



**HAL**  
open science

# Identifiability of total effects from abstractions of time series causal graphs

Charles K. Assaad, Emilie Devijver, Eric Gaussier, Gregor Gössler, Anouar Meynaoui

► **To cite this version:**

Charles K. Assaad, Emilie Devijver, Eric Gaussier, Gregor Gössler, Anouar Meynaoui. Identifiability of total effects from abstractions of time series causal graphs. 2024. hal-04250602v3

**HAL Id: hal-04250602**

**<https://hal.science/hal-04250602v3>**

Preprint submitted on 20 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Identifiability of total effects from abstractions of time series causal graphs

Charles K. Assaad<sup>1</sup>, Emilie Devijver<sup>2</sup>, Eric Gaussier<sup>2</sup>, Gregor Gössler<sup>3</sup>, and Anouar Meynaoui<sup>4</sup>

<sup>1</sup>R&D Department, EasyVista, 38000 Grenoble, France

<sup>2</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, 38000 Grenoble, France

<sup>3</sup>Univ. Grenoble Alpes, INRIA, CNRS, Grenoble INP, LIG, 38000 Grenoble, France

<sup>4</sup>Univ. Rennes, CNRS, IRMAR, 35000 Rennes, France

**Abstract**—We study the problem of identifiability of the total effect of an intervention from observational time series in the situation, common in practice, where one only has access to abstractions of the true causal graph. We consider here two abstractions: the extended summary causal graph, which conflates all lagged causal relations but distinguishes between lagged and instantaneous relations, and the summary causal graph which does not give any indication about the lag between causal relations. We show that the total effect is always identifiable in extended summary causal graphs and provide necessary and sufficient conditions for identifiability in summary causal graphs. We furthermore provide adjustment sets allowing to estimate the total effect whenever it is identifiable.

## I. INTRODUCTION

Over the last century and across numerous disciplines, experimentation has emerged as a potent methodology for estimating without bias the total effect of an intervention on a specific component of a given system [Neyman et al., 1990]. However, experimentation can be costly, unethical or even unfeasible. Both researchers and experts are thus interested in estimating the effect of an intervention directly from observational data. This can be done under some assumptions when relying on a complete causal graph [Pearl et al., 2000], and typically relies on two sequential steps: identifiability and estimation [Pearl, 2019]. The identifiability step involves distinguishing cases where a solution is possible and, when it exists, providing an estimand - an expression enabling the estimation of intervention effects from observational data. The subsequent step involves the actual estimation of this estimand from the available data.

The identifiability step received much attention for non-temporal causal graphs [Pearl, 1993, 1995a, Spirtes et al., 2000, Pearl et al., 2000, Shpitser and Pearl, 2008, Maathuis and Colombo, 2013]. For abstraction of causal graphs, Perkovic [2020] derived necessary and sufficient conditions for identifying total effects in MPDAG, and Anand et al. [2023] studied C-DAG, where relationships between clusters of variables are specified, but relationships between the variables within a cluster are left unspecified.

For temporal causal graph, Blondel et al. [2016] developed the do-calculus for the full-time causal graphs (FTCGs, Figure I.1a). However, in dynamic systems, experts have difficulties in

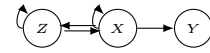
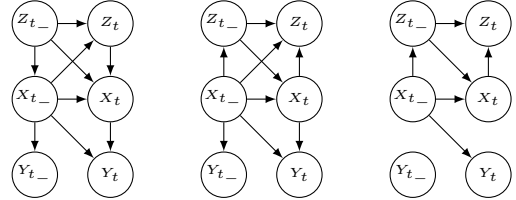
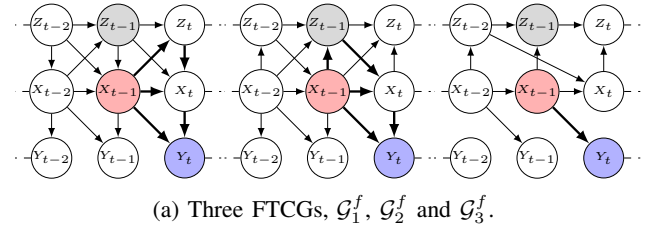


Figure I.1: Illustration: (a) three FTCGs, (b) three ESCGs derived from them, (c) the SCG which can be derived from any FTCG in (a) and any ESCG in (b). Consider  $f(y_t | do(x_{t-1}))$ , red vertex: the variable we intervene on, blue vertex: the response we are considering. Bold edges correspond to directed paths from  $X_{t-1}$  to  $Y_t$ , and gray vertices correspond to nodes with different status depending on the FTCG (see Definition 8).

building full time causal graphs, while they can usually build an abstraction of those graphs where temporal information is omitted (as ESCGs, Figure I.1b or SCGs, Figure I.1c). Assuming no instantaneous relations, Eichler and Didelez [2007] demonstrated that the total effect is identifiable from an ESCG or an SCG, and Assaad et al. [2023] established identifiability in the presence of instantaneous relations for acyclic SCGs. Ferreira and Assaad [2024] addressed the identifiability problem for general SCGs, including cycles and instantaneous relations for the direct effect; however, the identifiability of

total effects in this context remains unexplored.

Our main contribution consists in demonstrating, under causal sufficiency, when the total effect can be identified when working with such abstractions of a complete graph as an extended summary causal graph (ESCG, as in Figure I.1b) or a summary causal graph (SCG, as in Figure I.1c). The main difficulty lies here on the fact that these abstractions may correspond to complete graphs with potentially different skeletons and compatibility wrt to the true underlying probability distribution.

The remainder of the paper is structured as follows: Section II introduces the main notions, Section III presents the problem setup, identifiability in ESCGs and SCGs are respectively discussed in Sections IV and V, while Section VI concludes the paper. Omitted proofs can be found in the Supplement.

## II. PRELIMINARIES

a) *Graph notions:* For a graph  $\mathcal{G}$ , a *path* from  $X$  to  $Y$  in  $\mathcal{G}$  is a sequence of distinct vertices  $\langle X, \dots, Y \rangle$  in which every pair of successive vertices is adjacent. A *causal path* from  $X$  to  $Y$  is a path from  $X$  to  $Y$  in which all edges are directed towards  $Y$ , that is  $X \rightarrow \dots \rightarrow Y$ . A *backdoor path* between  $X$  and  $Y$  is a path between  $X$  and  $Y$  with an arrowhead into  $X$ . A causal path from  $X$  to  $Y$  and the edge  $Y \rightarrow X$  form a *directed cycle*. We denote by  $Cycles(X, \mathcal{G})$  the set of all directed cycles containing  $X$  in  $\mathcal{G}$ , and by  $Cycles^>(X, \mathcal{G})$  the subset of  $Cycles(X, \mathcal{G})$  with at least 2 vertices (i.e., excluding self-loops). If  $X \rightarrow Y$ , then  $X$  is a parent of  $Y$ . If there is a causal path from  $X$  to  $Y$ , then  $X$  is an *ancestor* of  $Y$ , and  $Y$  is a *descendant* of  $X$ . The sets of parents, ancestors and descendants of  $X$  in  $\mathcal{G}$  are denoted by  $Par(X, \mathcal{G})$ ,  $Anc(X, \mathcal{G})$  and  $Desc(X, \mathcal{G})$  respectively. If a path  $\pi$  contains  $X_i \rightarrow X_j \leftarrow X_k$  as a subpath, then  $X_j$  is a *collider* on  $\pi$ . A vertex  $X_j$  is a *definite non-collider* on a path  $\pi$  if the edge  $X_i \leftarrow X_j$ , or the edge  $X_j \rightarrow X_k$  is on  $\pi$ . A vertex is of *definite status* on a path if it is a collider, a definite non-collider or an endpoint on the path. A path  $\pi$  is of *definite status* if every vertex on  $\pi$  is of definite status. A path  $\pi$  from  $X$  to  $Y$  of definite status is *d-connected* given a vertex set  $\mathcal{Z}$ , with  $X, Y \notin \mathcal{Z}$  if every definite non-collider on  $\pi$  is not in  $\mathcal{Z}$ , and every collider on  $\pi$  has a descendant in  $\mathcal{Z}$ . Otherwise,  $\mathcal{Z}$  *blocks*  $\pi$ . Given an ordered pair of variables  $(X, Y)$  in a DAG  $\mathcal{G}$ , a set of variables  $\mathcal{Z}$  satisfies the *standard backdoor criterion* relative to  $(X, Y)$  if no vertex in  $\mathcal{Z}$  is a descendant of  $X$ , and  $\mathcal{Z}$  blocks every backdoor path between  $X$  and  $Y$ . By a slight abuse of notation, we denote  $\mathcal{G} \setminus \{Y\}$  as the subgraph of  $\mathcal{G}$  when removing the vertex  $Y$  and its corresponding edges. Lastly, the *skeleton* of a graph corresponds to all vertices and edges of the graph without considering edge orientations.

b) *Causal graphs in time series:* Consider  $\mathcal{V}$  a set of  $p$  observational time series and  $\mathcal{V}^f = \{\mathcal{V}_{t-\ell} | \ell \in \mathbb{Z}\}$  the set of temporal instances of  $\mathcal{V}$  where  $\mathcal{V}_{t-\ell}$  correspond to the variables of the time series at time  $t - \ell$ . We suppose that the time series are generated from an *unknown* dynamic structural causal model (DSCM, Rubenstein et al. [2018]), an extension of structural causal models (SCM, Pearl et al. [2000]) to time series. This DSCM defines a full-time causal graph (FTCG, see

below) which we call the *true* FTCG and a joint distribution  $P$  over its vertices which we call the *true* probability distribution.

The graph that is used to qualitatively represent causal relations described in a DSCM is known as the full-time causal graph (FTCG).

**Definition 1** (Full-time causal graph (FTCG), Figure I.1a). *Let  $\mathcal{V}$  be a set of  $p$  observational time series and  $\mathcal{V}^f = \{\mathcal{V}_{t-\ell} | \ell \in \mathbb{Z}\}$ . The full-time causal graph (FTCG)  $\mathcal{G}^f = (\mathcal{V}^f, \mathcal{E}^f)$  representing a given DSCM is defined by:  $X_{t-\gamma} \rightarrow Y_t \in \mathcal{E}^f$  if and only if  $X$  causes  $Y$  at time  $t$  with a time lag of  $\gamma > 0$  if  $X = Y$  and with a time lag of  $\gamma \geq 0$  for  $X \neq Y$ .*

As common in causality studies on time series, we consider in the remainder acyclic FTCGs with potential self-causes. Note that acyclicity is guaranteed for relations between variables at different time stamps and that self-causes, i.e., the fact that, for any time series  $X$ ,  $X_{t-1}$  causes  $X_t$ , are present in most time series. We furthermore assume causal sufficiency:

**Assumption 1** (Causal sufficiency). *There is no hidden common cause between any two observed variables.*

In practice, it is usually impossible to work with FTCGs and people have resorted to simpler causal graphs, exploiting the fact that causal relations between time series hold throughout time, as formalized in the following assumption which allows one to focus on a finite number of past slices, given by the maximum lag. We fix it to  $\gamma_{\max}$  in the remainder.

**Assumption 2** (Consistency throughout time). *An FTCG  $\mathcal{G}^f$  is said to be consistent throughout time if all the causal relationships remain constant in direction throughout time.*

Experts are used to working with abstractions of causal graphs which summarize the information into a smaller graph that is interpretable, often with the omission of precise temporal information. We consider in this study two known causal abstractions for time series, namely *extended summary causal graphs* and *summary causal graphs*. An extended summary causal graph [Assaad et al., 2022c] distinguishes between past time slices, denoted as  $\mathcal{V}_{t-}^e$ , and present time slices, denoted as  $\mathcal{V}_t^e$ , thus enabling the differentiation between lagged and instantaneous causal relations.

**Definition 2** (Extended summary causal graph (ESCG), Figure I.1b). *Let  $\mathcal{G}^f = (\mathcal{V}^f, \mathcal{E}^f)$  be an FTCG built from the set of time series  $\mathcal{V}$  satisfying Assumption 2 with maximal temporal lag  $\gamma_{\max}$ . The extended summary causal graph (ESCG)  $\mathcal{G}^e = (\mathcal{V}^e, \mathcal{E}^e)$  associated to  $\mathcal{G}^f$  is given by  $\mathcal{V}^e = (\mathcal{V}_{t-}^e, \mathcal{V}_t^e)$  and  $\mathcal{E}^e$  defined as follows:*

- for any  $X$  in  $\mathcal{V}$ , we define two vertices,  $X_{t-}$  and  $X_t$ , respectively in  $\mathcal{V}_{t-}^e$  and  $\mathcal{V}_t^e$ ;
- for all  $X_t, Y_t \in \mathcal{V}_t^e$ ,  $X_t \rightarrow Y_t \in \mathcal{E}^e$  if and only if  $X_t \rightarrow Y_t \in \mathcal{E}^f$ ;
- for all  $X, Y \in \mathcal{V}_{t-}^e$ ,  $X_{t-} \rightarrow Y_t \in \mathcal{E}^e$  if and only if there exists at least one temporal lag  $0 < \gamma \leq \gamma_{\max}$  such that  $X_{t-\gamma} \rightarrow Y_t \in \mathcal{E}^f$ .

*In that case, we say that  $\mathcal{G}^e$  is derived from  $\mathcal{G}^f$ .*

At a higher level of abstraction, a summary causal graph [Peters et al., 2013, Meng et al., 2020, Assaad et al., 2022a,b]

represents causal relationships among time series, regardless of the time delay between the cause and its effect.

**Definition 3** (Summary causal graph (SCG), Figure I.1c). *Let  $\mathcal{G}^f = (\mathcal{V}^f, \mathcal{E}^f)$  be an FTCCG built from the set of time series  $\mathcal{V}$  satisfying Assumption 2 with maximal temporal lag  $\gamma_{\max}$ . The summary causal graph (SCG)  $\mathcal{G}^s = (\mathcal{V}^s, \mathcal{E}^s)$  associated to  $\mathcal{G}^f$  is such that*

- $\mathcal{V}^s$  corresponds to the set of time series  $\mathcal{V}$ ,
- $X \rightarrow Y \in \mathcal{E}^s$  if and only if there exists at least one temporal lag  $0 \leq \gamma \leq \gamma_{\max}$  such that  $X_{t-\gamma} \rightarrow Y_t \in \mathcal{E}^f$ .

*In that case, we say that  $\mathcal{G}^s$  is derived from  $\mathcal{G}^f$  as well as from the ESCG derived from  $\mathcal{G}^f$ .*

Since the FTCCG is assumed to be acyclic, the ESCG is inherently acyclic. In contrast, SCGs may include directed cycles and even self-loops. For example, the three FTCCGs in Figure I.1a and the three ESCGs in Figure I.1b are acyclic, while the SCG in Figure I.1c has a cycle. We use the notation  $X \rightleftarrows Y$  to indicate situations where there are time lags where  $X$  causes  $Y$  and other lags where  $Y$  causes  $X$ . Additionally, if an SCG is an abstraction of an ESCG, in cases where there is no instantaneous relation, ESCGs and SCGs convey the same information.

It is worth noting that if there is a single ESCG or SCG derived from a given FTCCG, different FTCCGs, with possibly different orientations and skeletons, can yield the same ESCG or SCG. For example, the SCG in Figure I.1c can be derived from any FTCCG and any ESCG in Figures I.1a and I.1b, even though they may have different skeletons (for example,  $\mathcal{G}_1^f$  and  $\mathcal{G}_3^f$  or  $\mathcal{G}_1^e$  and  $\mathcal{G}_3^e$ ) and different orientations (for example,  $\mathcal{G}_1^f$  and  $\mathcal{G}_2^f$  or  $\mathcal{G}_1^e$  and  $\mathcal{G}_2^e$ ). In the remainder, for a given ESCG or SCG  $\mathcal{G}$ , we call any FTCCG from which  $\mathcal{G}$  can be derived as a *candidate FTCCG* for  $\mathcal{G}$ . For example, in Figure I.1,  $\mathcal{G}_1^f$ ,  $\mathcal{G}_2^f$  and  $\mathcal{G}_3^f$  are all candidate FTCCGs for  $\mathcal{G}^s$ . The set of all candidate FTCCGs for  $\mathcal{G}$  is denoted by  $\mathcal{C}(\mathcal{G})$ .

### III. PROBLEM SETUP

We focus in this paper on the *total effect* [Pearl et al., 2000] of the singleton variable  $X_{t-\gamma}$  on the singleton variable  $Y_t$ , written  $P(Y_t = y_t | do(X_{t-\gamma} = x_{t-\gamma}))$  (as well as  $P(y_t | do(x_{t-\gamma}))$  by a slight abuse of notation), when the only knowledge one has of the underlying DSCM consists in the ESCG or SCG derived from the unknown, true FTCCG.  $Y_t$  corresponds to the response and  $do(X_{t-\gamma} = x_{t-\gamma})$  represents an intervention (as defined in Pearl et al. [2000] and Eichler and Didelez [2007, Assumption 2.3]) on the variable  $X$  at time  $t - \gamma$ , with  $\gamma \geq 0$ .

The above setting is very common in practice and entails that one neither knows the true FTCCG nor the true probability distribution. Furthermore, even if one has access to observed data, in practice such observations are finite, which prevents one from discovering the true FTCCG, and even from identifying it in the set of candidate FTCCGs, as no existing causal discovery method is guaranteed to yield the true FTCCG in the finite data setting [Assaad et al., 2022b]. In the purely theoretical context of infinite data, such an identification is

only possible with additional assumptions, beyond the scope of this study.

Each candidate FTCCG proposes a particular decomposition of the true joint probability distribution which is given by the standard recursive decomposition that characterizes Bayesian networks. Not all decompositions are however correct wrt the true probability distribution  $P$ . We meet there the notion of compatibility, which we restrict here to  $P$  and candidate FTCCGs.

**Definition 4** (Compatibility with  $P$ , Pearl et al. [2000]). *If the probability distribution  $P$  admits the decomposition of a candidate FTCCG  $\mathcal{G}^f$ , we say that  $\mathcal{G}^f$  is compatible with  $P$ .*

Note that only the (unknown) true FTCCG is guaranteed to be compatible with  $P$ .

In general, a total effect  $P(y_t | do(x_{t-\gamma}))$  is said to be identifiable from a graph if it can be uniquely computed with a do-free formula from the observed distribution [Pearl, 1995a, Perkovic, 2020]. In our context, this means that the same do-free formula should hold in all candidate FTCCG so as to guarantee that it holds for the true one.

**Definition 5** (Identifiability of total effects in ESCGs and SCGs). *In a given ESCG or SCG  $\mathcal{G}$ ,  $P(y_t | do(x_{t-\gamma}))$  is identifiable iff it can be rewritten with a do-free formula that is valid for any FTCCG in  $\mathcal{C}(\mathcal{G})$ .*

do-free formulas are typically obtained through an adjustment set which contains variables that are sufficient to adjust for computing the total effect from  $x_{t-\gamma}$  to  $y_t$ . The standard backdoor criterion, introduced in Pearl [1995b] and which typically allows one to obtain adjustment sets in the true FTCCG, has been shown to be complete for identifiability under Assumption 1. We introduce here a new backdoor criterion specific to ESCGs and SCGs which we call the *summary backdoor criterion*.

**Definition 6** (Summary backdoor criterion). *Let  $X, Y$  be disjoint vertices in an ESCG or SCG  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ . A set of vertices  $\mathcal{Z}$  in  $\mathcal{V}$  satisfies the summary backdoor criterion relative to  $(X, Y)$  and  $\gamma$  if*

- (i)  $\mathcal{Z}$  blocks all backdoor paths between  $X_{t-\gamma}$  and  $Y_t$  in any FTCCG in  $\mathcal{C}(\mathcal{G})$ ,
- (ii)  $\mathcal{Z}$  does not block any directed path between  $X_{t-\gamma}$  and  $Y_t$  in any FTCCG in  $\mathcal{C}(\mathcal{G})$ ,
- (iii)  $\mathcal{Z}$  does not contain any descendant of  $X_t$  in any FTCCG in  $\mathcal{C}(\mathcal{G})$ .

Note that when there is no backdoor path between  $X_{t-\gamma}$  and  $Y_t$  in any FTCCG in  $\mathcal{C}(\mathcal{G})$ ,  $\mathcal{Z} = \emptyset$  satisfies the summary backdoor criterion.

The summary backdoor criterion is sound for the identification of the total effect  $P(y_t | do(x_{t-\gamma}))$  in an ESCG or SCG, as stated in the following proposition.

**Proposition 1.** *Let  $X$  and  $Y$  be distinct vertices in an ESCG or SCG  $\mathcal{G}$  of a DSCM with true (unknown) probability  $P$ . Under Assumptions 1 and 2 for  $\mathcal{G}$ , if there exists a set  $\mathcal{Z}$  satisfying the summary backdoor criterion relative to  $(X, Y)$*

and  $\gamma$ , then the total effect of  $X_{t-\gamma}$  on  $Y_t$  is identifiable in  $\mathcal{G}$ , and

$$P(y_t|do(x_{t-\gamma})) = \sum_{\mathbf{z}} P(y_t|x_{t-\gamma}, \mathbf{z})P(\mathbf{z}).$$

It turns out that the summary backdoor criterion is also complete for the identification of the total effect  $P(y_t|do(x_{t-\gamma}))$  in an ESCG or SCG, provided that the interventions considered are effective in the following sense.

**Assumption 3.** (Effective intervention) Let  $P$  be the true (unknown) distribution of a given DSCM. For any FTCCG  $\mathcal{G}^f$  compatible with  $P$ , for any set of variables  $\mathcal{Z}$  in  $\mathcal{G}^f$  which contains descendants of  $X_{t-\gamma}$  and/or does not block all backdoor paths between  $X_{t-\gamma}$  and  $Y_t$  in  $\mathcal{G}$ , we assume that:

$$\mathbb{E}_{\mathcal{Z}|do(x_{t-\gamma})}[P(y_t|\mathcal{Z}, do(x_{t-\gamma}))] \neq \mathbb{E}_{\mathcal{Z}}[P(y_t|\mathcal{Z}, x_{t-\gamma})].$$

This assumption requires that the effect of an intervention setting  $X_{t-\gamma}$  to  $x_{t-\gamma}$  has an impact on the expectation of the probability of observing  $y_t$  given  $\mathcal{Z}$  and  $x_{t-\gamma}$ , when  $\mathcal{Z}$  does not satisfy the backdoor criterion. As discussed in Appendix A, this assumption has strong justifications. Lastly, note that any FTCCG compatible with  $P$  is a candidate FTCCG for the ESCG or SCG associated with  $P$ .

We can now state the theorem related to the completeness of the summary backdoor criterion is complete for the identification of the total effect  $P(y_t|do(x_{t-\gamma}))$  in an ESCG or SCG.

**Theorem 1.** Let  $X$  and  $Y$  be distinct vertices in an ESCG or SCG  $\mathcal{G}$  of a DSCM with true (unknown) probability  $P$ . Under Assumptions 1 and 2 for  $\mathcal{G}$  and Assumption 3 for  $P$ , the total effect of  $X_{t-\gamma}$  on  $Y_t$  is identifiable in  $\mathcal{G}$  only if there exists a set  $\mathcal{Z}$  satisfying the summary backdoor criterion relative to  $(X, Y)$  and  $\gamma$ . Furthermore:

$$P(y_t|do(x_{t-\gamma})) = \sum_{\mathbf{z}} P(y_t|x_{t-\gamma}, \mathbf{z})P(\mathbf{z}).$$

## Remarks

- 1) Note that if  $\mathcal{X}$  or  $\mathcal{Y}$  are non-singleton sets, the summary backdoor criterion may fail to identify the total effect of  $\mathcal{X}$  on  $\mathcal{Y}$ . This is due to the fact that orientation (and thus causal ordering) may be different in different candidate FTCCGs, similarly to Perkovic [2020].
- 2) It is important to note that our context differs from previous studies in identifiability (for example Pearl [1995a], Shpitser and Pearl [2008], Blondel et al. [2016], Perkovic [2020]) as the graphs we have to consider for a given ESCG or SCG, namely the candidate FTCCGs, may have different skeletons and may not all be compatible with the true underlying distribution. This contrasts with the graphs associated with CPDAGs and MPDAGs studied in particular in Perkovic [2020]. This also explains why we need a specific assumption on the type of distributions on which completeness of the summary backdoor criterion can be established<sup>1</sup>.

<sup>1</sup>Note that the summary backdoor criterion is sound for any type of distributions.

We now turn to the identifiability problem in ESCGs and SCGs in the next two sections.

## IV. IDENTIFIABILITY IN ESCG

The total effect is always identifiable by adjustment in ESCGs, as stated in the following theorem.

**Theorem 2.** (Identifiability in ESCG) Consider an ESCG  $\mathcal{G}^e$ . Under Assumptions 1 and 2 for  $\mathcal{G}^e$ , the total effect  $P(y_t|do(x_{t-\gamma}))$  is identifiable in  $\mathcal{G}^e$  for any  $\gamma \geq 0$ . Furthermore, the set

$$\mathcal{B}_\gamma = \{(Z_{t-\gamma-\ell})_{1 \leq \ell \leq \gamma_{\max}} | Z_{t-} \in \text{Par}(X_t, \mathcal{G}^e)\} \\ \cup \{Z_{t-\gamma} | Z_t \in \text{Par}(X_t, \mathcal{G}^e)\},$$

is an adjustment set for  $P(y_t|do(x_{t-\gamma}))$ .

If  $\mathcal{B}_\gamma$  is a valid adjustment set, it may still be very large. Additional adjustment sets, potentially smaller than  $\mathcal{B}_\gamma$ , can however be obtained in the densest candidate FTCCG, which is the candidate FTCCG which contains all potential edges and is thus maximal in the number of edges.

**Proposition 2.** Consider an ESCG  $\mathcal{G}^e$  and a maximal lag  $\gamma_{\max}$  and let  $\gamma \geq 0$ . Any adjustment set  $\mathcal{B}'_\gamma$  for the total effect  $P(y_t|do(x_{t-\gamma}))$  valid with respect to the standard backdoor criterion on the densest candidate FTCCG is a valid adjustment set for the total effect in  $\mathcal{G}^e$  and

$$P(y_t|do(x_{t-\gamma})) = \sum_{\mathbf{b} \in \mathcal{B}'_\gamma} P(\mathbf{b})P(y_t|x_{t-\gamma}, \mathbf{b}).$$

In addition,  $\mathcal{B}_\gamma$  is a valid adjustment set with respect to the standard backdoor criterion on the densest candidate FTCCG.

Note however that smaller (in the number of variables) adjustment sets may exist in the true FTCCG when it is different from the densest candidate FTCCG.

## V. IDENTIFIABILITY IN SCG

SCGs are the most abstract version of causal graphs considered in this paper. Due to the potential cycles, we use the notion of  $\sigma$ -blocked path introduced in Forré and Mooij [2017], considered here in the context of causal sufficiency.

**Definition 7** ( $\sigma$ -blocked path with no hidden confounding). Consider an SCG  $\mathcal{G}^s = (\mathcal{V}^s, \mathcal{E}^s)$ . A path  $\pi = \langle X, \dots, Y \rangle$  in  $\mathcal{G}^s$  is  $\sigma$ -blocked by a set  $\mathcal{Z} \subseteq \mathcal{V}^s$  if:

- at least one of the endpoints  $X, Y$  is in  $\mathcal{Z}$ ; or
- there exists a vertex  $W \in \pi$ , such that  $W \notin \text{Anc}(\mathcal{Z}, \mathcal{G}^s)$  with two adjacent edges in  $\pi$  that form a collider at  $Z$ , (i.e.,  $\rightarrow W \leftarrow$ ); or
- there exists a vertex  $Z \in \pi \cap \mathcal{Z}$  with two adjacent edges in  $\pi$  that forms a non-collider at  $Z$ , and such that there exists a directed edge pointing from  $Z$  to a vertex not in the strongly connected component  $\text{Anc}(Z, \mathcal{G}^s) \cap \text{Desc}(Z, \mathcal{G}^s)$ .

If none of the above holds then the path  $\langle X, \dots, Y \rangle$  is  $\sigma$ -active by  $\mathcal{Z}$ .

Note that a backdoor path can start with the edge  $\rightleftarrows$ , and particularly  $X \rightleftarrows Y$  is the only  $\sigma$ -active backdoor path of size 2 in  $\mathcal{G}^s$  with  $\mathcal{Z} \subseteq \mathcal{V}^s \setminus \{X, Y\}$ .

We now present the main result of this section, which states necessary and sufficient conditions for the identifiability in SCG. Remind that  $Cycles(X, \mathcal{G}^s)$  is the set of all directed cycles containing  $X$  in  $\mathcal{G}^s$ , and  $Cycles^>(X, \mathcal{G}^s)$  is the subset where cycles have at least 2 vertices.

**Theorem 3. (Identifiability in SCG)** Consider an SCG  $\mathcal{G}^s = (\mathcal{V}^s, \mathcal{E}^s)$  associated with a DSCM with true (unknown) probability distribution  $P$ . Under Assumptions 1 and 2 for  $\mathcal{G}^s$  and Assumption 3 for  $P$ , the total effect  $P(y_t | do(x_{t-\gamma}))$ , with  $\gamma \geq 0$ , is identifiable if and only if  $X \notin Anc(Y, \mathcal{G}^s)$  or  $X \in Anc(Y, \mathcal{G}^s)$  and none of the following holds:

- 1)  $\gamma \neq 0$  and  $Cycles^>(X, \mathcal{G}^s \setminus \{Y\}) \neq \emptyset$ , or
- 2) there exists a  $\sigma$ -active backdoor path, with  $\mathcal{Z} = \emptyset$ ,

$$\pi^s = \langle V^1 = X, \dots, V^n = Y \rangle$$

from  $X$  to  $Y$  in  $\mathcal{G}^s$  such that  $\langle V^2, \dots, V^{n-1} \rangle \subseteq Desc(X, \mathcal{G}^s)$  and one of the following holds:

- a)  $n > 2$ , i.e.  $\langle V^2, \dots, V^{n-1} \rangle \neq \emptyset$ , or
- b)  $n = 2$  and  $\gamma \neq 1$ , or
- c)  $n = 2$ ,  $\gamma = 1$  and  $Cycles(Y, \mathcal{G}^s \setminus \{X\}) \neq \emptyset$ .

Note that Assumption 3 is needed to characterize the graph and the distribution for non-identifiable total effects, but is not necessary for graphical criterion of identification.

In the remainder, we prove the equivalence stated in the above theorem through Lemmas stated in Section V-A for the necessary conditions and in Section V-B for the sufficient conditions.

#### A. Necessary conditions

We first introduce the notion of ambiguous vertices, represented in gray in every figure, that will be useful for the proofs of most of the lemmas.

**Definition 8 (Ambiguous vertices).** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$ , for  $\gamma \geq 0$ . A vertex  $V_{t'}$  belonging to a backdoor path for  $(X_{t-\gamma}, Y_t)$  in a candidate FTCTG is ambiguous if there exists another candidate FTCTG in which  $V_{t'}$  belongs to a directed path from  $X_{t-\gamma}$  to  $Y_t$  or in which  $V_{t'}$  is a descendant of  $Y_t$ .

Ambiguous vertices are crucial for identifiability.

**Property 1.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$  for  $\gamma \geq 0$ . If there exists a backdoor path  $\pi^f$  in a candidate FTCTG between  $X_{t-\gamma}$  and  $Y_t$  in which all vertices are ambiguous then  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.

Property 1 provides a way to prove the necessary conditions given in Theorem 3 by exhibiting in a candidate FTCTG a path consisting only of ambiguous vertices. This situation, which is illustrated in Figures V.1-V.4, is the foundation for the proofs of Lemmas V.1 - V.4 which alternately consider all the conditions, starting with Condition 1 in Lemma V.1, of Theorem 3.

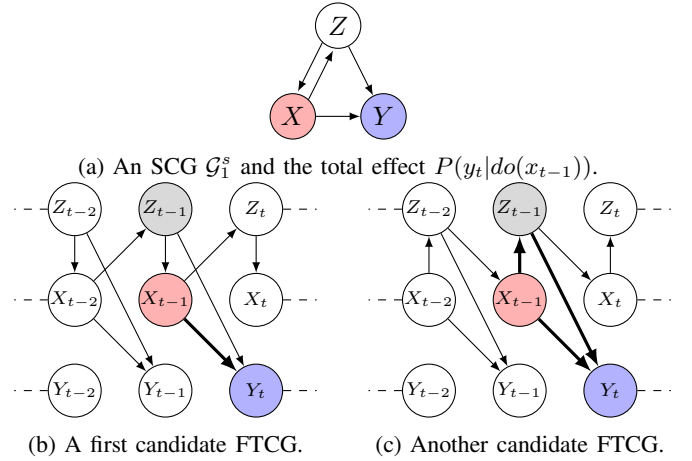


Figure V.1: An example of an SCG  $\mathcal{G}_1^s$  (a) satisfying the conditions of Lemma V.1 with two candidate FTCTGs in (b) and (c). Each pair of red and blue vertices in the FTCTGs represents the total effect we are interested in. Gray vertices are ambiguous: they constitute a backdoor path in (b) and belong to a directed path in (c) (bold edges indicate direct paths from  $X_{t-1}$  to  $Y_t$ ). The total effect is thus not identifiable given the SCG as stated in Property 1.

**Lemma V.1.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$  for  $\gamma > 0$ . If  $X \in Anc(Y, \mathcal{G}^s)$  and  $Cycles^>(X, \mathcal{G}^s \setminus \{Y\}) \neq \emptyset$  then  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.

**Example V.1.** Given the SCG in Figure V.1a, the total effect  $P(y_t | do(x_{t-1}))$  is non-identifiable since Lemma V.1 is satisfied. This can be illustrated by looking at the two candidate FTCTGs in Figure V.1b and V.1c. The path  $\langle X_{t-1}, Z_{t-1}, Y_t \rangle$  is a back-door path in the first FTCTG depicted in (b), but all vertices (except the endpoints) on this path are ambiguous, since  $Z_{t-1}$  belongs to the directed path  $\langle X_{t-1}, Z_{t-1}, Y_t \rangle$  in the second FTCTG depicted in (c). Thus, by Property 1,  $P(y_t | do(x_{t-1}))$  is non-identifiable.

The second lemma considers Condition 2(a).

**Lemma V.2.** Consider an SCG  $\mathcal{G}^s$ , and the total effect  $P(y_t | do(x_{t-\gamma}))$ , for  $\gamma \geq 0$ . If  $X \in Anc(Y, \mathcal{G}^s)$  and there exists a  $\sigma$ -active backdoor path

$$\pi^s = \langle V^1 = X, \dots, V^n = Y \rangle$$

from  $X$  to  $Y$  in  $\mathcal{G}^s$  with  $\mathcal{Z} = \emptyset$  such that  $\langle V^2, \dots, V^{n-1} \rangle \subseteq Desc(X, \mathcal{G}^s)$  and  $n > 2$  then  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.

**Example V.2.** Given the SCG in Figure V.2a, the total effect  $P(y_t | do(x_{t-1}))$  is non-identifiable since Lemma V.2 is satisfied. This can be illustrated by looking at the two candidate FTCTGs in Figures V.2b and V.2c. The path  $\langle X_{t-1}, Z_{t-1}, X_t, Y_t \rangle$  is a back-door path in the first FTCTG depicted in (b), but all vertices (except the endpoints) on this path are ambiguous, since  $Z_{t-1}$  and  $X_t$  belong to the directed path  $\langle X_{t-1}, Y_{t-1}, Z_{t-1}, X_t, Y_t \rangle$  in the second FTCTG

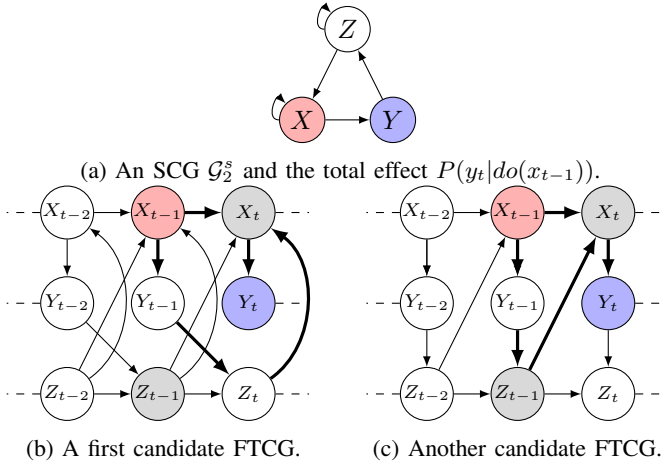


Figure V.2: An example of an SCG  $\mathcal{G}_2^s$  (a) satisfying the conditions of Lemma V.2 with two candidate FTCGs in (b) and (c). Each pair of red and blue vertices in the FTCGs represents the total effect we are interested in. Gray vertices are ambiguous: they constitute a backdoor path in (b) and belong to a directed path in (c) (bold edges indicate direct paths from  $X_{t-1}$  to  $Y_t$ ). The total effect is thus not identifiable given the SCG as stated in Property 1.

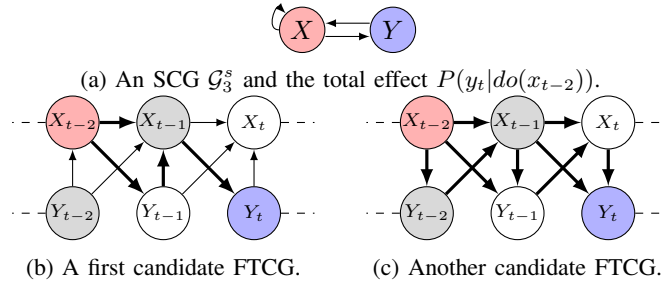


Figure V.3: An example of an SCG  $\mathcal{G}_3^s$  (a) satisfying the conditions of Lemma V.3 with two candidate FTCGs in (b) and (c). Each pair of red and blue vertices in the FTCGs represents the total effect we are interested in. Gray vertices are ambiguous: they constitute a backdoor path in (b) and belong to a directed path in (c) (bold edges indicate direct paths from  $X_{t-2}$  to  $Y_t$ ). The total effect is thus not identifiable given the SCG as stated in Property 1.

depicted in (c). Thus, by Property 1,  $P(y_t | do(x_{t-1}))$  is non-identifiable.

The third lemma considers Condition 2(b), where the only  $\sigma$ -active backdoor path from  $X$  to  $Y$  with  $n = 2$  is  $X \rightleftharpoons Y$ .

**Lemma V.3.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$ , with  $\gamma \neq 1$ . If  $X \rightleftharpoons Y \in \mathcal{G}^s$  then  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.

**Example V.3.** Given the SCG in Figure V.3a, the total effect  $P(y_t | do(x_{t-2}))$  is non-identifiable since Lemma V.3 is satisfied. This can be illustrated by looking at the two candidate FTCGs in Figures V.3b and V.3c. In the first FTCG depicted in (b),  $\langle X_{t-2}, Y_{t-2}, X_{t-1}, Y_t \rangle$  is a backdoor path, but all vertices on the path (except the endpoints) are ambiguous since in

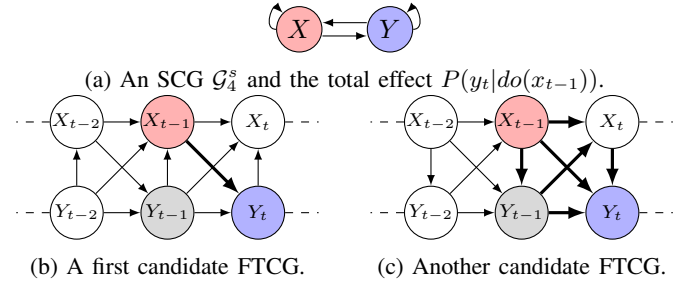


Figure V.4: An example of an SCG  $\mathcal{G}_4^s$  (a) satisfying the conditions of Lemma V.4 with two candidate FTCGs in (b) and (c). Each pair of red and blue vertices in the FTCGs represents the total effect we are interested in. Gray vertices are ambiguous: they constitute a backdoor path in (b) and belong to a directed path in (c) (bold edges indicate direct paths from  $X_{t-1}$  to  $Y_t$ ). The total effect is not identifiable given the SCG as stated in Property 1.

the other FTCG depicted in (c),  $Y_{t-2}$  and  $X_{t-1}$  belong to the directed path  $\langle X_{t-2}, Y_{t-2}, X_{t-1}, Y_t \rangle$ . Thus by Property 1,  $P(y_t | do(x_{t-2}))$  is not identifiable.

Lastly, the fourth lemma considers Condition 2(c) where again, the only  $\sigma$ -active backdoor path from  $X$  to  $Y$  with  $n = 2$  is  $X \rightleftharpoons Y$ .

**Lemma V.4.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-1}))$  ( $\gamma = 1$ ). If  $X \rightleftharpoons Y \in \mathcal{G}^s$  and  $Cycles(Y, \mathcal{G}^s \setminus \{X\}) \neq \emptyset$  then  $P(y_t | do(x_{t-1}))$  is not identifiable.

**Example V.4.** Given the SCG in Figure V.4a, the total effect  $P(y_t | do(x_{t-1}))$  is non-identifiable since Lemma V.4 is satisfied. This can be illustrated by looking at the two candidate FTCGs in Figures V.4b and V.4c. In the first FTCG depicted in (b),  $\langle X_{t-1}, Y_{t-1}, Y_t \rangle$  is a backdoor path, but all vertices on the path (except the endpoints) are ambiguous since in the other FTCG depicted in (c),  $Y_{t-1}$  belongs to the directed path  $\langle X_{t-1}, Y_{t-1}, Y_t \rangle$ . Thus, by Property 1,  $P(y_t | do(x_{t-1}))$  is not identifiable.

## B. Sufficient conditions

We now turn to Lemmas V.5-V.7 to prove that conditions in Theorem 3 are sufficient. To do so, for the total effect  $P(y_t | do(x_{t-\gamma}))$ , we consider the following set:

$$\mathcal{A}_\gamma = \{(Z_{t-\gamma-\ell})_{1 \leq \ell \leq \gamma_{\max}} | Z \in Desc(X; \mathcal{G}^s)\} \cup \{(Z_{t-\gamma-\ell})_{0 \leq \ell \leq \gamma_{\max}} | Z \in \mathcal{V}^s \setminus Desc(X, \mathcal{G}^s)\} \quad (1)$$

and we prove that it is an adjustment set when the total effect is identifiable. As one can note, it contains all possible parents of  $X_{t-\gamma}$  in all candidate FTCGs of  $\mathcal{G}^s$ . Thus,  $\mathcal{A}_\gamma$  blocks any backdoor path  $\pi$  between  $X_{t-\gamma}$  and  $Y_t$  in any candidate FTCG through the parent of  $X_{t-\gamma}$  on that path.

In addition to ambiguous vertices, one can also define ambiguous paths, as follows.

**Definition 9** (Ambiguous paths). Consider an SCG  $\mathcal{G}^s$  and a candidate FTCG  $\mathcal{G}^f$ . A path  $\pi^f \in \mathcal{G}^f$  between  $X_{t-\gamma}$  and  $Y_t$ ,

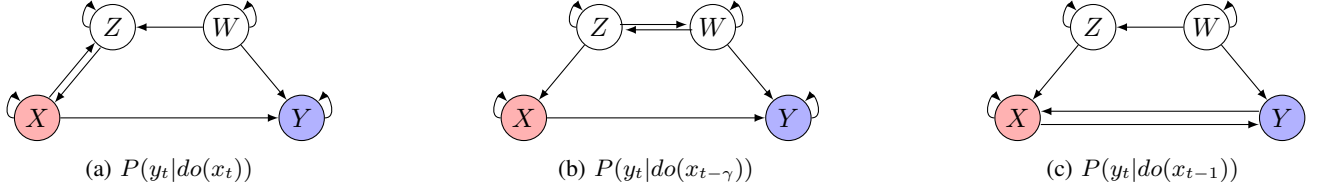


Figure V.5: Three SCGs and a total effect which is identifiable. Each pair of red and blue vertices in the FTCTGs represents the total effect we are interested in, and we precise the total effect and the lag considered in the caption. This illustrates Lemma V.2 (Figure a-b) and Lemma V.3 (Figure c).

for  $\gamma \geq 0$ , is an ambiguous path if it does not contain any vertex at time  $t - \gamma - \ell$  for  $\ell \geq 1$ . We note  $\Pi_\gamma^f$  the set of all ambiguous paths in  $\mathcal{G}^f$ .

When  $\pi$  is not an ambiguous path ( $\pi \notin \Pi_\gamma^f$ ), then the parent of  $X_{t-\gamma}$  is in the past of  $X_{t-\gamma}$  and thus cannot be ambiguous. One thus has the following property:

**Property 2.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$ , for  $\gamma \geq 0$ . Suppose  $\pi^f$  is a backdoor path between  $X_{t-\gamma}$  and  $Y_t$  in a candidate FTCTG  $\mathcal{G}^f$ . If  $\pi^f \notin \Pi_\gamma^f$ , then  $\pi^f$  is blocked by a subset of  $\mathcal{A}_\gamma$  containing at least one non-ambiguous vertex.

**Example V.5.** For example, in Figure V.4c,  $\pi_1^f = \langle X_{t-1}, X_{t-2}, Y_{t-1}, Y_t \rangle$  is not an ambiguous path between  $X_{t-1}$  and  $Y_t$  since  $X_{t-2}$  precedes  $X_{t-1}$  in time. On the other hand,  $\pi_2^f = \langle X_{t-1}, Y_{t-1}, Y_t \rangle$  is an ambiguous path between  $X_{t-1}$  and  $Y_t$ . The path  $\pi_1^f$  is blocked by  $X_{t-2}$ .

We now introduce the notion of compatible path that will allow us to relate backdoor paths in a given SCG and its candidate FTCTGs.

**Definition 10** (Compatible path). Consider an SCG  $\mathcal{G}^s$ , a candidate FTCTG  $\mathcal{G}^f$ , and the total effect  $P(y_t | do(x_{t-\gamma}))$ , for  $\gamma \geq 0$ . We say that a path  $\pi^f = \langle X_{t-\gamma}, W_{t_2}^2, \dots, W_{t_{m-1}}^{m-1}, Y_t \rangle$  in  $\mathcal{G}^f$  is compatible with a path  $\pi^s = \langle X, V^2, \dots, V^{n-1}, Y \rangle$  in  $\mathcal{G}^s$  if for all  $(W_{t_j}^j)_{2 \leq j \leq m-1}$ : either  $W_{t_j}^j \in \langle V^2, \dots, V^{n-1} \rangle$  or  $\exists V \in \langle V^2, \dots, V^{n-1} \rangle$  such that  $W_{t_j}^j \in \text{Cycles}(V, \mathcal{G}^s) \setminus \text{Cycles}(X, \mathcal{G}^s)$ .

The following property relates backdoor paths in a given SCG and in any of its candidate FTCTG.

**Property 3.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$  for  $\gamma \geq 0$ . Then (i)  $\Rightarrow$  (ii), where:

- (i)  $\gamma = 0$  or  $\text{Cycles}^>(X, \mathcal{G}^s \setminus \{Y\}) = \emptyset$ ,
- (ii) in any candidate FTCTG  $\mathcal{G}^f$ , there exists no backdoor path  $\pi^f \in \Pi_\gamma^f$  that is not compatible with any backdoor path in  $\mathcal{G}^s$ .

The two above properties allow one to prove the following lemmas, which, except for the first one, are the counterparts of the lemmas presented in the previous section and which prove that each condition Theorem 3 is sufficient. The first lemma is rather straightforward and concern the case where  $X \notin \text{Anc}(Y, \mathcal{G}^s)$  for a given SCG  $\mathcal{G}^s$ .

**Lemma V.5.** Consider an SCG  $\mathcal{G}^s$ ,  $\gamma \geq 0$  fixed and the total effect  $P(y_t | do(x_{t-\gamma}))$ . If  $X \notin \text{Anc}(Y, \mathcal{G}^s)$  then  $P(y_t |$

$do(x_{t-\gamma}))$  is identifiable, and  $P(y_t | do(x_{t-\gamma})) = P(y_t)$ .

The following lemma excludes both Conditions 1 and 2 of Theorem 3 by considering the negation of Condition 1 (in (i)) and the situation in which there is no  $\sigma$ -active backdoor path from  $X$  to  $Y$  with  $\mathcal{Z} = \emptyset$ .

**Lemma V.6.** Consider an SCG  $\mathcal{G}^s$ ,  $\gamma \geq 0$  fixed and the total effect  $P(y_t | do(x_{t-\gamma}))$ . If  $X \in \text{Anc}(Y, \mathcal{G}^s)$  and

- (i) either  $\gamma = 0$  or  $\text{Cycles}^>(X, \mathcal{G}^s \setminus \{Y\}) = \emptyset$  and
- (ii)  $\nexists \sigma$ -active backdoor path  $\pi^s = \langle V^1 = X, \dots, V^n = Y \rangle$  from  $X$  to  $Y$  in  $\mathcal{G}^s$  with  $\mathcal{Z} = \emptyset$  such that  $\langle V^2, \dots, V^{n-1} \rangle \subseteq \text{Desc}(X, \mathcal{G}^s)$ ,

then  $P(y_t | do(x_{t-\gamma}))$  is identifiable by  $\mathcal{A}_\gamma$ .

This lemma is illustrated in Figure V.5a - V.5b.

When there is a  $\sigma$ -active backdoor path from  $X$  to  $Y$  with  $\mathcal{Z} = \emptyset$ , the negation of Condition 2 of Theorem 3 is obtained with  $n = 2$ ,  $\gamma = 1$  and  $\text{Cycles}(Y, \mathcal{G}^s \setminus \{X\}) = \emptyset$ . The negation of Condition 1 of Theorem 3 is obtained in this setting with  $\text{Cycles}^>(X, \mathcal{G}^s \setminus \{Y\}) = \emptyset$ . Note that, as before, having a  $\sigma$ -active backdoor path from  $X$  to  $Y$  with  $\mathcal{Z} = \emptyset$  and  $n = 2$  is equivalent to  $X \rightleftharpoons Y$ .

**Lemma V.7.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-1}))$  ( $\gamma = 1$ ). If the only  $\sigma$ -active backdoor path from  $X$  to  $Y$  in  $\mathcal{G}^s$  with  $\mathcal{Z} = \emptyset$  is  $X \rightleftharpoons Y \in \mathcal{G}^s$  and

- (i)  $\text{Cycles}^>(X, \mathcal{G}^s \setminus \{Y\}) = \emptyset$  and
- (ii)  $\text{Cycles}(Y, \mathcal{G}^s \setminus \{X\}) = \emptyset$ ,

then  $P(y_t | do(x_{t-1}))$  is identifiable by  $\mathcal{A}_\gamma$ .

This lemma is illustrated in Figure V.5c.

### C. Adjustment set

When the total effect is identifiable, the set  $\mathcal{A}_\gamma$  defined in Equation (1) is an adjustment set, but it has a large size, so we provide a smaller adjustment set, defined as follows:

$$\mathcal{A}'_\gamma = \{V_{t'} \in \mathcal{A}_\gamma | V \in \text{Anc}(X, \mathcal{G}^s) \cup \text{Anc}(Y, \mathcal{G}^s)\}.$$

**Proposition 3.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$ , with  $\gamma \geq 0$ . Under assumptions of identifiability provided by Theorem 3, the set  $\mathcal{A}'_\gamma$  is an adjustment set for the total effect.

## VI. CONCLUSION

We have studied in this paper the identification of total effects between singleton variables, under causal sufficiency,



for both extended summary causal graphs and summary causal graphs. We have in particular shown that the total effect is always identifiable for extended summary causal graphs. The same does not hold for summary causal graphs for which the situation is slightly more complex. We have here established graphical conditions which are sufficient, in any underlying probability distribution, for the identifiability of the total effect; these conditions are furthermore necessary in all distributions satisfying the effective intervention assumption. In addition, we have provided a set of adjustment sets for estimating the total effect in extended summary causal graphs, and an adjustment set when considering summary causal graphs.

These results have significant implications, such as impact analysis in dynamic systems, particularly in scenarios where experts are unable to provide either a full temporal causal graph or an extended summary causal graph. They are also valuable in cases where the assumptions underlying causal discovery methods for inferring causal graphs with time lags are deemed overly restrictive.

#### References:

- Tara V. Anand, Adele H. Ribeiro, Jin Tian, and Elias Bareinboim. Causal effect identification in cluster dags. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(10):12172–12179, Jun. 2023.
- Charles K. Assaad, Emilie Devijver, and Eric Gaussier. Entropy-based discovery of summary causal graphs in time series. *Entropy*, 24(8), 2022a. ISSN 1099-4300. doi: 10.3390/e24081156.
- Charles K. Assaad, Emilie Devijver, and Eric Gaussier. Survey and evaluation of causal discovery methods for time series. *J. Artif. Int. Res.*, 73, may 2022b. ISSN 1076-9757. doi: 10.1613/jair.1.13428.
- Charles K. Assaad, Emilie Devijver, and Eric Gaussier. Discovery of extended summary graphs in time series. In James Cussens and Kun Zhang, editors, *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*, volume 180 of *Proceedings of Machine Learning Research*, pages 96–106. PMLR, 01–05 Aug 2022c.
- Charles K. Assaad, Imad Ez-Zejjari, and Lei Zan. Root cause identification for collective anomalies in time series given an acyclic summary causal graph with loops. In Francisco Ruiz, Jennifer Dy, and Jan-Willem van de Meent, editors, *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pages 8395–8404. PMLR, 25–27 Apr 2023.
- Gilles Blondel, Marta Arias, and Ricard Gavaldà. Identifiability and transportability in dynamic causal networks. *International Journal of Data Science and Analytics*, 3:131–147, 2016.
- Michael Eichler and Vanessa Didelez. Causal reasoning in graphical time series models. In *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence*, UAI’07, page 109–116, Arlington, Virginia, USA, 2007. AUAI Press. ISBN 0974903930.
- Simon Ferreira and Charles K. Assaad. Identifiability of direct effects from summary causal graphs. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38, 2024.
- Patrick Forré and Joris M. Mooij. Markov properties for graphical models with cycles and latent variables, 2017.
- Marloes Maathuis and Diego Colombo. A generalized backdoor criterion. *The Annals of Statistics*, 43, 07 2013. doi: 10.1214/14-AOS1295.
- Yuan Meng, Shenglin Zhang, Yongqian Sun, Ruru Zhang, Zhi-long Hu, Yiyin Zhang, Chenyang Jia, Zhaogang Wang, and Dan Pei. Localizing failure root causes in a microservice through causality inference. In *2020 IEEE/ACM 28th International Symposium on Quality of Service (IWQoS)*, pages 1–10, 2020. doi: 10.1109/IWQoS49365.2020.9213058.
- Jerzy Neyman, D. M. Dabrowska, and T. P. Speed. On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science*, 5(4):465 – 472, 1990. doi: 10.1214/ss/1177012031.
- Judea Pearl. [Bayesian Analysis in Expert Systems]: Comment: Graphical Models, Causality and Intervention. *Statistical Science*, 8(3):266 – 269, 1993. doi: 10.1214/ss/1177010894.
- Judea Pearl. Causal diagrams for empirical research. *Biometrika*, 82(4):669–688, 1995a.
- Judea Pearl. Causal diagrams for empirical research. *Biometrika*, 82(4):669–688, 12 1995b. ISSN 0006-3444. doi: 10.1093/biomet/82.4.669. URL <https://doi.org/10.1093/biomet/82.4.669>.
- Judea Pearl. *Causality: Models, reasoning and inference*. Cambridge University Press, 2000.
- Judea Pearl. The seven tools of causal inference, with reflections on machine learning. *Commun. ACM*, 62(3): 54–60, feb 2019. ISSN 0001-0782. doi: 10.1145/3241036.
- Judea Pearl et al. Models, reasoning and inference. *Cambridge, UK: CambridgeUniversityPress*, 19:2, 2000.
- Emilija Perkovic. Identifying causal effects in maximally oriented partially directed acyclic graphs. In Jonas Peters and David Sontag, editors, *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 124 of *Proceedings of Machine Learning Research*, pages 530–539. PMLR, 03–06 Aug 2020.
- Jonas Peters, D. Janzing, and B. Schölkopf. Causal inference on time series using restricted structural equation models. In *Advances in Neural Information Processing Systems 26*, pages 154–162, 2013.
- P. K. Rubenstein, S. Bongers, B. Schölkopf, and J. M. Mooij. From deterministic ODEs to dynamic structural causal models. In *Proceedings of the 34th Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 114–123, August 2018. URL <http://auai.org/uai2018/proceedings/papers/43.pdf>.
- Ilya Shpitser and Judea Pearl. Complete identification methods for the causal hierarchy. *Journal of Machine Learning Research*, 9:1941–1979, 2008.
- Peter Spirtes, Clark N Glymour, Richard Scheines, and David Heckerman. *Causation, prediction, and search*. MIT press, 2000.

#### APPENDIX

We first restate the effective intervention assumption.

**Assumption 3.** (Effective intervention) Let  $P$  be the true (unknown) distribution of a given DSCM. For any FTCC  $\mathcal{G}^f$  compatible with  $P$ , for any set of variables  $\mathcal{Z}$  in  $\mathcal{G}^f$  which contains descendants of  $X_{t-\gamma}$  and/or does not block all backdoor paths between  $X_{t-\gamma}$  and  $Y_t$  in  $\mathcal{G}$ , we assume that:

$$\mathbb{E}_{\mathcal{Z}|do(x_{t-\gamma})}[P(y_t|\mathcal{Z}, do(x_{t-\gamma}))] \neq \mathbb{E}_{\mathcal{Z}}[P(y_t|\mathcal{Z}, x_{t-\gamma})].$$

We consider here the augmented version  $\mathcal{G}_{aug}^f$  of  $\mathcal{G}$  [Pearl, 2000] obtained by adding to  $\mathcal{G}^f$  a parent node  $F_{X_{t-\gamma}}$  to  $X_{t-\gamma}$  representing the intervention  $do(x_{t-\gamma})$ .  $F_{X_{t-\gamma}}$  takes two values,  $\{do(x_{t-\gamma}), idle\}$ , and is such that:

$$\begin{aligned} P(x'_{t-\gamma}|\mathcal{P}(X_{t-\gamma})) \\ &= P(x'_{t-\gamma}|\mathcal{P}(X_{t-\gamma})) && \text{if } F_{X_{t-\gamma}} = idle, \\ &= 0 && \text{if } F_{X_{t-\gamma}} = x_{t-\gamma} \neq x'_{t-\gamma}, \\ &= 1 && \text{if } F_{X_{t-\gamma}} = x_{t-\gamma} = x'_{t-\gamma}. \end{aligned}$$

The distribution  $P$  is extended in  $\mathcal{G}_{aug}^f$  in such a way that, for any set  $\mathcal{W}$  of variables in  $\mathcal{G}^f$ ,

$$P(\mathcal{W}|do(x_{t-\gamma})) = P'(\mathcal{W}|F_{X_{t-\gamma}} = do(x_{t-\gamma}))$$

and that  $P$  can be viewed as the posterior distribution induced by conditioning  $F_{X_{t-\gamma}}$  on the value *idle*.

A set of variables  $\mathcal{Z}$  which does not satisfy the standard backdoor criterion for  $P(y_t | do(x_{t-\gamma}))$  in  $\mathcal{G}^f$  is such that it either contains a descendant of  $X_{t-\gamma}$ , or it does not block all backdoor paths between  $X_{t-\gamma}$  and  $Y_t$  in  $\mathcal{G}_{aug}^f$ , or both. If it contains a descendant of  $X_{t-\gamma}$ , then  $\mathcal{Z} \not\perp_{\mathcal{G}_{aug}^f} F_{X_{t-\gamma}}$  and  $P'(\mathcal{Z}|F_{X_{t-\gamma}}) \neq P'(\mathcal{Z})$ . If it does not block all backdoor paths between  $X_{t-\gamma}$  and  $Y_t$  in  $\mathcal{G}_{aug}^f$ , then  $Y_t \not\perp_{\mathcal{G}_{aug}^f} F_{X_{t-\gamma}}|X_{t-\gamma}, \mathcal{Z}$  (by definition of the standard backdoor criterion) meaning that:

$$\begin{aligned} P'(F_{X_{t-\gamma}}, Y_t|\mathcal{Z}, X_{t-\gamma}) &\neq P'(F_{X_{t-\gamma}}|\mathcal{Z}, X_{t-\gamma}) \\ &\quad \times P'(Y_t|\mathcal{Z}, X_{t-\gamma}). \end{aligned}$$

Thus:

$$\begin{aligned} P'(Y_t|\mathcal{Z}, X_{t-\gamma}, F_{X_{t-\gamma}}) &= \frac{P'(Y_t, \mathcal{Z}, X_{t-\gamma}, F_{X_{t-\gamma}})}{P'(\mathcal{Z}, X_{t-\gamma}, F_{X_{t-\gamma}})} \\ &= \frac{P'(Y_t, F_{X_{t-\gamma}}|\mathcal{Z}, X_{t-\gamma})}{P'(F_{X_{t-\gamma}}|\mathcal{Z}, X_{t-\gamma})} \\ &\neq P'(Y_t|\mathcal{Z}, X_{t-\gamma}). \end{aligned}$$

So, for a set of variables  $\mathcal{Z}$  which does not satisfy the backdoor criterion for  $P(y_t | do(x_{t-\gamma}))$  in  $\mathcal{G}^f$ , either  $P'(\mathcal{Z}|F_{X_{t-\gamma}}) \neq P'(\mathcal{Z})$  or  $P'(Y_t|\mathcal{Z}, X_{t-\gamma}, F_{X_{t-\gamma}}) \neq P'(Y_t|\mathcal{Z}, X_{t-\gamma})$ , or both. One can thus expect that the expectation of  $P'(y_t|\mathcal{Z}, x_{t-\gamma})$  wrt  $P'(\mathcal{Z})$  changes when conditioning on the intervention vertex  $F_{X_{t-\gamma}}$  as at least one of the two quantities change. This is what Assumption 3 states.

The following example provides a situation where Assumption 3 does not hold, which justifies the term *effective intervention*, and that it is necessary to prove Theorem 3.

**Example A.1.** Let us consider the following DSCM:

$$\begin{cases} X_0 = Y_0 = Z_0 = 0, \\ X_t = Z_{t-1} \text{ with prob. } p_x, (1 - Z_{t-1}) \text{ with prob. } (1 - p_x), \\ Z_t = X_t \text{ with prob. } p_z, (1 - X_t) \text{ with prob. } (1 - p_z), \\ Y_t = X_{t-1}^{Z_{t-1}} \text{ with prob. } p_y, (1 - X_{t-1})^{Z_{t-1}} \text{ with prob. } (1 - p_y). \end{cases}$$

The associated, true FTCC and SCG are represented in Figure V.1 (c) and (a) respectively. Set  $p_x = p_y = p_z = 0.5$  and consider the intervention  $P(Y_t = 1|do(X_{t-1} = 0))$  denoted here  $P(y_t|do(x_{t-1}))$  (any other value for  $Y_t$  and  $X_{t-1}$  gives similar results). Using the do-calculus results, we have quite straightforwardly, for  $z_{t-1} \in \{0, 1\}$ , that:

$$\begin{cases} P(z_{t-1}) = P(z_{t-1}|do(x_{t-1})), \\ P(y_t|z_{t-1}, x_{t-1}) = P(y_t|z_{t-1}, do(x_{t-1})). \end{cases}$$

Therefore,

$$\begin{aligned} &\sum_{z_{t-1}} P(z_{t-1}|do(x_{t-1}))P(y_t|z_{t-1}, do(x_{t-1})) \\ &= \sum_{z_{t-1}} P(z_{t-1})P(y_t|z_{t-1}, x_{t-1}). \end{aligned}$$

In other words, Assumption 3 is not met. The system is completely random with no way to distinguish between the two values 0 and 1. In this case, intervening or not on  $x_{t-1}$  has no impact on the observed values. Note however that most settings of  $p_x$ ,  $p_y$  and  $p_z$  lead to distributions that satisfy Assumption 3.

Furthermore, Lemma V.1, and thus Theorem 3, states that  $P(y_t|do(x_{t-1}))$  is not identifiable, as illustrated in Figure V.1 which displays, in addition to the true FTCC, a candidate FTCC (Figure V.1 (b)) in which the do-free formula for  $P(y_t|do(x_{t-1}))$  should be different from the one in the true FTCC. However, with our setting, both do-free formulas (obtained by the standard backdoor criterion in Figure V.1 (b) and by a direct rewriting in Figure V.1 (c)) are equivalent to  $P(y_t)$ . Lemma V.1, and thus Theorem 3, are not valid in this case.

#### A. Proofs of Section III

**Proposition 1.** Let  $X$  and  $Y$  be distinct vertices in an ESCG or SCG  $\mathcal{G}$  of a DSCM with true (unknown) probability  $P$ . Under Assumptions 1 and 2 for  $\mathcal{G}$ , if there exists a set  $\mathcal{Z}$  satisfying the summary backdoor criterion relative to  $(X, Y)$  and  $\gamma$ , then the total effect of  $X_{t-\gamma}$  on  $Y_t$  is identifiable in  $\mathcal{G}$ , and

$$P(y_t|do(x_{t-\gamma})) = \sum_{\mathbf{z}} P(y_t|x_{t-\gamma}, \mathbf{z})P(\mathbf{z}).$$

*Proof.* If there exists a set  $\mathcal{Z}$  satisfying the summary backdoor criterion relative to  $(X, Y)$ , then the total effect is identifiable as  $\mathcal{Z}$  satisfies the standard backdoor criterion in all candidate FTCCs; as we are assuming causal sufficiency, the same rewriting with a do-free formula of the total effect exists in all candidate FTCCs:

$$P(y_t|do(x_{t-\gamma})) = \sum_{\mathbf{z}} P(y_t|x_{t-\gamma}, \mathbf{z})P(\mathbf{z}).$$

□ shows that the set  $\mathcal{B}_\gamma$  blocks all back-door paths relatively to  $P(y_t | do(x_{t-\gamma}))$ .

**Theorem 1.** *Let  $X$  and  $Y$  be distinct vertices in an ESCG or SCG  $\mathcal{G}$  of a DSCM with true (unknown) probability  $P$ . Under Assumptions 1 and 2 for  $\mathcal{G}$  and Assumption 3 for  $P$ , the total effect of  $X_{t-\gamma}$  on  $Y_t$  is identifiable in  $\mathcal{G}$  only if there exists a set  $\mathcal{Z}$  satisfying the summary backdoor criterion relative to  $(X, Y)$  and  $\gamma$ . Furthermore:*

$$P(y_t | do(x_{t-\gamma})) = \sum_{\mathbf{z}} P(y_t | x_{t-\gamma}, \mathbf{z}) P(\mathbf{z}).$$

*Proof.* Let us assume that there is no set  $\mathcal{Z}$  satisfying the summary backdoor criterion relative to  $(X, Y)$ . Let denote by  $\mathcal{G}^f$  the true (unknown) FTCTG, and  $\mathcal{S}_a(\mathcal{G}^f)$  the set of all backdoor adjustment sets, according to the standard backdoor criterion, of  $P(y_t | do(x_{t-\gamma}))$  in the FTCTG  $\mathcal{G}^f$ . As  $\mathcal{Z}$  does not satisfy the summary backdoor criterion, there exist (at least) two candidate FTCTGs,  $\mathcal{G}_1^f$  and  $\mathcal{G}_2^f$ , such that  $\mathcal{S}_a(\mathcal{G}_1^f) \cap \mathcal{S}_a(\mathcal{G}_2^f) = \emptyset$ , which implies that either  $\mathcal{S}_a(\mathcal{G}_1^f) \cap \mathcal{S}_a(\mathcal{G}^f) = \emptyset$  or  $\mathcal{S}_a(\mathcal{G}_2^f) \cap \mathcal{S}_a(\mathcal{G}^f) = \emptyset$ , or both by completeness of the standard backdoor criterion in any candidate FTCTG<sup>2</sup>. Without loss of generality, let us assume that  $\mathcal{S}_a(\mathcal{G}_1^f) \cap \mathcal{S}_a(\mathcal{G}^f) = \emptyset$  and let  $\mathcal{Z} \in \mathcal{S}_a(\mathcal{G}_1^f)$ . One has:  $P(y_t | do(x_{t-\gamma})) = \sum_{\mathbf{z}} P(y_t | do(x_{t-\gamma}), \mathbf{z}) P(\mathbf{z} | do(x_{t-\gamma}))$ . As  $\mathcal{G}^f$  is compatible with  $P$ , by Assumption 3, the above quantity is different from  $\sum_{\mathbf{z}} P(y_t | x_{t-\gamma}, \mathbf{z}) P(\mathbf{z} | x_{t-\gamma})$  which is a do-free formula provided by  $\mathcal{G}_1^f$ . As the standard backdoor criterion is complete, there is no do-free formula common to both  $\mathcal{G}_1^f$  and  $\mathcal{G}^f$ , which concludes the proof. □

## B. Proofs of Section IV

**Theorem 2.** *(Identifiability in ESCG) Consider an ESCG  $\mathcal{G}^e$ . Under Assumptions 1 and 2 for  $\mathcal{G}^e$ , the total effect  $P(y_t | do(x_{t-\gamma}))$  is identifiable in  $\mathcal{G}^e$  for any  $\gamma \geq 0$ . Furthermore, the set*

$$\begin{aligned} \mathcal{B}_\gamma = & \{(Z_{t-\gamma-\ell})_{1 \leq \ell \leq \gamma_{\max}} | Z_t \in \text{Par}(X_t, \mathcal{G}^e)\} \\ & \cup \{Z_{t-\gamma} | Z_t \in \text{Par}(X_t, \mathcal{G}^e)\}, \end{aligned}$$

is an adjustment set for  $P(y_t | do(x_{t-\gamma}))$ .

*Proof.* If  $X \notin \text{Anc}(Y, \mathcal{G}^e)$ , then in every candidate FTCTG  $\mathcal{G}^f$ ,  $X_{t-\gamma} \notin \text{Anc}(Y_t, \mathcal{G}^f)$ . Thus,  $P(y_t | do(x_{t-\gamma}))$  is always identifiable in  $\mathcal{G}^e$ , and  $P(y_t | do(x_{t-\gamma})) = P(y_t)$ .

Assume now that  $X \in \text{Anc}(Y, \mathcal{G}^e)$ . Let  $\gamma_{\max}$  be the maximal lag, and  $\mathcal{G}^f$  be a candidate FTCTG. We prove that

$$\begin{aligned} \mathcal{B}_\gamma = & \{(Z_{t-\gamma-\ell})_{1 \leq \ell \leq \gamma_{\max}} | Z_t \in \text{Par}(X_t, \mathcal{G}^e)\} \\ & \cup \{Z_{t-\gamma} | Z_t \in \text{Par}(X_t, \mathcal{G}^e)\} \end{aligned}$$

is an adjustment set for  $P(y_t | do(x_{t-\gamma}))$  in  $\mathcal{G}^f$ .

First, we have to prove that  $\text{Par}(X_{t-\gamma}, \mathcal{G}^f) \subseteq \mathcal{B}_\gamma$ . Let  $Z_{t-\gamma-\ell} \in \text{Par}(X_{t-\gamma}, \mathcal{G}^f)$ . If  $\ell = 0$ , then  $Z_t$  causes  $X_t$  in  $\mathcal{G}^e$  by consistency throughout time, which means that  $Z_{t-\gamma} \in \mathcal{B}_\gamma$ . If  $\ell > 0$ , then  $Z_{t-\ell}$  causes  $X_t$  in  $\mathcal{G}^e$ , that is  $Z_{t-\ell-\gamma} \in \mathcal{B}_\gamma$ . This

<sup>2</sup>If both  $\mathcal{G}_1^f$  and  $\mathcal{G}_2^f$  have a non-empty intersection with  $\mathcal{G}^f$  resp. containing  $\mathcal{Z}_1$  and  $\mathcal{Z}_2$ , then  $\mathcal{Z}_1$  provides a valid do-free formula in  $\mathcal{G}_2^f$  which, by completeness of the standard backdoor criterion implies that  $\mathcal{Z}_1 \in \mathcal{S}_a(\mathcal{G}_2^f)$ .

Then, we have to prove  $\mathcal{B}_\gamma$  does not contain any descendant of  $X_{t-\gamma}$  in  $\mathcal{G}^f$ . If this is true, there exists  $W_{t-\gamma} \in \mathcal{B}_\gamma \cap \text{Desc}(X_{t-\gamma}, \mathcal{G}^f)$ , at time slice  $t - \gamma$  because it is a parent and a descendant of  $X_{t-\gamma}$ . By consistency throughout time,  $W_t \in \text{Desc}(X_t, \mathcal{G}^f)$ . However, by definition of  $\mathcal{B}_\gamma$ ,  $X_t \in \text{Desc}(W_t, \mathcal{G}^f)$ , which contradicts the acyclicity assumption of  $\mathcal{G}^e$ . It means that neither it blocks any directed path between  $X_{t-\gamma}$  and  $Y_t$ , nor it contains any descendant of  $Y_t$ . □

**Proposition 2.** *Consider an ESCG  $\mathcal{G}^e$  and a maximal lag  $\gamma_{\max}$  and let  $\gamma \geq 0$ . Any adjustment set  $\mathcal{B}'_\gamma$  for the total effect  $P(y_t | do(x_{t-\gamma}))$  valid with respect to the standard backdoor criterion on the densest candidate FTCTG is a valid adjustment set for the total effect in  $\mathcal{G}^e$  and*

$$P(y_t | do(x_{t-\gamma})) = \sum_{\mathbf{b} \in \mathcal{B}'_\gamma} P(\mathbf{b}) P(y_t | x_{t-\gamma}, \mathbf{b}).$$

In addition,  $\mathcal{B}_\gamma$  is a valid adjustment set with respect to the standard backdoor criterion on the densest candidate FTCTG.

*Proof.* Let  $\mathcal{G}_d^f$  be the densest candidate FTCTG, and  $\mathcal{B}'_\gamma$  be an adjustment set in  $\mathcal{G}_d^f$ . Let  $\mathcal{G}^f$  be another candidate FTCTG. By definition of  $\mathcal{G}_d^f$ , any back-door path in  $\mathcal{G}^f$  is also in  $\mathcal{G}_d^f$  (the last graph contains all possible edges). Then,  $\mathcal{B}'_\gamma$  blocks all back-door paths in  $\mathcal{G}^f$ . Moreover, since no vertex in  $\mathcal{B}'_\gamma$  is a descendant of  $X_{t-\gamma}$  in  $\mathcal{G}_d^f$ , the same holds for  $\mathcal{G}^f$ . Thus,  $\mathcal{B}'_\gamma$  is also an adjustment set in  $\mathcal{G}^f$ . □

## C. Proofs of Section V

**Property 1.** *Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$  for  $\gamma \geq 0$ . If there exists a backdoor path  $\pi^f$  in a candidate FTCTG between  $X_{t-\gamma}$  and  $Y_t$  in which all vertices are ambiguous then  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.*

*Proof.* Let's assume that the total effect is identifiable. Then, there exists a set of vertices  $\mathcal{W}$  which blocks all back-door paths between  $X_{t-\gamma}$  and  $Y_t$  of all candidate FTCTGs and which does not block any directed paths between  $X_{t-\gamma}$  and  $Y_t$  nor contains any descendant of  $Y_t$ . Let's consider a candidate FTCTG and a back-door path  $\pi^f$  between  $X_{t-\gamma}$  and  $Y_t$  in this FTCTG. All vertices which are in  $\mathcal{W}$  and  $\pi^f$  block  $\pi^f$ . As they are in  $\mathcal{W}$ , they cannot belong, in any candidate FTCTG, to a directed path from  $X_{t-\gamma}$  and  $Y_t$ , as otherwise they would block this path, and cannot be a descendant of  $Y_t$ . They are thus unambiguous according to Definition 8. □

**Lemma V.1.** *Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$  for  $\gamma > 0$ . If  $X \in \text{Anc}(Y, \mathcal{G}^s)$  and  $\text{Cycles}^>(X, \mathcal{G}^s \setminus \{Y\}) \neq \emptyset$  then  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.*

*Proof.* Let  $X \rightarrow W^2 \rightarrow \dots \rightarrow W^{m-1} \rightarrow Y$  be a directed path in  $\mathcal{G}^s$  and  $\langle X, V^2, \dots, V^{n-1}, X \rangle \in \text{Cycles}^>(X, \mathcal{G}^s \setminus \{Y\})$ .

There exists a candidate FTCTG  $\mathcal{G}_1^f$  such that  $\pi_1^f = X_{t-\gamma-1} \rightarrow \dots \rightarrow V_{t-\gamma}^{n-1} \rightarrow X_{t-\gamma} \rightarrow W_{t-\gamma}^2 \rightarrow \dots \rightarrow Y_t$ . There also exists another candidate FTCTG  $\mathcal{G}_2^f$  with the path

$\pi_2^f = X_{t-\gamma} \leftarrow V_{t-\gamma}^{n-1} \leftarrow \dots \leftarrow X_{t-\gamma-1} \rightarrow \dots \rightarrow Y_t$ . Thus, by Property 1,  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.  $\square$

**Lemma V.2.** Consider an SCG  $\mathcal{G}^s$ , and the total effect  $P(y_t | do(x_{t-\gamma}))$ , for  $\gamma \geq 0$ . If  $X \in Anc(Y, \mathcal{G}^s)$  and there exists a  $\sigma$ -active backdoor path

$$\pi^s = \langle V^1 = X, \dots, V^n = Y \rangle$$

from  $X$  to  $Y$  in  $\mathcal{G}^s$  with  $\mathcal{Z} = \emptyset$  such that  $\langle V^2, \dots, V^{n-1} \rangle \subseteq Desc(X, \mathcal{G}^s)$  and  $n > 2$  then  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.

*Proof.* Let  $\pi^s = \langle V^1 = X, \dots, V^n = Y \rangle$  be such a  $\sigma$ -active backdoor path.

Then there exists a candidate FTCCG  $\mathcal{G}_1^f$  such that  $\pi_1^f = X_{t-\gamma} \leftarrow \dots \leftarrow V_{t-\gamma}^i \rightarrow \dots \rightarrow Y_t$  is a backdoor path. As  $X \in Anc(Y, \mathcal{G}^s)$  and all  $V^j \subseteq Desc(X, \mathcal{G}^s)$ ,  $2 \leq j \leq n-1$ , all those variables also belong to a directed path from  $X_{t-\gamma}$  to  $Y_t$  in some other candidate FTCCG  $\mathcal{G}_2^f$ . Thus, by Property 1,  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.  $\square$

**Lemma V.3.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$ , with  $\gamma \neq 1$ . If  $X \rightleftharpoons Y \in \mathcal{G}^s$  then  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.

*Proof.* If  $\gamma = 0$ , there exists a candidate FTCCG  $\mathcal{G}_1^f$  such that  $X_t \in Par(Y_t, \mathcal{G}_1^f)$  and another candidate FTCCG  $\mathcal{G}_2^f$  such that  $Y_t \in Par(X_t, \mathcal{G}_2^f)$ , thus the total effect is not identifiable.

Let  $\gamma > 1$ . There exists a candidate FTCCG  $\mathcal{G}_1^f$  such that  $X_{t-\gamma} \leftarrow Y_{t-\gamma} \rightarrow X_{t-\gamma+1} \rightarrow \dots \rightarrow Y_t$  and  $X_{t-\gamma} \rightarrow Y_t$ . There also exists another candidate FTCCG  $\mathcal{G}_2^f$  such that  $X_{t-\gamma} \rightarrow Y_{t-\gamma} \rightarrow X_{t-\gamma+1} \rightarrow \dots \rightarrow Y_t$ . Then, by Property 1,  $P(y_t | do(x_{t-\gamma}))$  is not identifiable.  $\square$

**Lemma V.4.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-1}))$  ( $\gamma = 1$ ). If  $X \rightleftharpoons Y \in \mathcal{G}^s$  and  $Cycles(Y, \mathcal{G}^s \setminus \{X\}) \neq \emptyset$  then  $P(y_t | do(x_{t-1}))$  is not identifiable.

*Proof.* Let  $\langle Y, V^2, \dots, V^{n-1}, Y \rangle \in Cycles(Y, \mathcal{G}^s \setminus \{X\})$ . There exists a candidate FTCCG  $\mathcal{G}_1^f$  with a back-door path  $X_{t-1} \leftarrow Y_{t-1} \rightarrow \dots \rightarrow V_{t-1}^{n-1} \rightarrow Y_t$  and a directed path  $X_{t-1} \rightarrow Y_t$ . There also exists another candidate FTCCG  $\mathcal{G}_2^f$  with a directed  $X_{t-1} \rightarrow Y_{t-1} \rightarrow \dots \rightarrow V_{t-1}^{n-1} \rightarrow Y_t$ . Then, by Property 1,  $P(y_t | do(x_{t-1}))$  is not identifiable.  $\square$

**Property 3.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$  for  $\gamma \geq 0$ . Then (i)  $\Rightarrow$  (ii), where:

- (i)  $\gamma = 0$  or  $Cycles^>(X, \mathcal{G}^s \setminus \{Y\}) = \emptyset$ ,
- (ii) in any candidate FTCCG  $\mathcal{G}^f$ , there exists no backdoor path  $\pi^f \in \Pi_\gamma^f$  that is not compatible with any backdoor path in  $\mathcal{G}^s$ .

*Proof.* Assume first  $Cycles^>(X, \mathcal{G}^s \setminus \{Y\}) = \emptyset$ . Suppose  $\exists \pi^f = X_{t-\gamma} \leftarrow W_{t-\gamma} \dots \rightarrow Y_t \in \Pi_\gamma^f$  which is a back-door path between  $X_{t-\gamma}$  and  $Y_t$  that is not compatible with any back-door path  $\pi^s = \langle V^1 = X, V^2, \dots, V^{n-1}, V^n = Y \rangle$  in  $\mathcal{G}^s$ .

If  $n = 2$ , then the path compatible with the cycle  $\langle X, X \rangle$  is of the form  $X_{t-\gamma} \rightarrow X_{t-\gamma+i} \rightarrow \dots \rightarrow X_{t-\gamma+j} \rightarrow Y_t$ : it means that  $\pi^f$  cannot be a back-door path.

If  $n > 2$ ,  $W_{t-\gamma}$  is such that  $W \notin \{V^2, \dots, V^{n-1}\}$  and  $\nexists V \in \{V^2, \dots, V^{n-1}\}$  such that  $W \in Cycles(V, \mathcal{G}^s)$ . If the path between  $W_{t-\gamma}$  and  $Y_t$  does not pass by  $X_{t-\gamma+\ell}$  with  $\ell > 0$ , then there exists a back-door path between  $X$  and  $Y$  passing by  $W$  in  $\mathcal{G}^s$  as  $\pi^f$  lies in a candidate FTCCG, which contradicts our assumption. So the path necessarily passes by  $X_{t-\gamma+\ell}$ . Thus there is a cycle  $C_x$  on  $X$  such that  $size(C_x) > 2$ , which leads again to a contradiction. Thus, there does not exist a back-door path  $\pi^f \in \Pi_\gamma^f$  between  $X_{t-\gamma}$  and  $Y_t$  that is not compatible with any back-door path in  $\mathcal{G}^s$ .

The case  $\gamma = 0$  is treated in the same way, with the fact that the path considered cannot go back to  $X_t$  as this would create a cycle in the FTCCG.  $\square$

**Lemma V.5.** Consider an SCG  $\mathcal{G}^s$ ,  $\gamma \geq 0$  fixed and the total effect  $P(y_t | do(x_{t-\gamma}))$ . If  $X \notin Anc(Y, \mathcal{G}^s)$  then  $P(y_t | do(x_{t-\gamma}))$  is identifiable, and  $P(y_t | do(x_{t-\gamma})) = P(y_t)$ .

*Proof.* If  $X \notin Anc(Y, \mathcal{G}^s)$ , then in every candidate FTCCG  $\mathcal{G}^f$ ,  $X_{t-\gamma} \notin Anc(Y_t, \mathcal{G}^f)$ . Thus,  $P(y_t | do(x_{t-\gamma}))$  is always identifiable in  $\mathcal{G}^s$ , and  $P(y_t | do(x_{t-\gamma})) = P(y_t)$ .  $\square$

**Lemma V.6.** Consider an SCG  $\mathcal{G}^s$ ,  $\gamma \geq 0$  fixed and the total effect  $P(y_t | do(x_{t-\gamma}))$ . If  $X \in Anc(Y, \mathcal{G}^s)$  and

- (i) either  $\gamma = 0$  or  $Cycles^>(X, \mathcal{G}^s \setminus \{Y\}) = \emptyset$  and
- (ii)  $\nexists \sigma$ -active backdoor path  $\pi^s = \langle V^1 = X, \dots, V^n = Y \rangle$  from  $X$  to  $Y$  in  $\mathcal{G}^s$  with  $\mathcal{Z} = \emptyset$  such that  $\langle V^2, \dots, V^{n-1} \rangle \subseteq Desc(X, \mathcal{G}^s)$ ,

then  $P(y_t | do(x_{t-\gamma}))$  is identifiable by  $\mathcal{A}_\gamma$ .

*Proof.* We will prove that  $\mathcal{A}_\gamma$  is an adjustment set for  $P(y_t | do(x_{t-\gamma}))$  in any candidate FTCCG under conditions (i) and (ii). Let  $\mathcal{G}^f$  be a candidate FTCCG, and  $\Pi_\gamma^f$  the set of ambiguous paths. By Property 2, any back-door path  $\pi^f \notin \Pi_\gamma^f$  can be blocked by  $\mathcal{A}_\gamma$ . Furthermore, by definition, elements of  $\mathcal{A}_\gamma$  cannot block any directed path between  $X_{t-\gamma}$  and  $Y_t$  and cannot be descendants of  $Y_t$  as otherwise they would be descendant of  $X_{t-\gamma}$ .

We now turn our attention to paths in  $\Pi_\gamma^f$ . Let  $\pi^f \in \Pi_\gamma^f$  be a back-door path between  $X_{t-\gamma}$  and  $Y_t$ . Since  $\gamma = 0$  or  $Cycles^>(X, \mathcal{G}^s \setminus \{Y\}) = \emptyset$  then by Property 3, all back-door paths in  $\Pi_\gamma^f$  are compatible with back-door paths in  $\mathcal{G}^s$ . Let  $\pi^s = \langle V^1 = X, \dots, V^n = Y \rangle$  be a  $\sigma$ -active back-door path in  $\mathcal{G}^s$  compatible with  $\pi^f$ . By (ii), there exists  $m \geq 1$  vertices such that  $\{V^{i_1}, \dots, V^{i_m}\} \subseteq \langle V^2, \dots, V^{n-1} \rangle$  and  $\{V^{i_1}, \dots, V^{i_m}\} \not\subseteq Desc(X, \mathcal{G}^s)$ . Then,  $\forall V_{t-\gamma}$  such that  $V \in \{V^{i_1}, \dots, V^{i_m}\}$ ,  $V_{t-\gamma} \notin Desc(X_{t-\gamma}, \mathcal{G}^f)$  and since  $X \in Anc(Y, \mathcal{G}^s)$  then it must be the case that  $V \notin Desc(Y, \mathcal{G}^s)$  and by consequence  $V_{t-\gamma} \notin Desc(Y_t, \mathcal{G}^f)$ . Thus,  $V_{t-\gamma}$  cannot be an ambiguous vertex. Its parent in  $\pi^f$  furthermore blocks  $\pi^f$ , is not ambiguous (as otherwise  $V_{t-\gamma}$  would be ambiguous) and is a member of  $\mathcal{A}_\gamma$  by definition of  $\mathcal{A}_\gamma$ . Thus  $\mathcal{A}_\gamma$  blocks all back-door paths between  $X_{t-\gamma}$  and  $Y_t$  in any candidate FTCCG  $\mathcal{G}^f$ . Furthermore, no node in  $\mathcal{A}_\gamma$  can block a directed path between  $X_{t-\gamma}$  and  $Y_t$  or is a descendant of  $Y_t$  as nodes in  $\mathcal{A}_\gamma$  are either defined before  $t - \gamma$  or are not descendant of  $X_{t-\gamma}$ , and thus of  $Y_t$ . This concludes the proof.  $\square$

**Lemma V.7.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-1}))$  ( $\gamma = 1$ ). If the only  $\sigma$ -active backdoor path from  $X$  to  $Y$  in  $\mathcal{G}^s$  with  $\mathcal{Z} = \emptyset$  is  $X \rightleftharpoons Y \in \mathcal{G}^s$  and

- (i)  $Cycles^>(X, \mathcal{G}^s \setminus \{Y\}) = \emptyset$  and
- (ii)  $Cycles(Y, \mathcal{G}^s \setminus \{X\}) = \emptyset$ ,

then  $P(y_t | do(x_{t-1}))$  is identifiable by  $\mathcal{A}_\gamma$ .

*Proof.* We will prove that  $\mathcal{A}_1$  is an adjustment set for  $P(y_t | do(x_{t-1}))$  in any candidate FTTCG under conditions (i) and (ii). Let  $\mathcal{G}^f$  be a candidate FTTCG, and  $\Pi_1^f$  the set of ambiguous paths.

Since (ii) then by Property 3, all back-door paths in  $\Pi_1^f$  are compatible with back-door paths in  $\mathcal{G}^s$ . In addition, by Property 2, any path  $\pi^f \notin \Pi_1^f$  can be blocked by  $\mathcal{A}_1$ . Therefore, in the following, we focus on paths in  $\Pi_1^f$  compatible with back-door paths in  $\mathcal{G}^s$ .

Consider the  $\sigma$ -active back-door path  $\pi^s = \langle X, Y \rangle$ . As there cannot be a loop on  $Y$  by (i), the only path  $\pi^f \in \Pi_1^f$  from  $X_{t-1}$  to  $Y_t$  compatible with  $\pi^s$  that pass by  $Y_{t-1}$  is  $\pi_f = \langle X_{t-1}, Y_{t-1}, X_t, Y_t \rangle$ . Then, under consistency throughout time, acyclicity and temporal priority, the only choices are  $X_{t-1} \rightarrow Y_{t-1} \rightarrow X_t \rightarrow Y_t$  and  $X_{t-1} \leftarrow Y_{t-1} \rightarrow X_t \leftarrow Y_t$ . The first is a directed path, the second a back-door path already blocked due to the collider  $Y_{t-1} \rightarrow X_t \leftarrow Y_t$ . Thus, all potential back-door paths between  $X_{t-1}$  and  $Y_t$  in any candidate  $\mathcal{G}^f$  are blocked, and  $\mathcal{A}_1$  does not activate them.  $\square$

**Theorem 3.** (Identifiability in SCG) Consider an SCG  $\mathcal{G}^s = (\mathcal{V}^s, \mathcal{E}^s)$  associated with a DSCM with true (unknown) probability distribution  $P$ . Under Assumptions 1 and 2 for  $\mathcal{G}^s$  and Assumption 3 for  $P$ , the total effect  $P(y_t | do(x_{t-\gamma}))$ , with  $\gamma \geq 0$ , is identifiable if and only if  $X \notin Anc(Y, \mathcal{G}^s)$  or  $X \in Anc(Y, \mathcal{G}^s)$  and none of the following holds:

- 1)  $\gamma \neq 0$  and  $Cycles^>(X, \mathcal{G}^s \setminus \{Y\}) \neq \emptyset$ , or
- 2) there exists a  $\sigma$ -active backdoor path, with  $\mathcal{Z} = \emptyset$ ,

$$\pi^s = \langle V^1 = X, \dots, V^n = Y \rangle$$

from  $X$  to  $Y$  in  $\mathcal{G}^s$  such that  $\langle V^2, \dots, V^{n-1} \rangle \subseteq Desc(X, \mathcal{G}^s)$  and one of the following holds:

- a)  $n > 2$ , i.e.  $\langle V^2, \dots, V^{n-1} \rangle \neq \emptyset$ , or
- b)  $n = 2$  and  $\gamma \neq 1$ , or
- c)  $n = 2$ ,  $\gamma = 1$  and  $Cycles(Y, \mathcal{G}^s \setminus \{X\}) \neq \emptyset$ .

*Proof.* To prove this theorem, we present necessary conditions in Lemmas V.4-V.7 and sufficient conditions for the identifiability in Lemmas V.1-V.3.  $\square$

**Proposition 3.** Consider an SCG  $\mathcal{G}^s$  and the total effect  $P(y_t | do(x_{t-\gamma}))$ , with  $\gamma \geq 0$ . Under assumptions of identifiability provided by Theorem 3, the set  $\mathcal{A}'_\gamma$  is an adjustment set for the total effect.

*Proof.* Let  $\mathcal{G}^f$  be a candidate FTTCG. Consider  $V_{t'} \in \mathcal{A}'_\gamma \setminus \mathcal{A}'_\gamma$ : by definition of  $\mathcal{A}'_\gamma$ , it follows that  $V_{t'} \notin Anc(X_{t-\gamma}, \mathcal{G}^f) \cup Anc(Y_t, \mathcal{G}^f)$ . Therefore  $V_{t'}$  does not lie on any back-door path between  $X_{t-\gamma}$  and  $Y_t$ :  $V_{t'}$  is not necessary in the adjustment set, confirming that  $\mathcal{A}'_\gamma$  is also an adjustment set.  $\square$