{e}) &= 0 \,. \end{aligned}$$

Here we see that the causal mechanism f_2 *does not not depend on* x_1*, as one would expect from the latent projection (see Figure 12).*

Under the stronger conditions of Proposition 3.2.4 we can also prove the converse of Lemma 3.1.6, this gives:

Lemma 3.2.6 *Given an SCM* $\mathcal{M}$ *and two disjoint subsets* $\mathcal{L}_1, \mathcal{L}_2 \subset \mathcal{I}$ *such that* $\mathcal{M}$ *is uniquely solvable w.r.t.* $\mathcal{L}_1$*. Then* $\mathcal{M}_{\mathrm{marg}(\mathcal{L}_1)}$ *is uniquely solvable w.r.t.* $\mathcal{L}_2$ *if and only if* $\mathcal{M}$ *is*

(i) uniquely solvable w.r.t. $\mathcal{L}_1 \cup \mathcal{L}_2$*, and*

(ii) uniquely solvable w.r.t. $\widetilde{\mathrm{an}}_{\mathcal{G}^a(\mathcal{M})_{\mathcal{L}_1 \cup \mathcal{L}_2}}(\ell)$ *for every* $\ell \in \mathcal{L}_1 \cup \mathcal{L}_2$*, where* $\widetilde{\mathrm{an}}_{\mathcal{G}^a(\mathcal{M})_{\mathcal{L}_1 \cup \mathcal{L}_2}}(\ell)$ *are the ancestors of* ℓ *according to the induced subgraph* $\mathcal{G}^a(\mathcal{M})_{\mathcal{L}_1 \cup \mathcal{L}_2}$ *of the augmented functional graph* $\mathcal{G}^a(\mathcal{M})$ *on* $\mathcal{L}_1 \cup \mathcal{L}_2$*.*

Moreover $\mathrm{marg}(\mathcal{L}_2) \circ \mathrm{marg}(\mathcal{L}_1)(\mathcal{M}) \equiv \mathrm{marg}(\mathcal{L}_1 \cup \mathcal{L}_2)(\mathcal{M})$*.*

Proof. By Lemma 3.1.6 it suffices to show that $\mathcal{M}_{\mathrm{marg}(\mathcal{L}_1)}$ is uniquely solvable w.r.t. $\mathcal{L}_2$. If $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L}_1 \cup \mathcal{L}_2$, then there exists a mapping $\boldsymbol{h}_{\mathcal{L}_1 \cup \mathcal{L}_2}$ such that for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{h}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2) \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2)}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{h}_{\mathcal{L}_2}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2) \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2)}) \end{cases} \iff \begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{f}_{\mathcal{L}_1}(\boldsymbol{x}, \boldsymbol{e}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{f}_{\mathcal{L}_2}(\boldsymbol{x}, \boldsymbol{e}) \end{cases},$$

By Lemma 3.2.3 each component h_ℓ for $\ell \in \mathcal{L}_2$ can only depend on the endogenous variables $\mathrm{pa}(\mathcal{A}_\ell) \setminus \mathcal{A}_\ell$ and exogenous variables $\mathrm{pa}(\mathcal{A}_\ell)$ where $\mathcal{A}_\ell := \widetilde{\mathrm{an}}_{\mathcal{G}^a(\mathcal{M})_{\mathcal{L}_1 \cup \mathcal{L}_2}}(\ell)$. Moreover, we have for the endogenous variables $\mathrm{pa}(\mathcal{A}_\ell) \setminus \mathcal{A}_\ell \subseteq \widehat{\mathrm{pa}}(\mathcal{L}_2) \setminus \mathcal{L}_2$ and for the exogenous variables $\mathrm{pa}(\mathcal{A}_\ell) \subseteq \widehat{\mathrm{pa}}(\mathcal{L}_2)$. Take now the Ansatz for the mapping $\tilde{\boldsymbol{h}}_{\mathcal{L}_2} : \boldsymbol{\mathcal{X}}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)\setminus\mathcal{L}_2} \times \boldsymbol{\mathcal{E}}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)} \to \boldsymbol{\mathcal{X}}_{\mathcal{L}_2}$ given by $\tilde{\boldsymbol{h}}_{\mathcal{L}_2} := \boldsymbol{h}_{\mathcal{L}_2}$. Then, for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$:

$$
\begin{aligned}
&\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{h}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2)\setminus(\mathcal{L}_1 \cup \mathcal{L}_2)}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2)}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{h}_{\mathcal{L}_2}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2)\setminus(\mathcal{L}_1 \cup \mathcal{L}_2)}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2)}) \,, \end{cases} \\
\iff &\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{f}_{\mathcal{L}_1}(\boldsymbol{x}, \boldsymbol{e}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{f}_{\mathcal{L}_2}(\boldsymbol{x}, \boldsymbol{e}) \end{cases} \\
\iff &\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{f}_{\mathcal{L}_2}(\boldsymbol{x}, \boldsymbol{e}) \end{cases} \\
\iff &\begin{cases} \boldsymbol{x}_{\mathcal{L}_1} &= \boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}) \\ \boldsymbol{x}_{\mathcal{L}_2} &= \boldsymbol{f}_{\mathcal{L}_2}(\boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}), \boldsymbol{x}_{\mathcal{I}\setminus\mathcal{L}_1}, \boldsymbol{e}) \end{cases}
\end{aligned}
$$

This gives for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x}_{\mathcal{I}\setminus\mathcal{L}_1} \in \boldsymbol{\mathcal{X}}_{\mathcal{I}\setminus\mathcal{L}_1}$:

$$
\begin{aligned}
&\boldsymbol{x}_{\mathcal{L}_2} = \boldsymbol{h}_{\mathcal{L}_2}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2)\setminus(\mathcal{L}_1 \cup \mathcal{L}_2)}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1 \cup \mathcal{L}_2)}) \\
\iff &\boldsymbol{x}_{\mathcal{L}_2} = \boldsymbol{f}_{\mathcal{L}_2}(\boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}), \boldsymbol{x}_{\mathcal{I}\setminus\mathcal{L}_1}, \boldsymbol{e}) \,,
\end{aligned}
$$

And hence for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x}_{\mathcal{I}\setminus\mathcal{L}_1} \in \boldsymbol{\mathcal{X}}_{\mathcal{I}\setminus\mathcal{L}_1}$:

$$
\begin{aligned}
&\boldsymbol{x}_{\mathcal{L}_2} = \tilde{\boldsymbol{h}}_{\mathcal{L}_2}(\boldsymbol{x}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)\setminus\mathcal{L}_2}, \boldsymbol{e}_{\widehat{\mathrm{pa}}(\mathcal{L}_2)}) \\
\iff &\boldsymbol{x}_{\mathcal{L}_2} = \boldsymbol{f}_{\mathcal{L}_2}(\boldsymbol{g}_{\mathcal{L}_1}(\boldsymbol{x}_{\mathrm{pa}(\mathcal{L}_1)\setminus\mathcal{L}_1}, \boldsymbol{e}_{\mathrm{pa}(\mathcal{L}_1)}), \boldsymbol{x}_{\mathcal{I}\setminus\mathcal{L}_1}, \boldsymbol{e}) \,.
\end{aligned}
$$

and thus $\mathcal{M}_{\mathrm{marg}(\mathcal{L}_1)}$ is uniquely solvable w.r.t. $\mathcal{L}_2$. □

This leads to the following result:

Proposition 3.2.7 *Any marginalization of an SCM that is uniquely solvable w.r.t. every subset is uniquely solvable w.r.t. every subset.*

Proof. Take two disjoint subsets $\mathcal{L}_1$ and $\mathcal{L}_2$ in $\mathcal{I}$. Then, it suffices to show that $\mathcal{M}_{\mathrm{marg}(\mathcal{L}_1)}$ is uniquely solvable w.r.t. $\mathcal{L}_2$. But, this follows directly from Lemma 3.2.6. □

That is, one can perform on an SCM that is uniquely solvable w.r.t. every subset any number of marginalizations in any order. Moreover, all these marginalizations obey the latent projection (see Proposition 3.2.4) and each resulting marginal SCM has no self-loops (See Proposition 2.6.24).

From Lemma 2.4.4 and Proposition 2.6.18 and 3.2.4 together we recover the known result that the class of acyclic SCMs is closed under both intervention and marginalization. Similarly, from Proposition 2.6.27 and 3.2.7 it follows that the class of SCMs that are uniquely solvable w.r.t. every subset is also closed under both intervention and marginalization. This makes these classes convenient to work with.

Linear SCMs In Lemma 2.6.23 we saw that a linear SCM $\mathcal{M}$ is uniquely solvable w.r.t. $\mathcal{L} \subseteq \mathcal{I}$ if and only if the matrix $\mathbb{I}_{\mathcal{L}\mathcal{L}} - B_{\mathcal{L}\mathcal{L}}$ is invertible. This leads to the following result:

Proposition 3.2.8 *Given a linear SCM $\mathcal{M}$ and a subset $\mathcal{L} \subset \mathcal{I}$ of endogenous variables such that $\mathbb{I}_{\mathcal{L}\mathcal{L}} - B_{\mathcal{L}\mathcal{L}}$ is invertible. Then there exists a marginalization $\mathcal{M}_{\mathrm{marg}(\mathcal{L})}$ that is linear and with marginal causal mechanism $\tilde{\boldsymbol{f}} : \mathbb{R}^{\mathcal{O}} \times \mathbb{R}^{\mathcal{J}} \to \mathbb{R}^{\mathcal{O}}$ given by*

$$\begin{aligned}\tilde{\boldsymbol{f}}(\boldsymbol{x}_{\mathcal{O}}, \boldsymbol{e}) =&[B_{\mathcal{O}\mathcal{O}} + B_{\mathcal{O}\mathcal{L}}(\mathbb{I}_{\mathcal{L}\mathcal{L}} - B_{\mathcal{L}\mathcal{L}})^{-1} B_{\mathcal{L}\mathcal{O}}]\boldsymbol{x}_{\mathcal{O}} \\ &+ [B_{\mathcal{O}\mathcal{L}}(\mathbb{I}_{\mathcal{L}\mathcal{L}} - B_{\mathcal{L}\mathcal{L}})^{-1}\Gamma_{\mathcal{L}\mathcal{J}} + \Gamma_{\mathcal{O}\mathcal{J}}]\boldsymbol{e}\,.\end{aligned}$$

Proof. By the definition of marginalization and Lemma 2.6.23 we get for $\mathbb{P}_{\mathbb{R}^{\mathcal{J}}}$-almost every $\boldsymbol{e}$, for all $\boldsymbol{x}_{\mathcal{O}} \in \boldsymbol{\mathcal{X}}_{\mathcal{O}}$:

$$\begin{aligned}\tilde{\boldsymbol{f}}(\boldsymbol{x}_{\mathcal{O}}, \boldsymbol{e}) =&\boldsymbol{f}_{\mathcal{O}}(\boldsymbol{x}_{\mathcal{O}}, \boldsymbol{g}_{\mathcal{L}}(\boldsymbol{x}_{\mathcal{O}}, \boldsymbol{e}), \boldsymbol{e}) \\ =&B_{\mathcal{O}\mathcal{O}}\boldsymbol{x}_{\mathcal{O}} + B_{\mathcal{O}\mathcal{L}}\boldsymbol{g}_{\mathcal{L}}(\boldsymbol{x}_{\mathcal{O}}, \boldsymbol{e}) + \Gamma_{\mathcal{O}\mathcal{J}}\boldsymbol{e} \\ =&[B_{\mathcal{O}\mathcal{O}} + B_{\mathcal{O}\mathcal{L}}(\mathbb{I}_{\mathcal{L}\mathcal{L}} - B_{\mathcal{L}\mathcal{L}})^{-1} B_{\mathcal{L}\mathcal{O}}]\boldsymbol{x}_{\mathcal{O}} \\ &+ [B_{\mathcal{O}\mathcal{L}}(\mathbb{I}_{\mathcal{L}\mathcal{L}} - B_{\mathcal{L}\mathcal{L}})^{-1}\Gamma_{\mathcal{L}\mathcal{J}} + \Gamma_{\mathcal{O}\mathcal{J}}]\boldsymbol{e}\end{aligned}$$

□

Hence the class of linear SCMs is closed under marginalization. From Theorem 3.1.10 we know that $\mathcal{M}$ and its marginalization $\mathcal{M}_{\mathrm{marg}(\mathcal{L})}$ over $\mathcal{L}$ are interventionally equivalent. These results can also be found in Hyttinen et al. (2012).

Remark: In the latent projection w.r.t. $\mathcal{L}$ we had to replace every directed edge $k \to j$ for $j \in \mathcal{O}, k \in \mathcal{L}$ by the set of directed edges $i \to j$ with $i \in \mathcal{O} \cup \mathcal{J}$ whenever there is a directed path $i \to \ell \to \cdots \to k \to j$ where the subpath $\ell \to \cdots \to k$ is a sequence of directed edges between the nodes in $\mathcal{L}$. This substitution of the set of sequences of directed edges in $\mathcal{L}$ is precisely described by the weighted adjacency matrix $(\mathbb{I}_{\mathcal{L}\mathcal{L}} - B_{\mathcal{L}\mathcal{L}})^{-1}$. In particular, if the spectral radius of $B_{\mathcal{L}\mathcal{L}}$ is less than one, then $(\mathbb{I}_{\mathcal{L}\mathcal{L}} - B_{\mathcal{L}\mathcal{L}})^{-1} = \sum_{n=0}^{\infty}(B_{\mathcal{L}\mathcal{L}})^n$, i.e. the substitution of the set of sequences is described by the matrix that sums all the weighted adjacency matrices representing paths through latent variables of length n.

3.3 The direct causal graph w.r.t. a context

In Section 2.8 we discussed the direct causal graph for SCMs where we implicitly assumed the endogenous variables $\mathcal{I}$ as the context variable set. Suppose we want to define the direct causal graph w.r.t. a different context variable set, for example we only observe part of the endogenous variables, then fixing variables that are lying outside this context variable set is not possible and hence defining a direct cause between variables w.r.t. this context using Definition 2.8.3 is not a sensible thing to do.

Figure 13: The direct causal graph of the SCM $\mathcal{M}$ w.r.t. $\mathcal{I}$ (left) and w.r.t. $\{1,3\}$ (right) as in Example 3.3.1.

Example 3.3.1 *Consider the SCM* $\mathcal{M} = \langle \mathbf{3}, \emptyset, \mathbb{R}^3, 1, \boldsymbol{f}, \mathbb{P}_1 \rangle$ *with causal mechanism*

$$f_1(\boldsymbol{x}) = 0$$
$$f_2(\boldsymbol{x}) = x_1$$
$$f_3(\boldsymbol{x}) = x_2 .$$

The direct causal graph w.r.t. $\mathcal{I}$ *is shown in Figure 13 on the left. If we cannot observe variable* x_2 *then we can only intervene on* x_1 *and* x_3*. Performing different perfect interventions on* x_1 *will lead to a distributional change in* x_3*, however there is not a directed edge in* $\mathcal{G}^{dc}(\mathcal{M})$ *pointing from* x_1 *to* x_3*. As we will see below, the direct causal graph w.r.t.* $\{1,3\}$ *is depicted as in Figure 13 on the right.*

We can define the direct causal graph of an SCM w.r.t. a context variable set by using marginalization:

Definition 3.3.2 *Let* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *be an SCM and* $\mathcal{O} \subseteq \mathcal{I}$*. Assume that* $\mathcal{M}$ *is uniquely solvable w.r.t.* $\mathcal{I} \setminus \mathcal{O}$ *and* $\mathcal{M}_{\mathrm{marg}(\mathcal{I}\setminus\mathcal{O})}$ *is structurally uniquely solvable, then the* direct causal graph w.r.t. $\mathcal{O}$ *is the directed graph* $\mathcal{G}^{dc}_{\mathcal{O}}(\mathcal{M}) := \mathcal{G}^{dc}(\mathcal{M}_{\mathrm{marg}(\mathcal{I}\setminus\mathcal{O})})$*.*

Note that this definition is invariant under the equivalence relation $\equiv$ on SCMs. In Example 3.3.1 this yields the direct causal graph w.r.t. $\{1,3\}$ as depicted in Figure 13 on the right. Moreover, from the graph on the right we can exactly read the indirect causes between the two variables x_1 and x_3 of the original SCM.

Definition 3.3.3 *Let* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *be an SCM and consider two different* i *and* j *in* $\mathcal{I}$*. Assume that* $\mathcal{M}$ *is uniquely solvable w.r.t.* $\mathcal{I}\setminus\{i,j\}$ *and that* $\mathcal{M}_{\mathrm{marg}(\mathcal{I}\setminus\{i,j\})}$ *is structurally uniquely solvable, then* i *is an* indirect cause *of* j*, if there exists a directed edge* $i \to j$ *in* $\mathcal{G}^{dc}_{\{i,j\}}(\mathcal{M})$*.*

Similarly to the property that, in general, the marginalization of an SCM does not obey the latent projection w.r.t. its corresponding augmented functional graph, we have that, in general, the marginalization of an SCM does not obey the latent projection of its corresponding direct causal graph, as is illustrated by the following example:

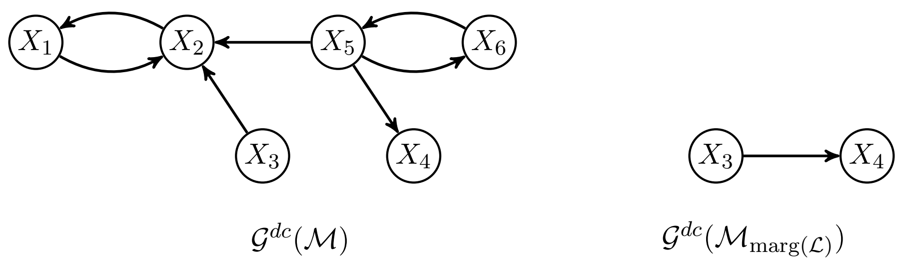

Figure 14: The direct causal graph of the SCM $\mathcal{M}$ (left) and $\mathcal{M}_{\text{marg}(\mathcal{L})}$ (right) as in Example 3.3.4.

Example 3.3.4 *Consider the SCM* $\mathcal{M} = \langle \mathbf{6}, \mathbf{1}, \mathbb{R}^4, \mathbb{R}, \boldsymbol{f}, \mathbb{P}_\mathbb{R}\rangle$ *with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}, e) &= x_2 \\ f_2(\boldsymbol{x}, e) &= x_1 \cdot (1 - \mathbf{1}_{\{0\}}(x_3 - x_5)) + 1 \\ f_3(\boldsymbol{x}, e) &= e \\ f_4(\boldsymbol{x}, e) &= x_5 \\ f_5(\boldsymbol{x}, e) &= x_6 \\ f_6(\boldsymbol{x}, e) &= x_5 \,, \end{aligned}$$

where $\mathbb{P}_\mathbb{R}$ *the standard-normal measure on* $\mathbb{R}$. *Note that* $\mathcal{M}$ *is structurally uniquely solvable and hence the direct causal graph with respect to the context variable set* $\mathcal{I}$ *is given in Figure 14 on the left. The marginal SCM* $\mathcal{M}_{\text{marg}(\mathcal{L})} = \langle \{3,4\}, \mathbf{1}, \mathbb{R}^2, \mathbb{R}\tilde{\boldsymbol{f}}, \mathbb{P}_\mathbb{R}\rangle$ *for* $\mathcal{L} = \{1,2,5,6\}$ *has as causal mechanism:*

$$\begin{aligned} \tilde{f}_3(\boldsymbol{x}) &= e \\ \tilde{f}_4(\boldsymbol{x}) &= x_3 \,. \end{aligned}$$

The direct causal graph of $\mathcal{M}_{\text{marg}(\mathcal{L})}$ *is depicted in Figure 14 on the right. Hence,* x_3 *is an indirect cause of* x_4*, although* x_4 *is not an ancestor of* x_3 *in the direct causal graph of* $\mathcal{M}$.

One may wonder if, under the unique solvability w.r.t. $\mathcal{L}$ condition, there exists an additional condition such that the identity $\mathcal{G}^{dc}(\mathcal{M}) \subseteq \text{marg}(\mathcal{L}) \circ \mathcal{G}^{dc}(\mathcal{M})$ does hold in general. We know that under the additional conditions of Proposition 3.2.4 we have that the marginalization of an SCM does obey the latent projection of its corresponding augmented functional graph, and together with the result of Proposition 2.8.6 we have

$$\begin{array}{ccc} \text{marg}(\mathcal{L}) \circ \mathcal{G}^{dc}(\mathcal{M}) & \subseteq & \text{marg}(\mathcal{L}) \circ \mathcal{G}^{a}(\mathcal{M})_\mathcal{I} \\ & & \cup\!| \\ \mathcal{G}^{dc}(\mathcal{M}_{\text{marg}(\mathcal{L})}) & \subseteq & \mathcal{G}^{a}(\mathcal{M}_{\text{marg}(\mathcal{L})})_{\mathcal{I}\setminus\mathcal{L}} \,. \end{array}$$

In the special case of $\mathcal{G}^{dc}(\mathcal{M}) = \mathcal{G}^{a}(\mathcal{M})_\mathcal{I}$ this gives $\mathcal{G}^{dc}(\mathcal{M}_{\text{marg}(\mathcal{L})}) \subseteq \text{marg}(\mathcal{L}) \circ \mathcal{G}^{dc}(\mathcal{M})$ and hence behavior as in Example 3.3.4 would in this case not occur. However, in general, we do not have that

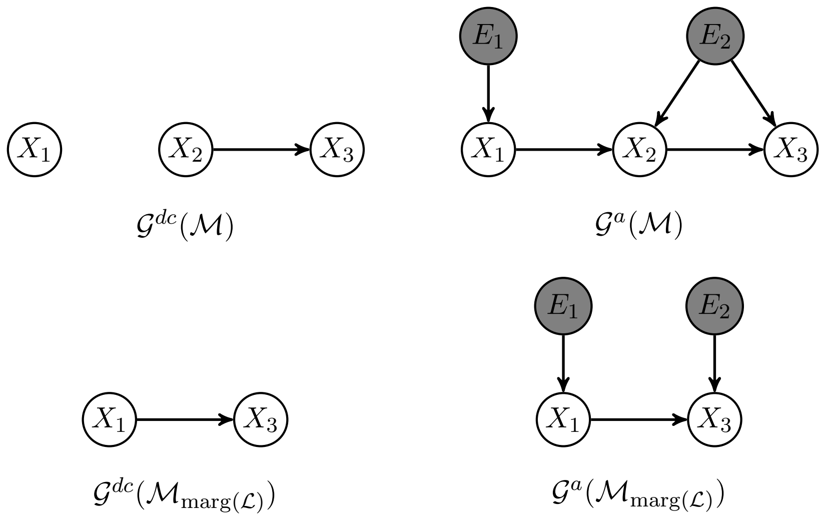

Figure 15: The augmented functional graphs (left) and direct causal graphs (right) of the SCM $\mathcal{M}$ (top) and $\mathcal{M}_{\mathrm{marg}(\mathcal{L})}$ (bottom) as in Example 3.3.5.

$\mathcal{G}^{dc}(\mathcal{M}) = \mathcal{G}^a(\mathcal{M})_{\mathcal{I}}$ and one may wonder if, in general, under the additional assumptions of Proposition 3.2.4 the identity $\mathcal{G}^{dc}(\mathcal{M}_{\mathrm{marg}(\mathcal{L})}) \subseteq \mathrm{marg}(\mathcal{L}) \circ \mathcal{G}^{dc}(\mathcal{M})$ still holds. The next example shows that this is also not the case:

Example 3.3.5 *Consider the acyclic SCM* $\mathcal{M} = \langle \mathbf{3}, \mathbf{2}, \{-1,1\}^2 \times \mathbb{R}, \{-1,1\}^2, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}, \boldsymbol{e}) &= e_1 \\ f_2(\boldsymbol{x}, \boldsymbol{e}) &= x_1 e_2 \\ f_3(\boldsymbol{x}, \boldsymbol{e}) &= x_2 + e_2 \,, \end{aligned}$$

where $\mathbb{P}_{\boldsymbol{\mathcal{E}}} = \mathbb{P}^{\boldsymbol{E}}$ *with* $E_1, E_2 \sim \mathcal{U}(\{-1,1\})$ *and* $E_1 \perp\!\!\!\perp E_2$. *Its augmented functional graph and direct causal graph are depicted in Figure 15 at the top. The marginal SCM* $\mathcal{M}_{\mathrm{marg}(\mathcal{L})} = \langle \{1,3\}, \mathbf{2}, \{-1,1\} \times \mathbb{R}, \{-1,1\}^2, \tilde{\boldsymbol{f}}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$ *for* $\mathcal{L} = \{2\}$ *has as causal mechanism:*

$$\begin{aligned} \tilde{f}_1(\boldsymbol{x}, \boldsymbol{e}) &= e_1 \\ \tilde{f}_3(\boldsymbol{x}, \boldsymbol{e}) &= (1 + x_1) e_2 \,. \end{aligned}$$

Its corresponding augmented functional graph and direct causal graph are depicted in Figure 15 at the bottom. From the figures one can read that $\mathcal{G}^{dc}(\mathcal{M}_{\mathrm{marg}(\mathcal{L})}) \nsubseteq \mathrm{marg}(\mathcal{L}) \circ \mathcal{G}^{dc}(\mathcal{M})$.

This shows that, even under the very strong assumption of acyclicity, the marginalization of an SCM does, in general, not obey the latent projection of its corresponding direct causal graph. This implies, that even for acyclic SCMs, an indirect cause does not necessarily have to come from a single direct cause or a concatenation of several direct causes.

3.4 Latent confounders

Latent confounders were not part of the direct causal graph of the SCM, because the direct causal graph describes only the direct causal relations among the endogenous variables and not among the endogenous and exogenous variables together. If we would like to treat exogenous variables as possible latent confounders, i.e. latent common causes, we should treat them instead as possible endogenous variables.

Definition 3.4.1 *Given an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$*, we define the* extended Structural Causal Model *as the SCM*

$$\hat{\mathcal{M}} := \langle \mathcal{I} \cup \hat{\mathcal{J}}, \mathcal{J}, \boldsymbol{\mathcal{X}} \times \hat{\boldsymbol{\mathcal{E}}}, \boldsymbol{\mathcal{E}}, \hat{\boldsymbol{f}}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle \,,$$

where $\hat{\mathcal{J}}$ *is a copy of* $\mathcal{J}$ *and the causal mechanism* $\hat{\boldsymbol{f}} : \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}}$ *is the measurable function defined by* $\hat{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{e}) = (\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{x}'), \boldsymbol{e})$*.*

This definition preserves the equivalence relation $\equiv$ on SCMs.

Example 3.4.2 *Consider the SCM* $\mathcal{M} = \langle \mathbf{2}, \mathbf{1}, \mathbb{R}^2, \mathbb{R}, \boldsymbol{f}, \mathbb{P}_{\mathbb{R}} \rangle$ *with the causal mechanism:*

$$\begin{aligned} f_1(\boldsymbol{x}) &= e_1 \\ f_2(\boldsymbol{x}) &= x_1 + e_1 \,, \end{aligned}$$

where we take for $\mathbb{P}_{\mathbb{R}}$ *the standard-normal measure on* $\mathbb{R}$*. The functional graph of* $\mathcal{M}$ *has a bidirected edge which could model a latent confounder, as one can see in the functional graph of the extended SCM* $\hat{\mathcal{M}}$ *of* $\mathcal{M}$*, see Figure 16. Note that the direct causal graph of* $\hat{\mathcal{M}}$ *equals the functional graph of* $\hat{\mathcal{M}}$ *in this case. Moreover, both* $\mathcal{M}$ *and* $\hat{\mathcal{M}}$ *have the same direct causal graph w.r.t.* $\mathbf{2}$*.*

Proposition 3.4.3 *Consider an SCM* $\mathcal{M} = \langle \mathcal{I}, \mathcal{J}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{E}}, \boldsymbol{f}, \mathbb{P}_{\boldsymbol{\mathcal{E}}} \rangle$*, then there always exists an SCM* $\hat{\mathcal{M}}$ *with additional endogenous variables* $\hat{\mathcal{J}}$ *such that*

1. $\mathcal{G}(\hat{\mathcal{M}})$ *has no bidirected edges, and*
2. $\hat{\mathcal{M}}_{\mathrm{marg}(\hat{\mathcal{J}})} \equiv \mathcal{M}$*, and*
3. $\mathcal{G}^{dc}_{\mathcal{I}}(\hat{\mathcal{M}}) = \mathcal{G}^{dc}(\mathcal{M})$*.*

Proof. Take for $\hat{\mathcal{M}}$ the extended SCM of $\mathcal{M}$. Then from the definition of $\hat{\mathcal{M}}$ it follows that $\mathcal{G}(\hat{\mathcal{M}})$ has no bidirected edges. Moreover, note that $\hat{\mathcal{M}}$ is always uniquely solvable w.r.t. $\hat{\mathcal{J}}$ and hence we can marginalize over $\hat{\mathcal{J}}$ by substituting $\boldsymbol{e}$ into $\boldsymbol{x}'$ which gives us the causal mechanism of $\mathcal{M}$. From this the last statement follows, that is $\mathcal{G}^{dc}_{\mathcal{I}}(\hat{\mathcal{M}}) = \mathcal{G}^{dc}(\hat{\mathcal{M}}_{\mathrm{marg}(\hat{\mathcal{J}})}) = \mathcal{G}^{dc}(\mathcal{M})$. □

We can interpret a bidirected edge $i \leftrightarrow j$ in the functional graph of $\mathcal{M}$ as representing latent confounders of i and j, i.e. one or more latent common causes k of i and j such that each such k is a direct cause of i and j.

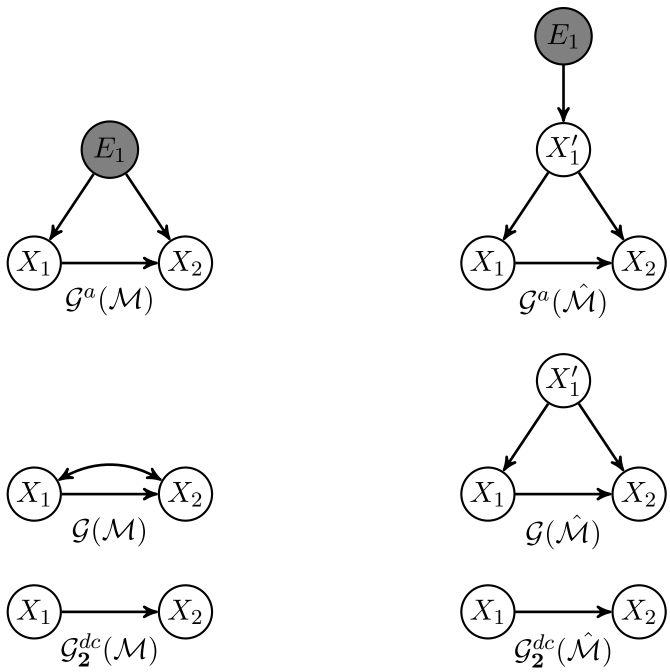

Figure 16: Comparison of the (augmented) functional and direct causal graph of the SCM $\mathcal{M}$ (left) and the extended SCM $\hat{\mathcal{M}}$ (right) of Example 3.4.2.

4 Discussion

To the best of our knowledge, this work provides the first rigorous treatment of cyclic SCMs that deals with measure-theoretic and various other complications that arise due to cycles. We have given a formal definition of SCMs, their solutions, various equivalence classes, their causal interpretation, various graphical representations, discussed their Markov properties and we have defined several operations that can be performed on SCMs.

The central topic of investigation was how to arrive at a parsimonious representation of an SCM on a subset of endogenous variables of interest. Under the condition that the SCM is uniquely solvable w.r.t. the subsystem, we can treat the subsystem effectively as a "black box" with a unique output for every possible input. This allows us to effectively remove this subsystem from the model, obtaining a marginal SCM which preserves the probabilistic, causal and counterfactual semantics on the subsystem. We would like to stress that this particular sufficient condition for the existence of a marginal SCM is not a necessary condition. It is possible to relax this condition further, but doing so is beyond the scope of this paper.

Regarding the use of measure theory in this work, the reader may wonder whether it was really necessary to invoke some of the more advanced results of measure theory here, while in most treatments of

SCMs measure-theoretical aspects are not discussed at all. That this seems necessary indeed is a somewhat unexpected consequence of our decision to develop a general theory that also incorporates cyclic SCMs. That measurability of solutions is not guaranteed in general is just one out of many technical complications that do not occur in the acyclic case. Another complication we encountered is that unique solvability is no longer guaranteed, nor is it preserved under interventions: solutions may either be ill-defined or allow for multiple different probability distributions which leads to ambiguity. On a more conceptual level, the causal interpretation of cyclic SCMs can be counter-intuitive, as we have seen, for example in the cyclic case, where the causal relations between variables need no longer coincide with their functional relations.

All these technical and conceptual complications may explain why in most of the SCM literature so far (except possibly for the linear SEM literature), acyclicity has been assumed, even though there is a real need for a theory of cyclic SCMs given that many causal systems in nature involve cycles. Actually, one may wonder why many systems appear to be acyclic at a macroscopic level, even though on a microscopic level, all particles interact with each other, leading to a fully connected causal graph with cycles on that microscopic level of detail. With this work, we hope to have provided a solid foundation to the theory of cyclic SCMs that will enable such models to be used for the purposes of causal discovery and prediction.

Future work consists of developing an analogue to marginalization for exogenous variables in order to obtain a "minimal" exogenous space that preserves the causal semantics of the model. This allows one to obtain more parsimonious representations of SCMs by reducing the space of latent variables whenever possible. Furthermore, we will investigate how one could extend the class of SCMs to causal models that can deal with constraints.

5 Acknowledgments

We thank Patrick Forré and Robin Evans for helpful discussions and Tineke Blom and Alexander Ly for providing helpful comments on earlier drafts. Stephan Bongers and Joris Mooij were supported by NWO, the Netherlands Organization for Scientific Research (VIDI grant 639.072.410 and VENI grant 639.031.036). Joris Mooij was also supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n$^{\mathrm{o}}$ 639466).

6 Appendix

Measurable selection theorems

Consider a measure space $(\boldsymbol{\mathcal{X}}, \boldsymbol{\Sigma}, \boldsymbol{\mu})$. A set $\boldsymbol{\mathcal{E}} \subseteq \boldsymbol{\mathcal{X}}$ is called a *$\boldsymbol{\mu}$-null set* if there exists a $\boldsymbol{\mathcal{A}} \in \boldsymbol{\Sigma}$ with $\boldsymbol{\mathcal{E}} \subseteq \boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mu}(\boldsymbol{\mathcal{A}}) = 0$. We denote the class of $\boldsymbol{\mu}$-null sets by $\boldsymbol{\mathcal{N}}$, and we denote the σ-algebra generated by $\boldsymbol{\Sigma} \cup \boldsymbol{\mathcal{N}}$ by $\bar{\boldsymbol{\Sigma}}$, and its members are called the *$\boldsymbol{\mu}$-measurable* sets. Note that each member of $\bar{\boldsymbol{\Sigma}}$ is of the form $\boldsymbol{\mathcal{A}} \cup \boldsymbol{\mathcal{E}}$ with $\boldsymbol{\mathcal{A}} \in \boldsymbol{\Sigma}$ and $\boldsymbol{\mathcal{E}} \in \boldsymbol{\mathcal{N}}$. The measure $\boldsymbol{\mu}$ is extended to a measure $\bar{\boldsymbol{\mu}}$ on $\bar{\boldsymbol{\Sigma}}$, by $\bar{\boldsymbol{\mu}}(\boldsymbol{\mathcal{A}} \cup \boldsymbol{\mathcal{E}}) = \boldsymbol{\mu}(\boldsymbol{\mathcal{A}})$ for any $\boldsymbol{\mathcal{A}} \in \boldsymbol{\Sigma}$ and $\boldsymbol{\mathcal{E}} \in \boldsymbol{\mathcal{N}}$, and is called its *completion*. A mapping $\boldsymbol{f} : \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{Y}}$ between measurable spaces is called *$\boldsymbol{\mu}$-measurable* if the inverse image $\boldsymbol{f}^{-1}(\boldsymbol{\mathcal{C}})$ of any measurable set $\boldsymbol{\mathcal{C}} \subseteq \boldsymbol{\mathcal{Y}}$ is $\boldsymbol{\mu}$-measurable.

Lemma 6.0.1 *Consider a $\boldsymbol{\mu}$-measurable mapping $\boldsymbol{f} : \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{Y}}$. If $\boldsymbol{\mathcal{Y}}$ is countably generated, then there exists a measurable mapping $\boldsymbol{g} : \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{Y}}$ such that $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x})$ holds $\boldsymbol{\mu}$-a.e..*

Proof. Let the σ-algebra of $\boldsymbol{\mathcal{Y}}$ be generated by the countably generating set $\{\boldsymbol{\mathcal{C}}_n\}_{n\in\mathbb{N}}$. The $\boldsymbol{\mu}$-measurable set $\boldsymbol{f}^{-1}(\boldsymbol{\mathcal{C}}_n) = \boldsymbol{\mathcal{A}}_n \cup \boldsymbol{\mathcal{E}}_n$ for some $\boldsymbol{\mathcal{A}}_n \in \boldsymbol{\Sigma}$ and some $\boldsymbol{\mathcal{E}}_n \in \boldsymbol{\mathcal{N}}$ and hence there is some $\boldsymbol{\mathcal{E}}_n \subseteq \boldsymbol{\mathcal{B}}_n \in \boldsymbol{\Sigma}$ such that $\boldsymbol{\mu}(\boldsymbol{\mathcal{B}}_n) = 0$. Let $\hat{\boldsymbol{\mathcal{B}}} = \cup_{n\in\mathbb{N}} \boldsymbol{\mathcal{B}}_n$, $\hat{\boldsymbol{\mathcal{A}}}_n = \boldsymbol{\mathcal{A}}_n \setminus \hat{\boldsymbol{\mathcal{B}}}$ and $\hat{\boldsymbol{\mathcal{A}}} = \cup_{n\in\mathbb{N}} \hat{\boldsymbol{\mathcal{A}}}_n$, then $\boldsymbol{\mu}(\hat{\boldsymbol{\mathcal{B}}}) = 0$, $\hat{\boldsymbol{\mathcal{A}}}$ and $\hat{\boldsymbol{\mathcal{B}}}$ are disjoint and $\boldsymbol{\mathcal{X}} = \hat{\boldsymbol{\mathcal{A}}} \cup \hat{\boldsymbol{\mathcal{B}}}$. Now define the mapping $\boldsymbol{g} : \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{Y}}$ by:

$$\boldsymbol{g}(\boldsymbol{x}) := \begin{cases} \boldsymbol{f}(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in \hat{\boldsymbol{\mathcal{A}}}, \\ \boldsymbol{y}_0 & \text{otherwise,} \end{cases}$$

where for $\boldsymbol{y}_0$ we can take an arbitrary point in $\boldsymbol{\mathcal{Y}}$. This mapping $\boldsymbol{g}$ is measurable since for each generator $\boldsymbol{\mathcal{C}}_n$ we have

$$\boldsymbol{g}^{-1}(\boldsymbol{\mathcal{C}}_n) = \begin{cases} \hat{\boldsymbol{\mathcal{A}}}_n & \text{if } \boldsymbol{y}_0 \notin \boldsymbol{\mathcal{C}}_n, \\ \hat{\boldsymbol{\mathcal{A}}}_n \cup \hat{\boldsymbol{\mathcal{B}}} & \text{otherwise.} \end{cases}$$

is in $\boldsymbol{\Sigma}$. Moreover, $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x})$ $\boldsymbol{\mu}$-almost everywhere. □

With this result at hand we can now prove a first measurable selection theorem:

Theorem 6.0.2 *Given a standard probability space $\boldsymbol{\mathcal{E}}$ with probability measure $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$, a standard measurable space $\boldsymbol{\mathcal{X}}$ and a measurable set $\boldsymbol{\mathcal{S}} \subseteq \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{X}}$ such that $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}})$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set, where $\boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}$ is the projection mapping on $\boldsymbol{\mathcal{E}}$. Then there exists a measurable mapping $\boldsymbol{g} : \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{X}}$ such that $(\boldsymbol{e}, \boldsymbol{g}(\boldsymbol{e})) \in \boldsymbol{\mathcal{S}}$ for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$.*

Proof. Take the subset $\hat{\boldsymbol{\mathcal{E}}} := \boldsymbol{\mathcal{E}} \setminus \boldsymbol{\mathcal{B}}$, for some $\boldsymbol{\mathcal{B}} \supseteq \boldsymbol{\mathcal{E}} \setminus \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}})$ and $\mathbb{P}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{B}}) = 0$, and note that $\hat{\boldsymbol{\mathcal{E}}}$ is a standard measurable space (see Corollary 13.4 in Kechris (1995)) and $\hat{\boldsymbol{\mathcal{E}}} \subseteq \boldsymbol{pr}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{S}})$. Let $\hat{\boldsymbol{\mathcal{S}}} = \boldsymbol{\mathcal{S}} \cap \hat{\boldsymbol{\mathcal{E}}} \times \boldsymbol{\mathcal{X}}$. Because the set $\hat{\boldsymbol{\mathcal{S}}}$ is measurable, it is in particular analytic (see Theorem 13.7 in Kechris (1995)). It follows by the Jankov-von Neumann Theorem (Theorem 29.9 in Kechris (1995)) that $\hat{\boldsymbol{\mathcal{S}}}$ has a universally measurable uniformizing function, that is, there exists a universally measurable mapping $\hat{\boldsymbol{g}} : \hat{\boldsymbol{\mathcal{E}}} \to \boldsymbol{\mathcal{X}}$ such that for all $\boldsymbol{e} \in \hat{\boldsymbol{\mathcal{E}}}$, $(\boldsymbol{e}, \hat{\boldsymbol{g}}(\boldsymbol{e})) \in \hat{\boldsymbol{\mathcal{S}}}$. Universal measurability of the mapping $\hat{\boldsymbol{g}}$ means that it is $\boldsymbol{\mu}$-measurable for any σ-finite measure $\boldsymbol{\mu}$ on $\hat{\boldsymbol{\mathcal{E}}}$. Hence, in particular, it is $\mathbb{P}_{\boldsymbol{\mathcal{E}}}\big|_{\hat{\boldsymbol{\mathcal{E}}}}$-measurable, where $\mathbb{P}_{\boldsymbol{\mathcal{E}}}\big|_{\hat{\boldsymbol{\mathcal{E}}}}$ is the restriction of $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$ to $\hat{\boldsymbol{\mathcal{E}}}$.

Now define the mapping $\boldsymbol{g}^* : \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{X}}$ by

$$\boldsymbol{g}^*(\boldsymbol{e}) := \begin{cases} \hat{\boldsymbol{g}}(\boldsymbol{e}) & \text{if } \boldsymbol{e} \in \hat{\boldsymbol{\mathcal{E}}} \\ \boldsymbol{x}_0 & \text{otherwise,} \end{cases}$$

where for $\boldsymbol{x}_0$ we can take an arbitrary point in $\boldsymbol{\mathcal{X}}$. Then this mapping $\boldsymbol{g}^*$ is $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-measurable. To see this, take any measurable set $\boldsymbol{\mathcal{C}} \subseteq \boldsymbol{\mathcal{X}}$, then

$$\boldsymbol{g}^{*-1}(\boldsymbol{\mathcal{C}}) = \begin{cases} \hat{\boldsymbol{g}}^{-1}(\boldsymbol{\mathcal{C}}) & \text{if } \boldsymbol{x_0} \notin \boldsymbol{\mathcal{C}} \\ \hat{\boldsymbol{g}}^{-1}(\boldsymbol{\mathcal{C}}) \cup \boldsymbol{\mathcal{B}} & \text{otherwise.} \end{cases}$$

Because $\hat{\boldsymbol{g}}^{-1}(\boldsymbol{\mathcal{C}})$ is $\mathbb{P}_{\boldsymbol{\mathcal{E}}}\big|_{\hat{\boldsymbol{\mathcal{E}}}}$-measurable it is also $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-measurable and thus $\boldsymbol{g}^{*-1}(\boldsymbol{\mathcal{C}})$ is $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-measurable.

By Lemma 6.0.1 and the fact that standard measurable spaces are countably generated, we prove the existence of a measurable mapping $\boldsymbol{g} : \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{X}}$ such that $\boldsymbol{g}^* = \boldsymbol{g}$ $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-a.e. and thus it satisfies $(\boldsymbol{e}, \boldsymbol{g}(\boldsymbol{e})) \in \boldsymbol{\mathcal{S}}$ for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$. □

This theorem rests on the assumption that the measurable space $\boldsymbol{\mathcal{E}}$ has a probability measure $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$. If this space becomes the product space $\boldsymbol{\mathcal{Y}} \times \tilde{\boldsymbol{\mathcal{E}}}$ for which only the space $\tilde{\boldsymbol{\mathcal{E}}}$ has a probability measure, then in general this theorem does not hold anymore. However, if we assume in addition that the fibers of $\boldsymbol{\mathcal{S}}$ in $\boldsymbol{\mathcal{Y}}$ at every point $(\boldsymbol{x}, \tilde{\boldsymbol{e}})$ are σ-compact, then we have the following measurable selection theorem:

Theorem 6.0.3 *Given a standard probability space $\boldsymbol{\mathcal{E}}$ with probability measure $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$, standard measurable spaces $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ and a measurable set $\boldsymbol{\mathcal{S}} \subseteq \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{Y}}$ such that $\boldsymbol{\mathcal{E}} \setminus \boldsymbol{\mathcal{K}}_\sigma$ is a $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-null set, where*

$$\boldsymbol{\mathcal{K}}_\sigma := \{\boldsymbol{e} \in \boldsymbol{\mathcal{E}} \,:\, \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}} (\boldsymbol{\mathcal{S}}_{(\boldsymbol{x},\boldsymbol{e})} \text{ is non-empty and } \sigma\text{-compact})\}\,,$$

with $\boldsymbol{\mathcal{S}}_{(\boldsymbol{x},\boldsymbol{e})}$ denoting the fiber over $(\boldsymbol{x}, \boldsymbol{e})$, that is

$$\boldsymbol{\mathcal{S}}_{(\boldsymbol{x},\boldsymbol{e})} := \{\boldsymbol{y} \in \boldsymbol{\mathcal{Y}} \,:\, (\boldsymbol{x}, \boldsymbol{e}, \boldsymbol{y}) \in \boldsymbol{\mathcal{S}}\}\,.$$

Then there exists a measurable mapping $\boldsymbol{g} : \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{Y}}$ such that for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ we have $(\boldsymbol{x}, \boldsymbol{e}, \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{e})) \in \boldsymbol{\mathcal{S}}$.

Proof. Take the subset $\hat{\boldsymbol{\mathcal{E}}} := \boldsymbol{\mathcal{E}} \setminus \boldsymbol{\mathcal{B}}$, for some $\boldsymbol{\mathcal{B}} \supseteq \boldsymbol{\mathcal{E}} \setminus \boldsymbol{\mathcal{K}}_\sigma$ and $\mathbb{P}_{\boldsymbol{\mathcal{E}}}(\boldsymbol{\mathcal{B}}) = 0$, and note that $\hat{\boldsymbol{\mathcal{E}}}$ is a standard measurable space, $\hat{\boldsymbol{\mathcal{E}}} \subseteq \boldsymbol{\mathcal{K}}_\sigma$ and $\hat{\boldsymbol{\mathcal{S}}} = \boldsymbol{\mathcal{S}} \cap \boldsymbol{\mathcal{X}} \times \hat{\boldsymbol{\mathcal{E}}} \times \boldsymbol{\mathcal{Y}}$ is measurable. By assumption, for each $(\boldsymbol{x}, \boldsymbol{e}) \in \hat{\boldsymbol{\mathcal{S}}}$ the fiber $\hat{\boldsymbol{\mathcal{S}}}_{(\boldsymbol{x},\boldsymbol{e})}$ is non-empty and σ-compact and hence by applying the Theorem of Arsenin-Kunugui (Theorem 35.46 in Kechris (1995)) it follows that the set $\hat{\boldsymbol{\mathcal{S}}}$ has a measurable uniformizing function, that is, there exists a measurable mapping $\hat{\boldsymbol{g}} : \boldsymbol{\mathcal{X}} \times \hat{\boldsymbol{\mathcal{E}}} \to \boldsymbol{\mathcal{Y}}$ such that for all $(\boldsymbol{x}, \boldsymbol{e}) \in \boldsymbol{\mathcal{X}} \times \hat{\boldsymbol{\mathcal{E}}}$, $(\boldsymbol{x}, \boldsymbol{e}, \hat{\boldsymbol{g}}(\boldsymbol{x}, \boldsymbol{e})) \in \hat{\boldsymbol{\mathcal{S}}}$.

Now define the mapping $\boldsymbol{g} : \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{Y}}$ by

$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{e}) := \begin{cases} \hat{\boldsymbol{g}}(\boldsymbol{x}, \boldsymbol{e}) & \text{if } \boldsymbol{e} \in \hat{\boldsymbol{\mathcal{E}}} \\ \boldsymbol{y}_0 & \text{otherwise,} \end{cases}$$

where for $\boldsymbol{y}_0$ we can take an arbitrary point in $\boldsymbol{\mathcal{Y}}$. This mapping $\boldsymbol{g}$ inherits the measurability from $\hat{\boldsymbol{g}}$ and it satisfies for $\mathbb{P}_{\boldsymbol{\mathcal{E}}}$-almost every $\boldsymbol{e}$ for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$, $(\boldsymbol{x}, \boldsymbol{e}, \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{e})) \in \boldsymbol{\mathcal{S}}$. □

A topological space is σ-compact, if it is the union of countably many compact subspaces. For example, all countable discrete spaces, any interval of the real line, and moreover all the Euclidean spaces are σ-compact spaces.

7 References

Balke, A. and Pearl, J. (1994). Probabilistic evaluation of counterfactual queries. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, volume 1, pages 230–237. MIT Press.

Bollen, K. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons, New York, USA.

Bongers, S. and Mooij, J. M. (2018). From random differential equations to structural causal models: the stochastic case. *arXiv.org preprint*, arXiv:1803.08784 [cs.AI].

Cohn, D. L. (2013). *Measure Theory*. Birkhäuser, Boston, USA, 2nd edition.

Dawid, A. (2002). Influence diagrams for causal modelling and inference. *International Statistical Review*, 70:161–189.

Duncan, O. (1975). *Introduction to Structural Equation Models*. Academic Press, New York.

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

Goldberger, A. and Duncan, O., editors (1973). *Structural Equation Models in the Social Sciences*. Seminar Press, New York.

Haavelmo, T. (1943). The statistical implications of a system of simultaneous equations. *Econometrica*, 11(1):1–12.

Hyttinen, A., Eberhardt, F., and Hoyer, P. (2012). Learning linear cyclic causal models with latent variables. *Journal of Machine Learning Research*, 13:3387–3439.

Kechris, A. (1995). *Classical Descriptive Set Theory*, volume 16 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, USA.

Lauritzen, S. (1996). *Graphical Models*, volume 17 of *Oxford Statistical Science Series*. Clarendon Press, Oxford.

Lauritzen, S., Dawid, A., Larsen, B., and Leimer, H.-G. (1990). Independence properties of directed Markov fields. *Networks*, 20:491–505.

Mason, S. (1953). Feedback theory - Some properties of signal flow graphs. In *Proceedings of the IRE*, volume 41.

Mason, S. (1956). Feedback theory - Further properties of signal flow graphs. In *Proceedings of the IRE*, volume 44.

Mooij, J. M., Janzing, D., and Schölkopf, B. (2013). From ordinary differential equations to structural causal models: the deterministic case. In Nicholson, A. and Smyth, P., editors, *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI-13)*, pages 440–448. AUAI Press.

Neal, R. (2000). On deducing conditional independence from *d*-separation in causal graphs with feedback. *Journal of Artificial Intelligence Research*, 12:87–91.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann Series in Representation and Reasoning. Morgan Kaufmann, 1st edition.

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

Pearl, J. and Dechter, R. (1996). Identifying independence in causal graphs with feedback. In *Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence (UAI-96)*, pages 420–426.

Richardson, T. and Spirtes, P. (2002). Ancestral graph Markov models. *The Annals of Statistics*, 30(4).

Spirtes, P. (1994). Conditional independence in directed cyclic graphical models for feedback. Technical Report CMU-PHIL-54, Carnegie Mellon University.

Spirtes, P. (1995). Directed cyclic graphical representations of feedback models. In *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence (UAI-95)*, pages 491–499.

Spirtes, P., Glymour, C., and Scheines, R. (2000). *Causation, Prediction, and Search*. The MIT Press, Cambridge, Massachusetts, 2nd edition.

Tian, J. (2002). Studies in causal reasoning and learning. Technical Report R-309, Cognitive Systems Laboratory, University of California, Los Angeles.

Verma, T. (1993). Graphical aspects of causal models. Technical Report R-191, Computer Science Department, University of California, Los Angeles.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, 20:557–585.