Proceedings of Machine Learning Research vol 275:1–14, 2025 4th Conference on Causal Learning and Reasoning

Algorithmic Syntactic Causal Identification

Dhurim Cakiqi DXC100@STUDENT.BHAM.AC.UK and **Max A. Little** M.LITTLE.1@BHAM.AC.UK
School of Computer Science, University of Birmingham, UK

Editors: Biwei Huang and Mathias Drton

Abstract

Causal identification in causal Bayes nets (CBNs) is an important tool in causal inference allowing the derivation of interventional distributions from observational distributions where this is possible in principle. However, most existing formulations of causal identification using techniques such as d-separation and do-calculus are expressed within the mathematical language of classical probability theory on CBNs. However, there are many causal settings where probability theory and hence current causal identification techniques are inapplicable such as relational databases, dataflow programs such as hardware description languages, distributed systems and most modern machine learning algorithms. We show that this restriction can be lifted by replacing the use of classical probability theory with the alternative axiomatic foundation of symmetric monoidal categories. In this alternative axiomatization, we show how an unambiguous and clean distinction can be drawn between the general *syntax* of causal models and any specific *semantic* implementation of that causal model. This allows a *purely syntactic* algorithmic description of general causal identification by a translation of recent formulations of the general ID algorithm through *fixing*. Our description is given entirely in terms of the non-parametric ADMG structure specifying a causal model and the algebraic signature of the corresponding monoidal category, to which a sequence of manipulations is then applied so as to arrive at a modified monoidal category in which the desired, purely syntactic interventional causal model, is obtained. We use this idea to derive purely syntactic analogues of classical back-door and front-door causal adjustment, and illustrate an application to a more complex causal model.

1. Introduction

Causal Bayes nets (CBNs) are probabilistic models in which *causal influences* between *random variables* are expressed via the use of *graphs* with *nodes* in these graphs being the random variables and *directed edges* indicating the direction of causality between them (J.Pearl, 2009; Bareinboim et al., 2022). Every such *directed acyclic graph* (DAG) with *latent* (*unobserved*) nodes has a corresponding *acyclic directed mixed graph* (ADMG) which is obtained from the DAG through *latent projection* which simplifies the DAG whilst preserving its *causal d-separation* properties (Richardson et al., 2012; J.Pearl, 2009).

The ADMG $\mathcal{G}=\left(\mathbf{V}^{\mathcal{G}},\mathbf{E}^{\mathcal{G}}\right)$ with variables $\mathbf{V}^{\mathcal{G}}$ and edges $\mathbf{E}^{\mathcal{G}}$ has bidirected edges indicating *unmeasured confounding* (Richardson et al., 2012). Excluding the bidirected edges, the topological properties of these graphs are given by their set-valued *parent* $\mathrm{pa}_{\mathcal{G}}$ and *child* functions $\mathrm{ch}_{\mathcal{G}}$. *Ancestors* $\mathrm{an}_{\mathcal{G}}$ and *descendents* $\mathrm{de}_{\mathcal{G}}$ are determined recursively from these. The *subgraph* $\mathcal{G}_{\mathbf{Y}}$ for $\mathbf{Y}\subset\mathbf{V}^{\mathcal{G}}$ is obtained by deleting from $\mathbf{V}^{\mathcal{G}}$ all nodes not in $\mathbf{Y}$ and edges which connect to those removed variables. A sequence of the subset of the nodes in $\mathbf{V}^{\mathcal{G}}$ such that every node in that sequence occurs before its children (or after its parents) in $\mathcal{G}$ through unidirectional edges, is called a *topological ordering* of the ADMG (Bareinboim et al., 2022). For such an ordering the function $\mathrm{pre}_{\mathcal{G}}\left(V\right)$ gives the set of all the nodes before V in the sequence, and $\mathrm{succ}_{\mathcal{G}}\left(V\right)$ gives all the nodes after it in the sequence. In the ADMG, sets of variables which are connected through a sequence of bidirected edges are called *districts* $\mathrm{dis}_{\mathcal{G}}$ (Richardson et al., 2012). Nodes $V \in \mathbf{V}^{\mathcal{G}}$ which have the property that $\mathrm{dis}_{\mathcal{G}}(V) \cap \mathrm{de}_{\mathcal{G}}(V) = \{V\}$ are called *fixable* nodes (Richardson et al., 2012).

For an ADMG with no bidirected edges (thus, no latent variables, equivalent to a CBN over a DAG), it is always possible to derive any interventional distribution from the joint distribution over the variables in the DAG using the *truncated factorization* (J.Pearl, 2009). However, more generally, in the presence of unobserved confounding (e.g. models having bidirected edges in the ADMG) this is no longer true and only certain interventional distributions can be derived from the observed variables (Shipster and Pearl, 2008; Bareinboim et al., 2022). Pearl's *do-calculus* (J.Pearl, 2009) is a set of three algebraic distribution transformations which it has been shown are necessary and sufficient for deriving the interventional distribution where this is possible (Shipster and Pearl, 2008). These algebraic transformations can be applied ad-hoc or, more systematically, using Shpitser's *ID algorithm* to derive a desired interventional distribution (Shipster and Pearl, 2008).

More recently, the specific conditions under which any particular interventional distribution can be determined from the observed variables using do-calculus or some other systematic algorithm, has been simplified in terms of *fixing operations* and *reachable subgraphs* in causal ADMGs (Richardson et al., 2012). Exploiting the same reasoning, Richardson et al. (2012) show how fixing operations can be combined in a simple algorithm which achieves the same result.

This algorithm, as with most algorithms for causal inference, is expressed in terms of CBNs using random variables and classical probabilities where probabilistic conditioning indicates the direction of causal inference in an ADMG. Such causal identification algorithms rely on simultaneous manipulation of the ADMG, tracking the consequence of such manipulations on the corresponding (joint) distribution over that graph. As long as the appropriate *Markov property* holds (Bareinboim et al., 2022), which guarantees the consistency of the distribution with the CBN, then this is a valid procedure for deriving the desired interventional distribution. Nonetheless, there are many practical settings where probabilistic modelling is inappropriate, such as relational databases (Patterson, 2017), hardware description languages, distributed systems modelled by Petri nets and most modern machine learning algorithms (Little, 2019). In these settings there is no such Markov property therefore it appears that the existing causal identification algorithms are inapplicable in these wider, non-probabilistic applications.

A different and more recently explored direction which might circumvent this limitation is to change the fundamental axiomatic basis of the modelling language to use (*monoidal*) *category theory* instead. This amounts to a fundamental reformulation of CBNs that, rather than organizing causal models around sets, measure theory and graph topology which requires the additional complexity of Markov properties to bind these together, instead views CBNs from the simpler and more abstract vantage point of *structured compositional processes*. Causal modelling and inference in terms of *string diagrams* representing such processes has shown considerable promise. Building on work by Fong (2013), Cho and Jacobs (2019) formulated the essential concepts of Bayesian reasoning as strings, following which Jacobs et al. (2021) provided an exposition of causal identification under a slightly extended form of the *front-door* causal scenario for *Markov categories* (Fritz, 2020). Since then, string analogues of do-calculus and d-separation have been described (Yin and Zhang, 2022; Fritz and Klingler, 2023) and explicit description of extensions of the categorical string diagram approach to causal modelling in non-probabilistic settings such as machine learning (Cakiqi and Little, 2022).

Nonetheless the full promise of this reformulation has yet to be realized. For instance, causal inference in string diagrams has, to date, only been described in probabilistic categories for single variable interventions in discrete sample spaces where interventions can be modelled by discrete uniform distributions (Jacobs et al., 2021), or more generally to string diagrams where causal identification beyond the slightly extended form of the front-door causal scenario is not carried out systematically (algorithmically) but instead requires manual string manipulation (Lorenz and Tull, 2023). Thus, our goal in this report is to provide the first, purely syntactic, algorithmic characterization of general causal identification by fixing which is applicable to the full range of causal models expressible as a structured, categorical compositional process.

2. Theory

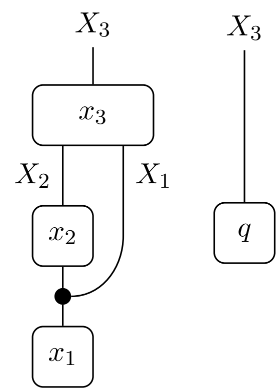
2.1. Symmetric monoidal categories and their algebraic signatures

Symmetric monoidal categories (SMCs) are algebraic structures which capture the notion of simultaneous *sequential* and (in our application) *parallel composition* of maps between types. Examples of such categories include ordinary sets and functions between these sets with the cartesian product indicating parallel composition, the category of sets and relations (Fong, 2013), *Markov categories* of sample spaces with conditional distributions modelled by sets and *probability monads* between them (Fritz, 2020) or other non-deterministic monads in arbitrary semifields (Cakiqi and Little, 2022). Following Selinger (2011), a symmetric monoidal category *signature* Σ provides all the information required to specify a particular SMC. For those unfamiliar with category theory, Riehl (2017) is an excellent introduction.

Definition 1 *(Symmetric monoidal signature). A symmetric monoidal signature* Σ*, consists of a set of* object terms Σ_0 *and* morphism variables*,* Σ_1*, along with a pair of functions* $\mathrm{dom}, \mathrm{cod} : \Sigma_1 \to \mathrm{Mon}\left(\Sigma_0\right)$ *which determine the domain and codomain of the morphism variables respectively. Here,* $\mathrm{Mon}\left(\Sigma_0\right) = \left(\Sigma_0, \otimes, 1\right)$ *is the* free commutative monoid *generated by the object terms in* Σ_0.

For brevity we will use exponential notation to indicate terms in $\mathrm{Mon}\left(\Sigma_0\right)$ with repeated objects, i.e. for $A, B \in \Sigma_0$ the object expression $(A \otimes B) \otimes (1 \otimes A)$ is written A^2B. A morphism with no input, $v : 1 \to A$, has monoidal unit domain 1; a morphism with deleted (empty) output has type $u : A \to 1$. We include the domain and codomain functions as a part of the signature i.e. $\Sigma = (\Sigma_0, \Sigma_1, \mathrm{dom}, \mathrm{cod})$. To save space, we will alternatively denote the information in $\mathrm{dom}, \mathrm{cod}$ through the more traditional type specifications of the morphisms, e.g. for the category containing the morphism $\mathrm{dom}\left(u\right) = A$, $\mathrm{dom}\left(v\right) = AB$, $\mathrm{cod}\left(u\right) = 1$ and $\mathrm{cod}\left(v\right) = V$ we write $\Sigma = \left(\left\{A, B\right\}, \left\{u, v\right\}, \left\{u : A \to 1, v : AB \to V\right\}\right)$. An (affine, symmetric) monoidal signature Σ determines a symmetric monoidal category whose morphisms are *generated* by combining morphism variables using sequential composition $\cdot$ and commutative monoidal product $\otimes$ along with identities $id_A : A \to A$, copies $\Delta_A : A \to A^2$ and deletions $del_A : A \to 1$ for every object in $A \in \Sigma_0$. Every expression formed this way is itself a morphism in the category specified by Σ.

However, the signature also determines a *specific* causal model which is the one of practical interest, in the following sense. Construct an expression in which all causal module morphism variable terms appear once, composed using sequential composition $\circ$ and monoidal product $\otimes$, inserting only the necessary identities and copies in order to ensure that the domains and codomains of these morphisms are matched. We call this expression the *maximal model* (quotiented by the



(a) (b)

Figure 1: Example string diagram representations of maximal causal models for monoidal signatures: (a) signature $\Sigma = \left(\{X_1, X_2, X_3\}, \{x_1, x_2, x_3\}, \left\{x_1 : 1 \to X_1^2, x_2 : X_1 \to X_2, x_3 : X_1X_2 \to X_3\right\}\right)$ with explicit internal causal mechanisms; (b) exterior signature $\mathrm{Ext}(\Sigma) = (\{X_3\}, \{q\}, \{q : 1 \to X_3\})$ hiding the internal causal mechanisms in (a).

identities, copies and deletions) determined by the signature. The maximal model is a (possibly composite) causal morphism in the monoidal category, with its own domain and codomain. For example, the signature

$$\Sigma = \left(\{X_1, X_2, X_3\}, \{x_1, x_2, x_3\}, \left\{x_1 : 1 \to X_1^2, x_2 : X_1 \to X_2, x_3 : X_1X_2 \to X_3\right\}\right), \quad (1)$$

has the maximal causal model expression $q = x_3 \cdot (x_2 \otimes id_{X_1}) \cdot x_1$ with type

$$\mathrm{dom}(q) = 1 \overset{x_1}{\to} X_1X_1 \overset{x_2 \otimes id_{X_1}}{\longrightarrow} X_2X_1 \overset{x_3}{\to} X_3 = \mathrm{cod}(q). \quad (2)$$

While the expression q is itself a morphism in the category generated by Σ, below we will find it useful to isolate this as a separate causal morphism which generates a category with composite signature $\Sigma' = (\{X_3\}, \{q\}, \{q : 1 \to X_3\})$. This signature hides the *internal* details of how q was obtained; the detailed signature can always be reconstructed from information in the expression from which it is formed. We call the signature $\Sigma' = \mathrm{Ext}(\Sigma)$ the *exterior* signature Σ.

Because of the modular nature of causal ADMGs, in the corresponding monoidal signature there is a one-one mapping between domain object labels and morphism labels.

Definition 2 *(Causal module labels). For a given monoidal signature Σ representing an ADMG with object and morphism labels Σ_0, Σ_1, the function $\mathrm{Module} : \Sigma_0 \to \Sigma_1$, returns the morphism label associated with a causal module.*

As an example, in the case of the exterior signature above, $\mathrm{Module}(X_3) = q$ because in $\mathrm{Ext}(\Sigma)$, $\mathrm{cod}(q) = X_3$.

A convenient graphical notational device for representing such (maximal) causal models are *string diagrams* (Selinger, 2011) which have a rigorous algebraic meaning which coincides with that of monoidal categories, see Figure 1. We recommend Coecke and Kissinger (2017) as a background to string diagrams for those unfamiliar with the concept.

In the next section we build a bridge between latent CBNs represented as ADMGs, and their representation as (affine) SMC signatures specifying a causal model.

2.2. Monoidal category signatures from ADMGs

The monoidal category signature $\Sigma^{\mathcal{G}}$ generated by the ADMG $\mathcal{G}$ is $\Sigma^{\mathcal{G}} = \left(\Sigma_0^{\mathcal{G}}, \Sigma_1^{\mathcal{G}}, \mathrm{dom}, \mathrm{cod}\right)$, where $\Sigma_0^{\mathcal{G}} = \mathbf{V}^{\mathcal{G}}$, $\Sigma_1^{\mathcal{G}} = \left\{\mathrm{Module}\left(V'\right) | V' \in \mathbf{V}^{\mathcal{G}}\right\}$ and for each node (signature object) $V \in \mathbf{V}^{\mathcal{G}}$ with corresponding causal module (signature morphism),

$$\begin{aligned} \mathrm{dom}\left(\mathrm{Module}\left(V\right)\right) &= \bigotimes_{V' \in \mathrm{pa}_{\mathcal{G}}(V)} V' \\ \mathrm{cod}\left(\mathrm{Module}\left(V\right)\right) &= V^{|\mathrm{ch}_{\mathcal{G}}(V)|+1}. \end{aligned} \tag{3}$$

Our formulation of syntactic causal identification can only be applied where the domain and codomain of the causal module $v = \mathrm{Module}\left(V\right)$ is explicit in a signature. Therefore, in practice, syntactic causal identification requires the undirectional part of the ADMG to be available in *chain-factored* form. This is obtained using any topological ordering of the ADMG as $\Sigma^{\mathcal{F}} = \left(\Sigma_0^{\mathcal{G}}, \Sigma_1^{\mathcal{G}}, \mathrm{dom}, \mathrm{cod}\right)$ where,

$$\begin{aligned} \mathrm{dom}\left(\mathrm{Module}\left(V\right)\right) &= \bigotimes_{V' \in \mathrm{pre}_{\mathcal{G}}(V)} V' \\ \mathrm{cod}\left(\mathrm{Module}\left(V\right)\right) &= V^{|\mathrm{succ}_{\mathcal{G}}(V)|+1}. \end{aligned} \tag{4}$$

It is useful to have access to parent and child information from a signature. The set of parent modules of a causal module $v = \mathrm{Module}\left(V\right)$ are given by $\mathrm{pa}_{\Sigma}\left(V\right) = \left\{v' \in \Sigma_1 : \mathrm{cod}\left(v'\right) \cap \mathrm{dom}\left(\mathrm{Module}\left(V\right)\right) \neq \emptyset\right\}$ and child modules by $\mathrm{ch}_{\Sigma}\left(V\right) = \left\{v' \in \Sigma_1 : \mathrm{dom}\left(v'\right) \cap \mathrm{cod}\left(\mathrm{Module}\left(V\right)\right) \neq \emptyset\right\}$.

2.3. Syntactic causal identification

Here we present our main result. Richardson et al. (2012, Theorem. 49) is a re-formulation of the *ID algorithm* (Shipster and Pearl, 2008) for causal identification in general causal models with latent variables, in terms of fixing operations on conditional ADMGs (CADMGs). In this section we provide a purely syntactic description of the same algorithm which uses only the structural information in the ADMG.

Theorem 3 *In the ADMG $\mathcal{G}$, consider the set of* cause $\mathbf{A} \subset \mathbf{V}^{\mathcal{G}}$ *and effect variables* $\mathbf{Y} \subset \mathbf{V}^{\mathcal{G}}$, *where $\mathbf{A}$ and $\mathbf{Y}$ do not intersect. Now consider the set of variables $\mathbf{Y}^{\star} = \mathrm{an}_{\mathcal{G}_{\mathbf{V}^{\mathcal{G}} \setminus \mathbf{A}}}\left(\mathbf{Y}\right)$ and $\mathbf{D}^{\star}$ the set of districts of the subgraph $\mathcal{G}_{\mathbf{Y}^{\star}}$. The signature of the* syntactic causal effect, $\Sigma^{\mathcal{G}}_{\mathbf{Y}|\mathrm{do}(\mathbf{A})}$, *of $\mathbf{A}$ on $\mathbf{Y}$ is* identifiable *if, for every district $\mathbf{D}' \in \mathbf{D}^{\star}$ the set of nodes $\mathbf{V}^{\mathcal{G}} \setminus \mathbf{D}'$ is a valid fixing sequence. If identifiable, this causal effect is given by the following composite signature manipulation,*

$$\Sigma^{\mathcal{G}}_{\mathbf{Y}|\mathrm{do}(\mathbf{A})} = \mathrm{Hide}_{\mathbf{Y}^{\star} \setminus \mathbf{Y}}\left(\bigcup_{\mathbf{D}' \in \mathbf{D}^{\star}} \mathrm{Simple}\left(\mathrm{Fixseq}_{\mathbf{V}^{\mathcal{G}} \setminus \mathbf{D}'}\left(\Sigma^{\mathcal{F}}\right)\right)\right). \tag{5}$$

2.4. Signature manipulations

This section details the manipulations required to implement the syntactic identification algorithm of (5). The syntactic analogue of marginalization is captured in the following definition.

Definition 4 *(Marginalization/hiding). Given a symmetric monoidal signature* $\Sigma^{\mathcal{F}} = (\Sigma_0^{;}\Sigma_1, \mathrm{dom}, \mathrm{cod})$*, and an object variable* $V \in \Sigma_0$*. The function* $\mathrm{Hide}_V\left(\Sigma^{\mathcal{F}}\right) = \left(\Sigma_0^{\mathcal{F}}, \Sigma_1^{\mathcal{F}}, \mathrm{dom}, \mathrm{cod}'\right)$*, manipulates the signature in the following way*

$$\mathrm{cod}'\left(\mathrm{Module}\left(V\right)\right) = V^{|\mathrm{succ}_{\mathcal{G}}(V)|} \tag{6}$$

and $\mathrm{cod}' = \mathrm{cod}$ *otherwise. For the set* $\mathbf{W} = \{V_1, \ldots, V_k\} \subset \mathbf{V}^{\mathcal{G}}$*, we extend this function to the composite* $\mathrm{Hide}_{\mathbf{W}} = \mathrm{Hide}_{V_1} \circ \cdots \circ \mathrm{Hide}_{V_k}$.

Pearl's causal interventions in ADMGs requires deleting parent edges (J.Pearl, 2009; Richardson et al., 2012; Bareinboim et al., 2022). For affine SMC signatures, this entails replacing causal modules with identity/copy morphisms and deleting any wires connected to that module. This is captured in the following definition.

Definition 5 *(Causal control). Given a symmetric monoidal signature* $\Sigma = (\Sigma_0, \Sigma_1, \mathrm{dom}, \mathrm{cod})$*, and an object* $V \in \Sigma_0$*. The function* $\mathrm{Control}_V\left(\Sigma\right) = \left(\Sigma_0, \Sigma_1, \mathrm{dom}', \mathrm{cod}'\right)$*, manipulates the signature in the following way*

$$\mathrm{dom}'\left(\mathrm{Module}\left(V\right)\right) = V, \tag{7}$$

which replaces the module with a copied identity morphism, and for all other $v' \in \Sigma_1$ *such that* $v' \neq \mathrm{Module}\left(V\right)$*,*

$$\begin{aligned} \mathrm{dom}'\left(v'\right) &= \mathrm{dom}\left(v'\right) \\ \mathrm{cod}'\left(v'\right) &= \mathrm{cod}\left(v'\right) \backslash \mathrm{dom}\left(\mathrm{Module}\left(V\right)\right), \end{aligned} \tag{8}$$

where the second line deletes incoming wires using the multiset difference*. This operation is extended to controlling a set as with marginalization above.*

Definition 6 *(Causal fixing). Given a symmetric monoidal signature* $\Sigma = (\Sigma_0, \Sigma_1, \mathrm{dom}, \mathrm{cod})$*, and an object* $V \in \Sigma_0$*. The syntactic fixing operation is the composition of marginalization and control functions*

$$\mathrm{Fix}_V = \mathrm{Control}_V \circ \mathrm{Hide}_V, \tag{9}$$

The above definition is exactly the syntactic analogue of fixing in ADMGs (Richardson et al., 2012). To ensure identifiability, fixing can only be applied to objects which are fixable relative to some signature, $\Sigma^{\mathcal{G}}$ derived from an ADMG $\mathcal{G}$. Given a set of objects $\mathbf{W} = \{V_1, \ldots, V_k\}$ to fix, we need to determine a valid sequence of fixing operations, $\mathrm{Fixseq}_{\mathbf{W}}$, for this set (Richardson et al., 2012).

To initialize the recursion, set $\mathbf{W}' = \mathbf{W}$, initialize $\Sigma = \Sigma^{\mathcal{G}}$, initialize the fixing sequence operation $\mathrm{Fixseq}_{\mathbf{W}} = id$ (the identity operation). The recursion step is as follows: choose any $V \in \mathbf{W}'$ such that V is fixable in Σ. If there exists no such V then the sequence $\mathbf{W}$ cannot have a valid fixing sequence and the recursion terminates. Otherwise, if $\mathrm{ch}_{\Sigma}\left(V\right) = \emptyset$ then update $\mathrm{Fixseq}_{\mathbf{W}} \mapsto \mathrm{Hide}_V \circ \mathrm{Fixseq}_{\mathbf{W}}$, otherwise update $\mathrm{Fixseq}_{\mathbf{W}} \mapsto \mathrm{Fix}_V \circ \mathrm{Fixseq}_{\mathbf{W}}$ instead. Now, apply this operation to obtain the updated signature $\Sigma \mapsto \mathrm{Fixseq}_{\mathbf{W}}\left(\Sigma^{\mathcal{G}}\right)$ and delete V from the fixing set, $\mathbf{W}' \mapsto \mathbf{W}' \backslash V$. If all objects to fix have been exhausted, i.e. $\mathbf{W}' = \emptyset$, then the recursion terminates with the fixing operation sequence $\mathrm{Fixseq}_{\mathbf{W}}$, otherwise the process returns to the recursion step.

Manipulating a signature can lead to a causal module $v = \mathrm{Module}(V)$ being equivalent to the identity morphism i.e. where $\mathrm{dom}(\mathrm{Module}(V)) = \mathrm{cod}(\mathrm{Module}(V)) = V$. These can be deleted from the signature, simplifying the causal model it specifies. Writing the set of modules which are not equivalent to a identity as $\mathbf{W} = \left\{V \in \Sigma_0^{\mathcal{G}} | \left(\mathrm{dom}(\mathrm{Module}(V)) \neq V\right) \vee \left(\mathrm{cod}(\mathrm{Module}(V)) \neq V\right)\right\}$, then $\mathrm{DeleteId}(\Sigma) = \left(\mathbf{W}, \{\mathrm{Module}(V') | V' \in \mathbf{W}\}, \mathrm{dom}, \mathrm{cod}\right)$ is the simplified signature. Furthermore, signature manipulation can lead to a causal module having no downstream effects. Such modules can also be deleted from the signature, further simplifying the causal model.

Definition 7 *(Signature simplification). Consider the set* $\mathbf{W} = \{V \in \Sigma_0 | \mathrm{cod}(\mathrm{Module}(V)) \neq 1\}$ *of objects whose causal modules do not have marginalized outputs, then the simplified signature can be written,*

$$\mathrm{Simplify}(\Sigma) = \left(\mathbf{W}, \left\{\mathrm{Module}\left(V'\right) | V' \in \mathbf{W}\right\}, \mathrm{dom}, \mathrm{cod}'\right) \tag{10}$$

with modified codomain $\mathrm{cod}'(\mathrm{Module}(V')) = \mathrm{cod}(\mathrm{Module}(V')) \backslash \mathrm{dom}(\mathrm{Module}(W'))$ *for all* $V \in \Sigma_0 \backslash \mathbf{W}$ *and* $W' \in \mathbf{W}$.

Deleting a module might lead to other causal modules having no downstream effects, therefore it is necessary to iterate (10) until a fixed point signature is reached. Formally, starting at $\Sigma^0 = \Sigma$, the iteration $\Sigma^{n+1} = \mathrm{Simplify}(\Sigma^n)$ is repeated until some N is obtained such that $\mathrm{Simplify}\left(\Sigma^N\right) = \Sigma^N$, whereupon we use $\mathrm{Simple}(\Sigma) = \mathrm{DeleteId}\left(\Sigma^N\right)$ as the fully simplified signature.

Syntactic causal identification requires combining exterior morphisms of manipulated signatures.

Definition 8 *(Combining exterior signatures). For signatures* Σ^1 *and* Σ^2*, their combination is,*

$$\Sigma = \mathrm{Ext}\left(\Sigma^1\right) \cup \mathrm{Ext}\left(\Sigma^2\right) = \left(\Sigma_0^1 \cup \Sigma_0^2, \Sigma_1^1 \cup \Sigma_1^2, \mathrm{dom}^1 \cup \mathrm{dom}^2, \mathrm{cod}'\right) \tag{11}$$

where $\mathrm{Ext}\left(\Sigma^1\right) = \left(\Sigma_0^1, \Sigma_1^1, \mathrm{dom}^1, \mathrm{cod}^1\right)$, $\mathrm{Ext}\left(\Sigma^2\right) = \left(\Sigma_0^2, \Sigma_1^2, \mathrm{dom}^2, \mathrm{cod}^2\right)$ *are the signatures of the exteriors of* Σ^1 *and* Σ^2 *respectively, and their combined codomain is,*

$$\mathrm{cod}'(\mathrm{Module}(V)) = V^{|\mathrm{ch}_\Sigma(V)|+1} \tag{12}$$

for all $V \in \Sigma_0^1 \cup \Sigma_0^2$ *for which a causal module is assigned through* cod^1 *and* cod^2.

3. Applications

3.1. Back-door adjustment: simple case

As first application, we show how to derive a purely syntactic account of back-door adjustment. Consider a simple, fully observed model with one confounder. The ADMG is defined by $\mathbf{V}^{\mathcal{G}} = \{X, Y, U\}$ and edges $\mathbf{E}^{\mathcal{G}} = \{X \to Y, U \to X, U \to Y\}$, and we want the interventional signature $\Sigma^{\mathcal{G}}_{Y|\mathrm{do}(X)}$. In this situation, the model is equivalent to a fully chain-factored graph so we can directly fix X in the signature without the need for explicit district decomposition. The corresponding signature $\Sigma^{\mathcal{G}} = \left(\{X,Y,U\},\{x,y,u\},\left\{u:1\to U^3, x:U\to X^2, y:XU\to Y\right\}\right)$ so that, on application of fixing, we obtain the manipulated signature,

$$\begin{aligned}
\Sigma &= \mathrm{Fix}_X\left(\{X,Y,U\},\{x,y,u\},\left\{u:1\to U^3, x:U\to X^2, y:XU\to Y\right\}\right)\\
&= \left(\mathrm{Control}_X\circ\mathrm{Hide}_X\right)\left(\{X,Y,U\},\{x,y,u\},\left\{u:1\to U^3, x:U\to X^2, y:XU\to Y\right\}\right)\\
&= \left(\mathrm{Control}_X\right)\left(\{X,Y,U\},\{x,y,u\},\left\{u:1\to U^3, x:U\to X, y:XU\to Y\right\}\right)\\
&= \left(\{X,Y,U\},\{x,y,u\},\left\{u:1\to U^2, x:X\to X, y:XU\to Y\right\}\right),
\end{aligned} \tag{13}$$

which, when simplified becomes,

$$\begin{aligned}
\mathrm{Simple}\left(\Sigma\right) &= \left(\{X,Y,U\},\{y,u\},\left\{u:1\to U^2, y:XU\to Y\right\}\right)\\
&= \Sigma^{\mathcal{G}}_{YU|\mathrm{do}(X)},
\end{aligned} \tag{14}$$

and then finally marginalizing out U we obtain the desired interventional signature,

$$\begin{aligned}
\Sigma^{\mathcal{G}}_{Y|\mathrm{do}(X)} &= \mathrm{Hide}_U\left(\{X,Y,U\},\{y,u\},\left\{u:1\to U^2, y:XU\to Y\right\}\right)\\
&= \left(\{X,Y,U\},\{y,u\},\{u:1\to U, y:XU\to Y\}\right).
\end{aligned} \tag{15}$$

This is the purely syntactic categorical analogue of the well-known *back-door adjustment formula* in this simple case. When this signature is interpreted as a Markov category (Fritz, 2020) with continuous sample spaces $X\mapsto\Omega_X$, $U\mapsto\Omega_U$ and $Y\mapsto\Omega_Y$ with composition $\cdot$ corresponding to the *Chapman-Kolmogorov equation* (Little, 2019; Jacobs et al., 2021) and the causal morphisms are *conditional distributions* $u\mapsto p\left(U\right)$ and $y\mapsto p\left(Y|X,U\right)$, then (15) is the interventional distribution,

$$p\left(Y=y|\mathrm{do}\left(X=x\right)\right) = \int_{\Omega_U} p\left(Y=y|X=x,U=u\right)p\left(U=u\right)du. \tag{16}$$

Another interesting interpretation is in terms of the *min-plus semifield Markov category* (Cakiqi and Little, 2022) for which

$$q\left(y|\mathrm{do}\left(x\right)\right) = \min_{u\in U}\left[q\left(y|x,u\right)+q\left(u\right)\right], \tag{17}$$

where $u\mapsto q\left(u\right)$ is an *inferential bias* and $y\mapsto q\left(y|x,u\right)$ a *clique potential* widely encountered in machine learning. For these functions to be biases/potentials, they must be *normalizable* in the min-plus semifield, i.e. $\min_{u\in U} q\left(u\right)=0$ and $\min_{y\in Y} q\left(y|x,u\right)=0$ (Little, 2019).

3.2. Front-door adjustment

The front-door graph (J.Pearl, 2009) is an ADMG $\mathcal{G}$ given by $\mathbf{V}^{\mathcal{G}}=\{X,Y,Z\}$ with edges $\mathbf{E}^{\mathcal{G}}=\{X\to Z\to Y, X\leftrightarrow Y\}$. We want the pure syntactic interventional signature $\Sigma^{\mathcal{G}}_{Y|\mathrm{do}(X)}$, which represents the causal effect on $\mathbf{Y}=\{Y\}$ of $\mathbf{A}=\{X\}$.
Here, $\mathbf{V}^{\mathcal{G}}\backslash\mathbf{A}=\{Y,Z\}$ so that $\mathbf{Y}^{\star}=\mathrm{an}_{\mathcal{G}_{\{Y,Z\}}}\left(\{Y\}\right)=\{Y,Z\}$ with corresponding districts $\mathbf{D}^{\star}=\left\{\{Y\},\{Z\}\right\}$ of subgraph $\mathcal{G}_{\{Y,Z\}}$ (See Figure 2).The monoidal signature of $\mathcal{G}$ is, $\Sigma^{\mathcal{G}}=$

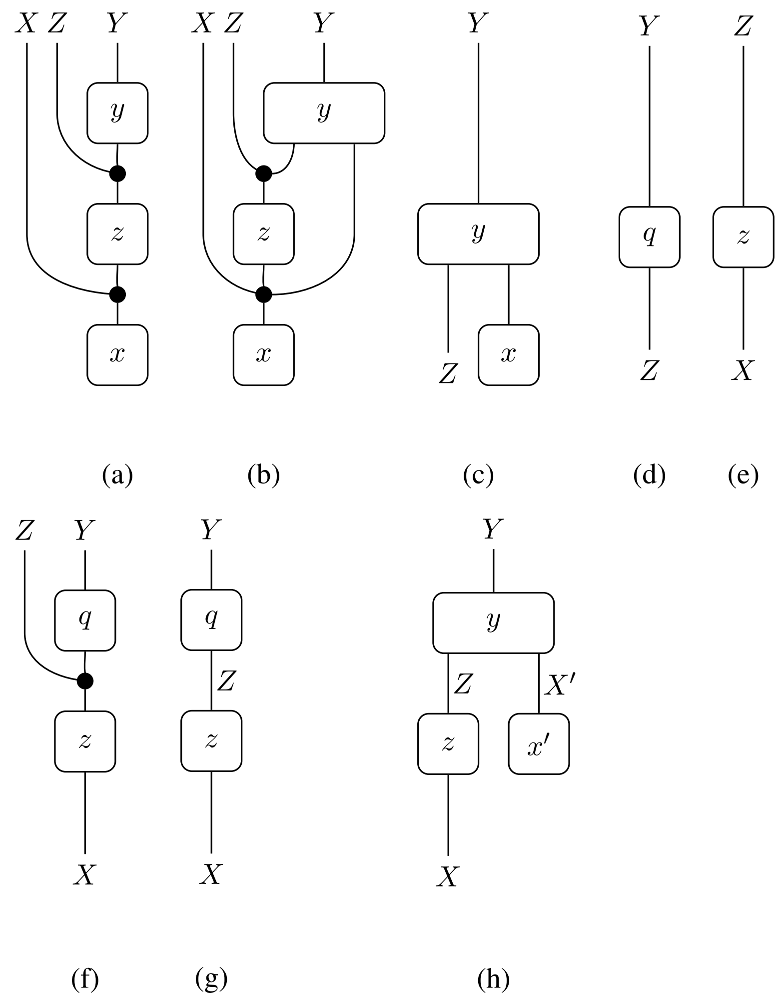


Figure 2: String diagram representations of maximal causal models for monoidal signatures obtained during the derivation of the purely syntactic front-door adjustment interventional signature $\Sigma^{\mathcal{G}}_{Y|\mathrm{do}(X)}$. Signatures are as follows: (a) $\Sigma^{\mathcal{G}}$ (front-door ADMG, observable variables only); (b) $\Sigma^{\mathcal{F}}$ (chain-factored observable ADMG); (c) $\Sigma^1 = \left(\mathrm{Simple} \circ \mathrm{Fixseq}_{\{X,Z\}}\right)\left(\Sigma^{\mathcal{F}}\right)$ (fixed district $\{Y\}$); (d) $\mathrm{Ext}\left(\Sigma^1\right)$ (exterior signature of Σ^1); (e) $\Sigma^2 = \left(\mathrm{Simple} \circ \mathrm{Fixseq}_{\{X,Y\}}\right)\left(\Sigma^{\mathcal{F}}\right)$ (fixed district $\{Z\}$); (f) $\Sigma^{\mathcal{G}}_{ZY|\mathrm{do}(X)} = \mathrm{Ext}\left(\Sigma^1\right) \cup \mathrm{Ext}\left(\Sigma^2\right)$ (interventional signature); (g) $\Sigma^{\mathcal{G}}_{Y|\mathrm{do}(X)}$ (marginalized interventional signature) and (h) $\Sigma^{\mathcal{G}}_{Y|\mathrm{do}(X)}$ with internal mechanisms made explicit.

$\left(\{X, Y, Z\}, \{x, y, z\}, \left\{x : 1 \to X^2, z : X \to Z^2, y : Z \to Y\right\}\right)$, with corresponding chain-factored signature, $\Sigma^{\mathcal{F}} = \left(\{X, Y, Z\}, \{x, y, z\}, \left\{x : 1 \to X^3, z : X \to Z^2, y : XZ \to Y\right\}\right)$.

For the district $\mathbf{D}' = \{Y\}$, the fixing set is $\mathbf{W} = \{X, Z\}$, for which we obtain the sequence $\mathrm{Fixseq}_{\{X,Z\}} = \mathrm{Hide}_X \circ \mathrm{Fix}_Z$ applied to $\Sigma^{\mathcal{F}}$,

$$\begin{aligned}\Sigma^1 &= (\mathrm{Simple} \circ \mathrm{Hide}_X \circ \mathrm{Fix}_Z)\left(\Sigma^{\mathcal{F}}\right)\\ &= (\mathrm{Simple} \circ \mathrm{Hide}_X \circ \mathrm{Fix}_Z)\left(\{X,Y,Z\},\{x,y,z\},\left\{x:1\to X^3, z:X\to Z^2, y:XZ\to Y\right\}\right)\\ &= (\mathrm{Simple} \circ \mathrm{Hide}_X)\left(\{X,Y,Z\},\{x,y,z\},\left\{x:1\to X^2, z:Z\to Z, y:XZ\to Y\right\}\right)\\ &= \mathrm{Simple}\left(\{X,Y,Z\},\{x,y,z\},\{x:1\to X, z:Z\to Z, y:XZ\to Y\}\right)\\ &= \left(\{X,Y,Z\},\{x,y\},\{x:1\to X, y:XZ\to Y\}\right),\end{aligned} \tag{18}$$

so that $\mathrm{Ext}\left(\Sigma^1\right) = (\{Y,Z\},\{q\},\{q:Z\to Y\})$ with $q = y\cdot(id_Z\otimes x)$ and $\mathrm{Module}\,(Y) = q$.

For the district $\mathbf{D'} = \{Z\}$, the fixing set is $\mathbf{W} = \{X,Y\}$, for which we obtain the sequence $\mathrm{Fixseq}_{\{X,Y\}} = \mathrm{Fix}_X \circ \mathrm{Fix}_Y$,

$$\begin{aligned}\Sigma^2 &= (\mathrm{Simple} \circ \mathrm{Fix}_X \circ \mathrm{Fix}_Y)\left(\Sigma^{\mathcal{F}}\right)\\ &= (\mathrm{Simple} \circ \mathrm{Fix}_X \circ \mathrm{Fix}_Y)\left(\{X,Y,Z\},\{x,y,z\},\left\{x:1\to X^3, z:X\to Z^2, y:XZ\to Y\right\}\right)\\ &= (\mathrm{Simple} \circ \mathrm{Fix}_X)\left(\{X,Y,Z\},\{x,y,z\},\left\{x:1\to X^2, z:X\to Z, y:Y\to 1\right\}\right)\\ &= \mathrm{Simple}\left(\{X,Y,Z\},\{x,y,z\},\{x:X\to X, z:X\to Z, y:Y\to 1\}\right)\\ &= \left(\{X,Z\},\{z\},\{z:X\to Z\}\right)\\ &= \mathrm{Ext}\left(\Sigma^2\right),\end{aligned} \tag{19}$$

with $\mathrm{Module}\,(Z) = z$, where this signature is already exterior because z is not composite. Combining these two exterior signatures, we obtain,

$$\Sigma^{\mathcal{G}}_{ZY|\mathrm{do}(X)} = \mathrm{Ext}\left(\Sigma^1\right) \cup \mathrm{Ext}\left(\Sigma^2\right) = \left(\{X,Y,Z\},\{q,z\},\left\{z:X\to Z^2, q:Z\to Y\right\}\right), \tag{20}$$

and since $\mathbf{Y}^{\star}\backslash\mathbf{Y} = \{Y,Z\}\setminus\{Y\} = \{Z\}$, we obtain the desired interventional distribution by marginalization,

$$\Sigma^{\mathcal{G}}_{Y|\mathrm{do}(X)} = \mathrm{Hide}_{\{Z\}}\left(\Sigma^{\mathcal{G}}_{ZY|\mathrm{do}(X)}\right) = \left(\{X,Y,Z\},\{q,z\},\{z:X\to Z, q:Z\to Y\}\right). \tag{21}$$

In practice, it may be useful to expose the interior of q and to do this we will need to relabel $x \to x'$(and correspondingly, $X \mapsto X'$) inside q to avoid a naming clash with the interventional input X,

$$\Sigma^{\mathcal{G}}_{Y|\mathrm{do}(X)} = \left(\left\{X,X',Y,Z\right\},\{x,y,z\},\left\{x':1\to X', z:X\to Z, y:X'Z\to Y\right\}\right). \tag{22}$$

As with the back-door model, this is the purely syntactic categorical analogue of the *front-door adjustment formula*. As an example interpretation, consider the Markov category with discrete sample spaces $X' \mapsto \Omega_X$, $Z \mapsto \Omega_Y$, $Y \mapsto \Omega_Y$ and with conditional distributions $x' \mapsto p\,(X')$, $z \mapsto p\,(Z|X)$ and $y \mapsto p\,(Y|X,Z)$, then (22) is the familiar discrete interventional distribution (J.Pearl, 2009),

$$p\,(Y=y|\mathrm{do}\,(X=x)) = \sum_{z\in\Omega_Z} p\,(Z=z|X=x) \sum_{x'\in\Omega_X} p\left(Y=y|X'=x', Z=z\right) p\left(X'=x'\right). \tag{23}$$

Another useful interpretation are *deterministic causal models* in the SMC of sets and functions, in which composition $\cdot$ is ordinary function composition and $\otimes$ is the pairing (bifunctor), with the identity 1 corresponding to the empty pair $(\,)$. Then, interpreting X' with the set of possible values for the constant $f_{X'} : 1 \to X'$, the functions $f_Z : X \to Z$ and $f_Y : X'Z \to Y$, the front-door interventional model corresponding to the signature $\Sigma^{\mathcal{G}}_{Y|\mathrm{do}(X)}$ is

$$f_{Y|\mathrm{do}(X)}\left(x\right) = \pi_2\left(f_Z\left(x\right), f_Y\left(f_{X'}\left(\,\right), f_Z\left(x\right)\right)\right) = f_Y\left(f_{X'}\left(\,\right), f_Z\left(x\right)\right), \tag{24}$$

where π_2 is projection onto the second item of a pair.

3.3. A more complex example

Richardson et al. (2012, Example. 50) describe an application of their fixing theorem to a more complex causal model with four variables and a single bidirected edge whose latent projection ADMG $\mathcal{G}$ is given by $\mathbf{V}^{\mathcal{G}} = \{X_1, X_2, X_3, X_4\}$ and by edge set

$$\mathbf{E}^{\mathcal{G}} = \{X_3 \leftarrow X_1 \rightarrow X_2, X_2 \rightarrow X_3, X_3 \rightarrow X_4, X_2 \leftrightarrow X_4\}. \tag{25}$$

In this example they identify the interventional distribution $p\left(X_4|\mathrm{do}\left(X_2\right)\right)$. To illustrate our syntactic fixing algorithm for this example, we want the monoidal signature of the pure syntactic causal effect on $\mathbf{Y} = \{X_4\}$ of $\mathbf{A} = \{X_2\}$, i.e. $\Sigma^{\mathcal{G}}_{X_4|\mathrm{do}(X_2)}$. Here, $\mathbf{V}^{\mathcal{G}}\backslash\mathbf{A} = \{X_1, X_3, X_4\}$ so that $\mathbf{Y}^{\star} = \mathrm{an}_{\mathcal{G}_{\mathbf{V}^{\mathcal{G}}\backslash\mathbf{A}}}\left(\{X_4\}\right) = \{X_1, X_3, X_4\}$ with corresponding subgraph districts $\mathbf{D}^{\star} = \mathcal{G}_{\mathbf{Y}^{\star}} = \{\{X_1\}, \{X_3\}, \{X_4\}\}$. The monoidal signature of $\mathcal{G}$ is,

$$\begin{aligned}\Sigma^{\mathcal{G}} = (\{X_1, X_2, X_3, X_4\}, \{x_1, x_2, x_3, x_4\}, \\ \left\{x_1 : 1 \to X_1^3, x_2 : X_1 \to X_2^2, x_3 : X_1X_2 \to X_3^2, x_4 : X_1X_2X_3 \to X_4\right\})\end{aligned} \tag{26}$$

with corresponding chain-factored signature,

$$\begin{aligned}\Sigma^{\mathcal{F}} = (\{X_1, X_2, X_3, X_4\}, \{x_1, x_2, x_3, x_4\}, \\ \left\{x_1 : 1 \to X_1^4, x_2 : X_1 \to X_2^3, x_3 : X_1X_2 \to X_3^2, x_4 : X_1X_2X_3 \to X_4\right\})\end{aligned} \tag{27}$$

For the district $\mathbf{D}' = \{X_1\}$, the fixing set is $\mathbf{W} = \{X_2, X_3, X_4\}$, for which we obtain,

$$\begin{aligned}\Sigma^1 &= \left(\mathrm{Simple} \circ \mathrm{Fix}_{X_2} \circ \mathrm{Fix}_{X_3} \circ \mathrm{Fix}_{X_4}\right)\left(\Sigma^{\mathcal{F}}\right) \\ &= \left(\{X_1\}, \{x_1\}, \{x_1 : 1 \to X_1\}\right) \\ &= \mathrm{Ext}\left(\Sigma^1\right), \mathrm{Module}\left(X_1\right) = x_1.\end{aligned} \tag{28}$$

For the district $\mathbf{D}' = \{X_3\}$, the fixing set $\mathbf{W} = \{X_1, X_2, X_4\}$, leading to,

$$\begin{aligned}\Sigma^2 &= \left(\mathrm{Simple} \circ \mathrm{Fix}_{X_1} \circ \mathrm{Fix}_{X_2} \circ \mathrm{Fix}_{X_4}\right)\left(\Sigma^{\mathcal{F}}\right) \\ &= \left(\{X_1, X_2, X_3\}, \{x_3\}, \{x_3 : X_1X_2 \to X_3\}\right) \\ &= \mathrm{Ext}\left(\Sigma^2\right), \mathrm{Module}\left(X_3\right) = x_3.\end{aligned} \tag{29}$$

Finally, for the district $\mathbf{D}' = \{X_4\}$, the fixing set $\mathbf{W} = \{X_1, X_2, X_3\}$ leads to,

$$\begin{aligned}\Sigma^3 &= \left(\text{Simple} \circ \text{Hide}_{X_2} \circ \text{Fix}_{X_3} \circ \text{Fix}_{X_1}\right)\left(\Sigma^{\mathcal{F}}\right) \\ &= \left(\{X_1, X_2, X_3, X_4\}, \{x_4, x_2\}, \left\{x_1 : X_1 \to X_1^2, x_2 : X_1 \to X_2, x_4 : X_1X_2X_3 \to X_4\right\}\right),\end{aligned} \tag{30}$$

with exterior signature

$$\begin{aligned}\text{Ext}\left(\Sigma^3\right) = \left(\{X_1, X_3, X_4\}, \{q\}, \{q : X_1X_3 \to X_4\}\right), \text{Module}\left(X_4\right) = q \\ q = x_4 \cdot \left(x_2 \otimes id_{X_1} \otimes id_{X_3}\right) \cdot \left(x_1 \otimes id_{X_3}\right),\end{aligned} \tag{31}$$

where the maximal exterior morphism q has type,

$$\text{dom}\left(q\right) = X_1X_3 \to X_1X_1X_3 \to X_2 \to X_1X_3 \to X_4 = \text{cod}\left(q\right). \tag{32}$$

Combining the exterior signatures obtains

$$\begin{aligned}\Sigma^{\mathcal{G}}_{X_1X_3X_4|\text{do}(X_2)} &= \text{Ext}\left(\Sigma^1\right) \cup \text{Ext}\left(\Sigma^2\right) \cup \text{Ext}\left(\Sigma^3\right) \\ &= \left(\{X_1, X_2, X_3, X_4\}, \right. \\ &\left.\{x_1, x_3, q\}, \left\{x_1 : 1 \to X_1^3, x_3 : X_1X_2 \to X_3^2, q : X_1X_3 \to X_4\right\}\right),\end{aligned} \tag{33}$$

from which we can compute the desired, syntactic interventional signature by marginalization,

$$\begin{aligned}\Sigma^{\mathcal{G}}_{X_4|\text{do}(X_2)} &= \text{Hide}_{\{X_1, X_3\}}\left(\Sigma^{\mathcal{G}}_{X_1X_3X_4|\text{do}(X_2)}\right) \\ &= \left(\{X_1, X_2, X_3, X_4\}, \{x_1, x_3, q\}, \left\{x_1 : 1 \to X_1^2, x_3 : X_1X_2 \to X_3, q : X_1X_3 \to X_4\right\}\right).\end{aligned} \tag{34}$$

4. Discussion

In this paper, we have shown that purely syntactic causal identification can be performed using relatively simple steps. We observe that the simplicity of this approach largely arises from the process-centric formulation of directed causal modelling and the fact that manipulations of this model this can be expressed in terms of functions of the signature of the category in which this model is represented. These steps are unambiguous and therefore easily implemented in software.

Although our approach relies on chain factorization of the observed process, we note that this is more of a mathematical convenience than a restriction. An alternative development of our approach can use *comb disintegration* in place of the fixing operator described here. This would lead to different forms of the resulting interventional signatures which are, nonetheless equivalent exterior processes. Furthermore, this approach could be extended to edge interventions, and we believe this would be quite simple to implement.

Finally, application to other, more elaborate forms of causal identification such as conditional causal effects, those arising through edge interventions, and more general forms of causal identification by combining multiple causal models, would be valuable. It would be interesting to see the extent to which the signature-based approach also simplifies the formulation of existing algorithms for these problems.

References

E. Bareinboim, J.D. Correa, D.Ibeling, and T.Icard. *On Pearl's Hierarchy and the Foundations of Causal Inference*, page 507–556. Association for Computing Machinery, New York, NY, USA, 1 edition, 2022. doi: 10.1145/3501714.3501743. URL https://doi.org/10.1145/3501714.3501743.

D. Cakiqi and M. A. Little. Non-probabilistic markov categories for causal modeling in machine learning. In *ACT 2022: Applied Category Theory*, ACT 2022, 2022. doi: 10.48550/arXiv.2307.15519.

K. Cho and B. Jacobs. Disintegration and bayesian inversion via string diagrams. *Mathematical Structures in Computer Science*, 29(7):938–971, 2019. doi: 10.1017/S0960129518000488.

Bob Coecke and Aleks Kissinger. *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning*. Cambridge University Press, 2017.

B. Fong. Causal theories: A categorical perspective on bayesian networks. Master's thesis, 2013.

T. Fritz. A synthetic approach to markov kernels, conditional independence and theorems on sufficient statistics. *Advances in Mathematics*, 370:107239, 2020. doi: 10.1016/j.aim.2020.107239. URL https://www.sciencedirect.com/science/article/pii/S0001870820302656.

T. Fritz and A. Klingler. The d-separation criterion in categorical probability. *Journal of Machine Learning Research*, 24(46):1–49, 2023. doi: 10.48550/arXiv.2207.05740. URL http://jmlr.org/papers/v24/22-0916.html.

B. Jacobs, A. Kissinger, and F.Zanasi. Causal inference via string diagram surgery: A diagrammatic approach to interventions and counterfactuals. *Mathematical Structures in Computer Science*, 31 (5):553–574, 2021. doi: 10.1017/S096012952100027X.

J.Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2009. doi: 10.1017/CBO9780511803161.

M.A. Little. *Machine Learning for Signal Processing: Data Science, Algorithms, and Computational Statistics*. Oxford University Press, 08 2019. doi: 10.1093/oso/9780198714934.001.0001.

R. Lorenz and S. Tull. Causal models in string diagrams, 2023.

E. Patterson. Knowledge representation in bicategories of relations, 2017.

T.S Richardson, J.M Robins, and I. Shpitser. Nested markov properties for acyclic directed mixed graphs. In *UAI'12: Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*, UAI'12, page 13. AUAI Press, 2012. doi: 10.5555/2789272.2886819.

E. Riehl. *Category Theory in Context*. Aurora: Dover Modern Math Originals. Dover Publications, 2017. ISBN 9780486820804. URL https://books.google.co.uk/books?id=6B9MDgAAQBAJ.

P. Selinger. *A Survey of Graphical Languages for Monoidal Categories*, pages 289–355. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. doi: 10.1007/978-3-642-12821-9_4.

I. Shipster and J. Pearl. Complete identification methods for the causal hierarchy. *J. Mach. Learn. Res.*, 9(64):1941–1979, Jun 2008. doi: 10.5555/1390681.1442797. URL http://jmlr.org/papers/v9/shpitser08a.html.

Y. Yin and J. Zhang. Markov categories, causal theories, and the do-calculus, 2022.