Orientability of Causal Relations in Time Series Using Summary Causal Graphs and Faithful Distributions

Timothée Loranchet
Sorbonne Université, INSERM,
Institut Pierre Louis d'Epidémiologie
et de Santé Publique,
F75012, Paris, France

Charles K. Assaad
Sorbonne Université, INSERM,
Institut Pierre Louis d'Epidémiologie
et de Santé Publique,
F75012, Paris, France

Abstract

Understanding causal relations between temporal variables is a central challenge in time series analysis, particularly when the full causal structure is unknown. Even when the full causal structure cannot be fully specified, experts often succeed in providing a high-level abstraction of the causal graph, known as a summary causal graph, which captures the main causal relations between different time series while abstracting away micro-level details. In this work, we present conditions that guarantee the orientability of micro-level edges between temporal variables given the background knowledge encoded in a summary causal graph and assuming having access to a faithful and causally sufficient distribution with respect to the true unknown graph. Our results provide theoretical guarantees for edge orientation at the micro-level, even in the presence of cycles or bidirected edges at the macro-level. These findings offer practical guidance for leveraging SCGs to inform causal discovery in complex temporal systems and highlight the value of incorporating expert knowledge to improve causal inference from observational time series data.

1 INTRODUCTION

Time series data arise in a wide range of domains, including epidemiology, intensive care units (ICU) monitoring, econometrics, neuroscience, and information technology systems. In these settings, understanding the underlying causal structure is crucial: it enables predictions under interventions, guides decision-making, and supports robust explanations of observed dynamics. For example, determining whether an observed association between two physiological signals in the ICU reflects a genuine causal influence or merely a common response to another process can directly affect patient management. Similarly, in epidemiology, orienting the causal relations between incidence rates, and behavioral factors is key to forecasting and policy design.

Directed acyclic graphs (DAGs) provide a powerful framework [Pearl, 2009] for representing causal relations. In the context of time-indexed systems [Peters et al., 2013, Runge et al., 2019, Runge, 2020], such graphs are commonly referred to as full-time DAGs (FT-DAGs), where each node corresponds to a time-indexed variable. In many applications, however, an FT-DAG cannot be specified directly from background knowledge [Aït-Bachir et al., 2023]. As a result, researchers often rely either on abstractions of the underlying FT-DAG, such as summary causal graphs (SCGs) [Assaad et al., 2023, 2024, Ferreira and Assaad, 2024] where each node corresponds to an entire time series, or on causal discovery algorithms that recover partially oriented structures such as full time completed partially directed acyclic graphs (FT-CPDAGs) [Spirtes et al., 2001, Chickering, 2002]. Each of these two representations (SCG and FT-CPDAG) provides partial but complementary information about the underlying causal structure.

Since both SCGs and FT-CPDAGs contain less information than the true FT-DAG, it is natural to seek a representation that integrates their strengths. A natural candidate is the maximally oriented partially directed acyclic graph (FT-MPDAG) [Perkovic, 2020], which retains all orientations guaranteed either by discovery algorithms or by the background knowledge en-

Proceedings of the 29th International Conference on Artificial Intelligence and Statistics (AISTATS) 2026, Tangier, Morocco. PMLR: Volume 300. Copyright 2026 by the author(s).

coded in the SCG, thereby bringing the representation as close as possible to the FT-DAG. However, relying on causal discovery algorithms comes at a cost: they are computationally demanding and require strong assumptions, such as faithfulness and causal sufficiency. This motivates a fundamental question: can we determine in advance—before running the causal discovery algorithm—whether a specific edge is guaranteed to be oriented, given the SCG and the available background knowledge? Answering this question allows researchers to focus discovery efforts where they are most informative.

This paper investigates the conditions under which causal relations in time series, particularly instantaneous ones, can be oriented by combining information from SCGs with independence constraints extracted from data. Furthermore it discusses the implications of these results on causal reasoning and causal discovery.

The remainder of the paper is organized as follows. Section 2 introduces the necessary background and related concepts. Section 3 presents our main theoretical contributions. Section 4 presents the direct implications these results on causal reasoning. Finally, Section 5 concludes with a discussion and outlines directions for future work.

2 BACKGROUND

We consider multivariate time series $\mathbb{V}$ evolving according to an *unknown* discrete-time dynamic structural causal model (DT-DSCM) [Peters et al., 2013, Runge, 2020, Gerhardus and Runge, 2020, Ferreira and Assaad, 2025], observed from an initial time t_0 up to a final time $t_{\max}$, *i.e.*, $\forall Y_t \in \mathbb{V}$, $Y_t := f^{y_t}(\mathbb{X}, L_t^{y_t})$, where $\mathbb{X}$ represents the observed direct causes of Y_t in $\mathbb{V}$ and $L_t^{y_t}$ a latent noise term for Y_t. The DT-DSCM governs the temporal and causal dependencies among the variables in $\mathbb{V}$. It respects the temporal priority between causes and effects, ensuring that a variable $X_t \in \mathbb{V}$ cannot influence a variable $Y_{t-\ell}$ with $\ell > 0$, while allowing for instantaneous relations, so that X_t may directly affect Y_t. The DT-DSCM induces a full time[1] directed acyclic graph (FT-DAG) $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ with vertices $\mathbb{V}$ and edges $\mathbb{E}$. In $\mathcal{G}$, a parent of $Y_t \in \mathbb{V}$ is any $X_{t'} \in \mathbb{V}$ s.t. $X_{t'} \to Y_t$ is in $\mathbb{E}$. The set of neighbors of Y_t is denoted $Ne(Y_t, \mathcal{G})$ and the set of parents of Y_t is denoted $Pa(Y_t, \mathcal{G})$. A node Y_t in $\mathcal{G}$ is considered a collider on a path p if there exists a subpath $Z_{t''} \to X_{t'} \leftarrow Y_t$ within p. In this context, we will interchangeably refer to the triple $Z_{t''} \to X_{t'} \leftarrow Y_t$ and the node $X_{t'}$ as the collider. Furthermore, $Z_{t''} \to X_{t'} \leftarrow Y_t$ is termed an unshielded collider if $Z_{t''}$ and Y_t are not adjacent.

[1]The FT-DAG is a DAG; we add the prefix "FT" to emphasize that it represents the full-time structure of the dynamic causal model, capturing all temporal and instantaneous dependencies across the observed time series.

The DT-DSCM also induces a distribution compatible with $\mathcal{G}$ according to the causal Markov condition [Spirtes et al., 2001].

Definition 1 (Causal Markov condition, Spirtes et al. [2001])**.** *Consider an FT-DAG $\mathcal{G}$ compatible with a distribution P. For any variable $Y_t \in \mathbb{V}$, Y_t is independent in P of its non-descendant conditional on its parents in $\mathcal{G}$.*

In many applications, the FT-DAG cannot be fully specified from background knowledge [Assaad et al., 2023, Aït-Bachir et al., 2023]. Therefore, researchers in different fields focus on specifying based on background knowledge a summary causal graph, which is a summary of the FT-DAG where each vertex represent a time series.

Definition 2 (Summary Causal Graph (SCG), Assaad et al. [2024])**.** *Consider an FT-DAG $\mathcal{G} = (\mathbb{V}, \mathbb{E})$. The* summary causal graph (SCG) $\mathcal{G}^{\mathbb{s}} = (\mathbb{S}, \mathbb{E}^{\mathbb{s}})$ *compatible with $\mathcal{G}$ is defined in the following way:*

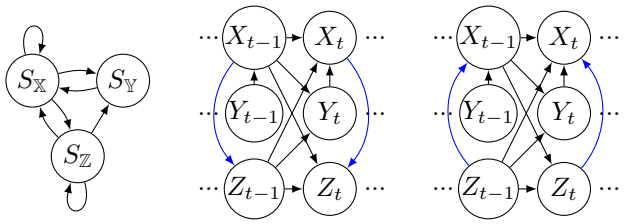
- $\mathbb{S} := \{S_{\mathbb{Y}} | \forall S_{\mathbb{Y}} = \{Y_{t_0}, \cdots, Y_{t_{max}}\} \in \mathbb{V}\}$,
- $\mathbb{E}^{\mathbb{s}} := \{S_{\mathbb{X}} \to S_{\mathbb{Y}} | \forall S_{\mathbb{X}}, S_{\mathbb{Y}} \in \mathbb{S}, \exists t' \leq t \in [t_0, t_{max}]$ *s.t.* $X_{t'} \to Y_t \in \mathbb{E}\}$.

All graphical notions introduced for FT-DAGs, such as kinship and paths, are also applicable to SCGs. Then, as in an FT-DAG, a node $S_{\mathbb{Z}}$ is a collider between $S_{\mathbb{X}}$ and $S_{\mathbb{Y}}$ if it is their common child. Since the SCG can contains bidirected edges, we emphasize that configurations such as $S_{\mathbb{X}} \rightleftarrows S_{\mathbb{Z}} \rightleftarrows S_{\mathbb{Y}}$ or $S_{\mathbb{X}} \to S_{\mathbb{Z}} \rightleftarrows S_{\mathbb{Y}}$ are colliders. However, unlike FT-DAGs, SCGs can contain cycles of various sizes. We denote the special case of cycles which involve only one vertex in the SCG as *self-loops*.

Definition 3 (Self-loop)**.** *Let $\mathcal{G}^{\mathbb{s}} = (\mathbb{S}, \mathbb{E}^{\mathbb{s}})$ be an SCG and $S_{\mathbb{X}} \in \mathbb{S}$ a variable. We say that $\mathcal{G}^{\mathbb{s}}$ contains a* self-loop *on $S_{\mathbb{X}}$ if $(S_{\mathbb{X}}, S_{\mathbb{X}}) \in \mathbb{E}^{\mathbb{s}}$.*

Many different FT-DAGs can be compatible with the same SCG. Examples of an SCG together with two of its compatible FT-DAGs are shown in Figure 1. We denote the set of all FT-DAGs compatible with a given SCG $\mathcal{G}^{\mathbb{s}}$ as *compatible*$(\mathcal{G}^{\mathbb{s}})$.

SCGs can be very useful in many applications but they contain less information than the FT-DAG therefore it is not surprising that the set of causal effects identifiable from an SCG or from an FT-CPDAG is generally smaller than the set identifiable from FT-DAG [Assaad



(a) SCG $\mathcal{G}^{\mathbb{s}}$. (b) FT-DAG $\mathcal{G}_1$. (c) FT-DAG $\mathcal{G}_2$.

Figure 1: An SCG $\mathcal{G}^{\mathbb{s}}$ and two FT-DAGs $\mathcal{G}_1, \mathcal{G}_2 \in$ *compatible*$(\mathcal{G}^{\mathbb{s}})$: The blue edge is oriented differently in the two FT-DAGs.

et al., 2024, Ferreira and Assaad, 2024]. This difficulty in SCG arises because of cycles.

Therefore, in contexts where it needed to acquire more information than the ones provided by the SCG, *e.g.*we would like to know if $X_t \to Y_t$, it can be useful to apply causal discovery algorithms to obtain a representation closer to the underlying DAG. However, using causal discovery algorithms comes with a cost. Causal discovery algorithms can be computationally expensive, and they require additional assumptions. In this paper, we focus on algorithms that require the following three assumptions.

Assumption 1 (Causal sufficiency, Spirtes et al. [2001])**.** *All noise terms in the DT-DSCM are mutually independent and each noise term can affect only one observed variable.*

Assumption 2 (Faithfulness, Spirtes et al. [2001])**.** *All conditional independencies in the distribution are entails by the causal Markov condition.*

Assumption 3 (Stationarity)**.** *For any two temporal variables Y_t and $Y_{t'}$ belonging to the same time series in a DT-DSCM, we have $f^{y_t} = f^{y_{t'}}$. In contrast, this equality does not necessarily hold for temporal variables Y_t and $X_{t'}$ coming from two different time series in the DT-DSCM.*

Assumption 3 implies that, in an FT-DAG induced by a stationary DT-DSCM, if an edge $X_{t'} \to Y_t$ exists, then all edges $X_{t'-\ell} \to Y_{t-\ell}$ also exist for every $\ell \geq 0$ (s.t. $t' - \ell$ and $t - \ell$ remain within the interval $[t_0, t_{\max}]$). Consequently, FT-DAGs can be represented by focusing on a time window ranging from the variables at time t to the variables at time $t - \gamma_{\max}$, where $\gamma_{\max}$ denotes the maximal lag for which a causal link exists between a variable at that lag and t. In the following examples, we set $\gamma_{\max} = 1$ to simplify the FT-DAGs, although larger values are possible. To emphasize that we only depict a fragment of the FT-DAG rather than the complete graph, we always add "…" to indicate that, due to stationarity, the FT-DAG extends both into the past and the future beyond the displayed window of the FT-DAG. It is possible to relax the stationarity assumption by using independent ensembles of time series realizations (i.e., multiple multivariate time series). However, in this paper, we mainly focus on the case where only a single realization of the multivariate time series is available, which makes Assumption 3 necessary. A weaker form of stationarity that accommodates multiple realizations is discussed in Section 5.

Under Assumptions 1, 2, and 3, the FT-DAG $\mathcal{G}$ cannot be recovered; instead, one can guarantee identification of its FT-CPDAG [Chickering, 2002], which represents the Markov equivalence class of all FT-DAGs, denoted as $MEC(\mathcal{G})$, sharing the same set of conditional independencies [Verma and Pearl, 1990]. All DAGs in $MEC(\mathcal{G})$ are characterized by the same skeleton (i.e., the same unoriented graph over the variables) and the same unshielded colliders [Verma and Pearl, 1990]. When additional background knowledge is available, it becomes possible to orient more edges beyond those oriented in the FT-CPDAG. The resulting graph, which augments the FT-CPDAG with these orientations, is called a maximally oriented partially directed acyclic graph (MPDAG) [Perkovic, 2020]. In this paper, the only source of background knowledge we assume is the information encoded in the summary causal graph (SCG). Accordingly, we define the FT-MPDAG[2] in terms of SCGs in the following.

Definition 4 (FT-MPDAG using SCG)**.** *Consider an SCG $\mathcal{G}^{\mathbb{s}}$ and Markov equivalence class $MEC(\mathcal{G})$. A full-time maximally oriented partially acyclic directed graph (FT-MPDAG) $\mathcal{C} = (\mathbb{V}, \mathbb{E}')$ defined using $\mathcal{G}^{\mathbb{s}}$ and $MEC(\mathcal{G})$ is a graph with the same vertices as the original FT-DAG, with edges defined as follows:*

- $\mathbb{E}^{\to} := \{ X_{t'} \to Y_t \mid \forall X_{t'}, Y_t \in \mathbb{V},$ *if*
  - $\forall \mathcal{G}^i \in MEC(\mathcal{G}), X_{t'} \to Y_t \in \mathcal{G}^i$ *or*
  - *both the following holds:*
    - $\exists \mathcal{G}^i, \mathcal{G}^j \in MEC(\mathcal{G})$, *s.t.* $X \to Y \in \mathcal{G}^i$ *and* $Y \to X \in \mathcal{G}^j$, *and*
    - $S_{\mathbb{X}} \to S_{\mathbb{Y}} \in \mathcal{G}^{\mathbb{s}}$ *and* $S_{\mathbb{Y}} \to S_{\mathbb{X}} \notin \mathcal{G}^{\mathbb{s}}\}$,
- $\mathbb{E}^{-} := \{ X_{t'} - Y_t \mid \forall X_{t'}, Y_t \in \mathbb{V},$ *if*
  - $\exists \mathcal{G}^i, \mathcal{G}^j \in MEC(\mathcal{G})$, *s.t.* $X \to Y \in \mathcal{G}^i$ *and* $Y \to X \in \mathcal{G}^j$ *and*
  - *one of the following holds:*
    - $S_{\mathbb{X}} \to S_{\mathbb{Y}} \notin SCG$ *and* $S_{\mathbb{Y}} \to S_{\mathbb{X}} \notin \mathcal{G}^{\mathbb{s}}$, *or*
    - $S_{\mathbb{X}} \to S_{\mathbb{Y}} \in SCG$ *and* $S_{\mathbb{Y}} \to S_{\mathbb{X}} \in \mathcal{G}^{\mathbb{s}}\}$,

[2]While the cited literature refers to CPDAGs and MPDAGs constructed from DAGs, we here use FT-CPDAGs and FT-MPDAGs to emphasize that they are built from FT-DAGs and thus encode temporal structure; this does not change their definitions or any results about them.

where $\mathbb{E}' = \mathbb{E}^{\rightarrow} \cup \mathbb{E}^{-}$.

As in the case of SCGs, all graphical notions introduced for FT-DAGs, applies to FT-MPDAG. Under Assumptions 1, 2 and 3, causal discovery algorithms can recover the FT-MPDAG from observational data. One of the most widely used algorithms is the PC algorithm [Spirtes et al., 2001, Colombo and Maathuis, 2014], which can be naturally extended to time series. We refer to this temporal extension as tPC. Several variants of tPC have been proposed in the literature [Spirtes et al., 2001, Runge et al., 2019, Runge, 2020, Assaad et al., 2022a, Bang and Didelez, 2023], each offering different advantages. For clarity and focus, we restrict our attention to the simplest version, which forms the foundation of these extensions. The only modification we introduce is that we present tPC explicitly in the presence of background knowledge provided by the SCG.

The tPC algorithm proceeds as follows: It starts with a fully connected unoriented graph. It orients all lagged relations from past to present, and orients all instantaneous edges that can be directly resolved using the SCG (*i.e.*, if $S_{\mathbb{X}} \rightarrow S_{\mathbb{Y}}$ but $S_{\mathbb{Y}} \not\rightarrow S_{\mathbb{X}}$, then $X_t \rightarrow Y_t$ is oriented). Using conditional independence tests, it prunes edges from the graph: an edge between $X_{t'}$ and Y_t is removed if there exists a conditioning set $\mathbb{Z}$ s.t. Y_t and $X_{t'}$ are statistically independent given $\mathbb{Z}$. The above step guarantees detecting the correct skeleton of the graph. At this step, because of the temporal lag orientations (i.e. $X_{t'} \rightarrow Y_t$ is oriented whenever $t' < t$), the only edges which are unoriented are instantaneous edges (between X_t and Y_t at the same instant t). Then, for each unshielded triple $Z_{t'} - X_t - Y_t$, it identifies an unshielded collider $Z_{t'} \rightarrow X_t \leftarrow Y_t$ if the middle node X_t does not belong to the separating set that rendered Y_t and Z_t statistically independent. For clarity, we will refer to this rule as the **UC-Rule**. Then, it iteratively applies Meek's four orientation rules [Meek, 1995] to orient as much as possible the remaining unoriented edges. For clarity, in the following we only recall Rules 1 and 2 of Meek, which will be useful for the remainder of the paper.

Meek-Rule 1: If $Z_{t'} \rightarrow X_t - Y_t$ is unshielded, then $Z_{t'} \rightarrow X_t \rightarrow Y_t$.

Meek-Rule 2: If there is a directed path from X_t to Y_t and $X_t - Y_t$, then $X_t \rightarrow Y_t$.

Finally, a subtle background knowledge rule implied by the SCG can be applied to all discovered FT-MPDAGs. By definition, an edge $S_{\mathbb{X}} \rightarrow S_{\mathbb{Y}}$ is present in the SCG if and only if there exist t' and t s.t. $X_{t'} \rightarrow Y_t$ in the FT-DAG. Hence, if there is a remaning unoriented instantaneous edge $X_t - Y_t$ at this step and no other oriented edge from $X_{t'}$ to Y_t, then by this definition of the SCG (and stationarity) we must have $X_t \rightarrow Y_t$; otherwise, the edge $S_{\mathbb{X}} \rightarrow S_{\mathbb{Y}}$ in the SCG would be contradicted.

The FT-DAG and the FT-MPDAG are guaranteed to share the same skeleton and the same unshielded colliders [Verma and Pearl, 1990], along with additional edge orientations implied either by constrains from conditional independencies or by background knowledge. Moreover, as we mentionned, due to the temporal priority of causes over their effects, all lagged relations are oriented identically in both the FT-DAG and the FT-MPDAG. By definition, the only edges that may remain unoriented in the FT-MPDAG are the instantaneous ones. So if the goal is to determine whether an edge such as $X_{t-1} \rightarrow Y_t$ exists in the underlying DAG, then applying these algorithms is certainly worthwhile. However, when the question concerns edges of the form $X_t \rightarrow Y_t$, the situation becomes less straightforward as after running causal discovery, it is possible that the FT-MPDAG only indicates an unoriented edge $X_t - Y_t$, leaving the causal direction unresolved.

In this paper, we aim to characterize the orientations of these instantaneous edges by exploiting the background knowledge provided by an SCG. To this end, we introduce a simple operator that captures the status of the causal orientation between two nodes in an FT-DAG or in an FT-MPDAG:

$$\text{status}(X_t, Y_t; \mathcal{C}) = \begin{cases} X_t \rightarrow Y_t & \text{if } X_t \rightarrow Y_t \in \mathcal{C}, \\ Y_t \rightarrow X_t & \text{if } Y_t \rightarrow X_t \in \mathcal{C}, \\ 1 & \text{if } X_t - Y_t \text{ in } \mathcal{C}, \\ 0 & \text{otherwise.} \end{cases}$$

The central concept underlying our characterization is the orientability of edges from an SCG.

Definition 5 (Orientability of an edge from SCGs and faithful distributions compatible with an FT-DAG)**.** *Let* $\mathcal{G}^{\mathbb{s}}$ *be an SCG, let* $\mathcal{G} \in$ *compatible*$(\mathcal{G}^{\mathbb{s}})$*, and let* $\mathbb{P}$ *denote the set of faithful distributions compatible with* $\mathcal{G}$*. The direct relation between* X_t *and* Y_t*, i.e., status*$(X_t, Y_t; \mathcal{G})$*, is said to be* orientable *from* $(\mathcal{G}^{\mathbb{s}}, \mathbb{P})$ *if*

$$\forall P \in \mathbb{P}, \quad \textit{status}(X_t, Y_t; \mathcal{C}) \neq 1,$$

where $\mathcal{C}$ *denotes the FT-MPDAG constructed from* P *and* $\mathcal{G}^{\mathbb{s}}$*.*

Since there might exists many FT-DAGs compatible with an SCG the set of distributions compatible with an SCG is larger than the set of distributions compatible with a DAG. Consequently, the standard definition of orientability is not sufficient for our purposes. In this paper, we assume that the underlying FT-DAG is unknown, and thus we cannot classify a distribution as

being compatible with any specific FT-DAG. Instead, we must account for all distributions compatible with the SCG. This distinction motivates the notion of *s-orientability.*

Definition 6 (s-orientability of an edge from an SCG and all compatible faithful distributions)**.** *Let $\mathcal{G}^{\mathbb{s}}$ be an SCG, let $\mathcal{G} \in compatible(\mathcal{G}^{\mathbb{s}})$, and let $\mathbb{P}^*$ denote the set of faithful distributions compatible with at least one $\mathcal{G} \in compatible(\mathcal{G}^{\mathbb{s}})$. The direct relation between X_t and Y_t, i.e., $status(X_t, Y_t; \mathcal{G})$, is said to be* s-orientable *from $(\mathcal{G}^{\mathbb{s}}, \mathbb{P}^*)$ if for every $\mathcal{G} \in compatible(\mathcal{G}^{\mathbb{s}})$ and every set of distribution $\mathbb{P} \subset \mathbb{P}^*$ compatible with $\mathcal{G}$, $status(X_t, Y_t; \mathcal{G})$ is orientable from $(\mathcal{G}^{\mathbb{s}}, \mathbb{P})$, i.e.,*

$$\forall P \in \mathbb{P}^*, \quad status(X_t, Y_t; \mathcal{C}) \neq 1,$$

where $\mathcal{C}$ denotes the FT-MPDAG constructed from P and $\mathcal{G}^{\mathbb{s}}$.

We emphasize that we define s-orientability as *s-orientability from the pair* $(\mathcal{G}^{\mathbb{s}}, \mathbb{P}^*)$ to highlight that edge orientations can only be determined given a faithful distribution compatible with $\mathcal{G}^{\mathbb{s}}$. However, since the set of faithful distributions compatible with any FT-DAG compatible with $\mathcal{G}^{\mathbb{s}}$ depends solely on $\mathcal{G}^{\mathbb{s}}$, s-orientability is fundamentally a property of $\mathcal{G}^{\mathbb{s}}$ itself, independent of any specific distribution. While we retain the original notation to stress the role of faithfulness throughout the paper, one could equivalently refer to it as *s-orientability from* $\mathcal{G}^{\mathbb{s}}$ alone.

3 MAIN RESULTS

Starting from Definition 2, it is clear that an oriented (non-bidirected) edge in the SCG is highly informative. Indeed, it indicates that any corresponding edge in the FT-MPDAG is oriented in the same direction. Thus, if an instantaneous edge exists between X_t and Y_t, it is oriented consistently with the corresponding macro-variables in the SCG. Lemma 1 formalizes this intuitive result.

Lemma 1. *Let $\mathcal{G}^{\mathbb{s}}$ be an SCG and $\mathbb{P}^*$ the set of distributions compatible with $\mathcal{G}^{\mathbb{s}}$. Under Assumptions 1, 2 and 3, if $S_{\mathbb{X}} \to S_{\mathbb{Y}} \in \mathcal{G}^{\mathbb{s}}$, then $\forall \mathcal{G} \in compatible(\mathcal{G}^{\mathbb{s}})$, $status(X_t, Y_t; \mathcal{G})$ is s-orientable from $(\mathcal{G}^{\mathbb{s}}, \mathbb{P}^*)$.*

Proof. Immediate consequence of Definition 4. □

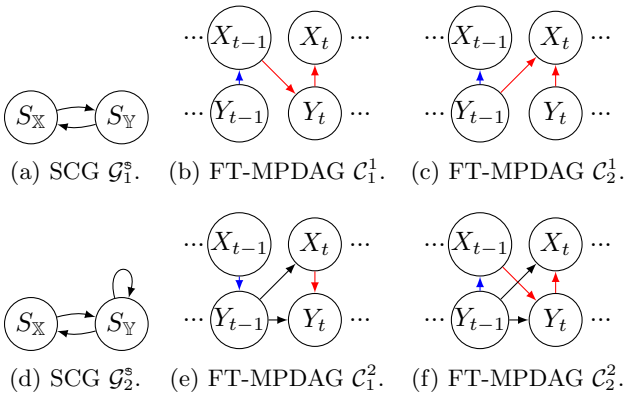
We thus obtain that the only potentially unoriented edges in the FT-MPDAG are instantaneous edges of the form $X_t - Y_t$, whose associated macro-variables are bidirected in the SCG, i.e., $S_{\mathbb{X}} \rightleftarrows S_{\mathbb{Y}}$. The following lemma further reduces these cases by introducing a condition on the self-loops of the macro-variables.

Lemma 2. *Let $\mathcal{G}^{\mathbb{s}}$ be an SCG and $\mathbb{P}^*$ the set of distributions compatible with $\mathcal{G}^{\mathbb{s}}$. Under Assumptions 1, 2 and 3, if $S_{\mathbb{X}} \rightleftarrows S_{\mathbb{Y}} \in \mathcal{G}^{\mathbb{s}}$ and there is no self-loop on* ***both*** *$S_{\mathbb{X}}$* ***and*** *$S_{\mathbb{Y}}$* ***simultaneously****, then $\forall \mathcal{G} \in compatible(\mathcal{G}^{\mathbb{s}})$, $status(X_t, Y_t; \mathcal{G})$ is s-orientable from $(\mathcal{G}^{\mathbb{s}}, \mathbb{P}^*)$.*

Proof. If there is no edge between X_t and Y_t in $\mathcal{G}$, then no edge exists between them in $\mathcal{C}$, since $\mathcal{C}$ shares the same skeleton as $\mathcal{G}$.

Suppose there is an edge between X_t and Y_t in $\mathcal{C}$. We show it must be oriented. First, assume there are no self-loops. Since $S_{\mathbb{X}} \rightleftarrows S_{\mathbb{Y}} \in \mathcal{G}^{\mathbb{s}}$, there exist t_1, t_2 s.t. $t - \gamma_{\max} \leq t_1, t_2 \leq t$, and $Y_{t_1} \to X_t \in \mathcal{C}$ and $X_{t_2} \to Y_t \in \mathcal{C}$; and since $\mathcal{C}$ is acyclic, $t_1 < t$ or $t_2 < t$. Assume without loss of generality that $t_1 < t$. Then, because of temporal ordering $Y_{t_1} \to X_t \in \mathcal{C}$ and (Y_{t_1}, X_t, Y_t) is an unshielded triple, since $S_{\mathbb{Y}}$ has no self-loop and thus Y_{t_1} and Y_t are not adjacent. This triple is either an unshielded collider $Y_{t_1} \to X_t \leftarrow Y_t$, which is oriented in $\mathcal{C}$ by UC-rule, or a chain $Y_{t_1} \to X_t \to Y_t$, which is oriented in $\mathcal{C}$ by Meek-Rule 1. Now, suppose there is a self-loop on one of the nodes. Without loss of generality, let $S_{\mathbb{Y}}$ have a self-loop. Then, there exists $t_1 < t$ s.t. $Y_{t_1} \to Y_t$ in the FT-MPDAG. Two cases arise: (i) there is no t_2 s.t. $t - \gamma_{\max} \leq t_2 < t$ and $X_{t_2} \to Y_t$ so necessarily $X_t \to Y_t$ to satisfy the SCG constraint $S_{\mathbb{X}} \rightleftarrows S_{\mathbb{Y}}$, which requires at least one edge between a time point of $\{X_{t_0}, \cdots, X_{t_{\max}}\}$ and one of $\{Y_{t_0}, \cdots, Y_{t_{\max}}\}$; or (ii) there exists t_3 s.t. $t - \gamma_{\max} \leq t_3 < t$ and $X_{t_3} \to Y_t$ so (X_{t_3}, X_t, Y_t) forms an unshielded triplet (since $S_{\mathbb{X}}$ has no self-loop), and the edge X_t–Y_t is oriented either by Meek-Rule 1 if $X_t \leftarrow Y_t \leftarrow X_{t_3}$, or by the UC-rule if $X_t \to Y_t \leftarrow X_{t_3}$. □

Figure 2 provides an intuitive visual illustration of Lemma 2. The figures are separated into two cases: one corresponds to an SCG $\mathcal{G}_1^{\mathbb{s}}$ without any self-loop (top) and one corresponds to an SCG $\mathcal{G}_2^{\mathbb{s}}$ with a single self-loop (bottom). In each case, the SCG is shown alongside two examples of compatible FT-MPDAGs, ($\mathcal{C}_1^1$ and $\mathcal{C}_2^1$ for the case without self-loop, $\mathcal{C}_1^2$ and $\mathcal{C}_2^2$, for the case with self-loop). In $\mathcal{C}_1^1$, the edge between X_t and Y_t is necessarily oriented according to Meek-Rule 1, $Y_{t-1} \to Y_t \to X_t$ (given that $Y_{t-1} \to Y_t$ is oriented by temporal order). In $\mathcal{C}_2^1$, the edge between X_t and Y_t is oriented due to the unshielded collider $Y_{t-1} \to X_t \leftarrow Y_t$. Note that this Meek rule and the unshielded collider are applied because the triplet (Y_{t-1}, X_t, Y_t) is always unshielded due to the absence of self-loops. In $\mathcal{C}_1^2$, if the edge between X_t and Y_t were oriented in the opposite direction, $Y_t \leftarrow X_t$, this would contradict the SCG $\mathcal{G}^{\mathbb{s}}$, as no edge would then point from $S_{\mathbb{X}}$ to $S_{\mathbb{Y}}$. Therefore, the edge must be ori-



(a) SCG $\mathcal{G}_1^s$. (b) FT-MPDAG $\mathcal{C}_1^1$. (c) FT-MPDAG $\mathcal{C}_2^1$.

(d) SCG $\mathcal{G}_2^s$. (e) FT-MPDAG $\mathcal{C}_1^2$. (f) FT-MPDAG $\mathcal{C}_2^2$.

Figure 2: Visual illustration of Lemma 2, with the top FT-MPDAGs based on the top SCG $\mathcal{G}_1^s$ (no self-loop) and the bottom FT-MPDAGs based on the bottom SCG $\mathcal{G}_2^s$ (with a self-loop). Red illustrates the edges used to orient the edge $X_t - Y_t$, blue is the repetition of this edge by stationarity.

ented as $X_t \to Y_t$ in all compatible DAGs and hence in the FT-MPDAG. In $\mathcal{C}_2^2$, the situation is more similar to the case without self-loops: the edge between X_{t-1} and Y_t is oriented according to Meek-Rule 1, $X_{t-1} \to Y_t \to X_t$.

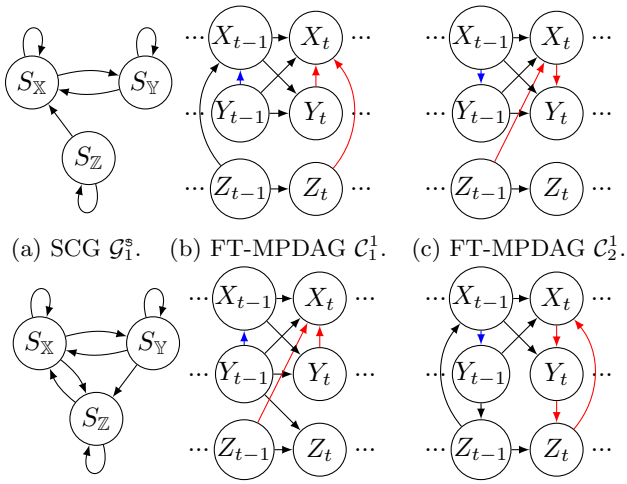
From Lemma 1 and 2, it immediately follows that for an SCG $\mathcal{G}^s$ without self-loops, any FT-MPDAG $\mathcal{C}$ based on $\mathcal{G}^s$ contains no unoriented edges (it is an FT-DAG).

At this stage, the only potentially unoriented edges in the FT-MPDAG are instantaneous edges whose corresponding macro-variables that are bidirected in the SCG and each contain a self-loop. Lemma 3 further restricts the possibility of unoriented edges in the FT-MPDAG by stating that such edges are oriented whenever there exists a node that is a parent of one macro-node but not of the other.

Lemma 3. *Let $\mathcal{G}^s$ be an SCG and $\mathbb{P}^*$ the set of distributions compatible with $\mathcal{G}^s$. Under Assumptions 1, 2 and 3, if $S_{\mathbb{X}} \rightleftarrows S_{\mathbb{Y}} \in \mathcal{G}^s$ and $\exists S_{\mathbb{Z}} \in Pa(S_{\mathbb{X}}, \mathcal{G}^s) \setminus \{S_{\mathbb{X}}\}$ s.t. $S_{\mathbb{Z}} \notin Pa(S_{\mathbb{Y}}, \mathcal{G}^s)$, then $\forall \mathcal{G} \in compatible(\mathcal{G}^s)$, $status(X_t, Y_t; \mathcal{G})$ is s-orientable from $(\mathcal{G}^s, \mathbb{P}^*)$.*

Proof. If there is no edge between X_t and Y_t in $\mathcal{G}$, then no edge exists between them in $\mathcal{C}$, since $\mathcal{C}$ shares the same skeleton as $\mathcal{G}$.

Suppose that there is an edge between X_t and Y_t in $\mathcal{C}$. We show that such an edge must be oriented. First, since $S_{\mathbb{Z}} \in Pa(S_{\mathbb{X}}, \mathcal{G}^s)$, there exists $t' \leq t$ such that $Z_{t'} \to X_t$. If $t' < t$, the triple $(Z_{t'}, X_t, Y_t)$ is unshielded because $Y_t \to Z_{t'}$ is impossible due to the temporal order. Hence this triple either forms a chain, which is oriented by Meek's rule 1 in $\mathcal{C}$, or an unshielded collider, which is also oriented by the UC-Rule in $\mathcal{C}$. If $t' = t$ and there is no $t'' < t$ s.t. $Z_{t''} \to X_t$, the edge $Z_t \to X_t$ is necessarily oriented in $\mathcal{C}$ (otherwise it would contradict $S_{\mathbb{Z}} \in Pa(S_{\mathbb{X}}, \mathcal{G}^s)$). Then we distinguish two case. If Y_t is not adjacent to Z_t in $\mathcal{C}$, the unshielded triple (Z_t, X_t, Y_t) can be oriented by the same argument as in the case $t' < t$. If instead Y_t is adjacent to Z_t, then we necessarily have $Y_t \to Z_t$ since $S_{\mathbb{Z}} \notin Pa(S_{\mathbb{Y}}, \mathcal{G}^s)$. To avoid creating a cycle (Meek's rule 2), since $Y_t \to Z_t \to X_t$, the edge $Y_t \to X_t$ must be oriented in $\mathcal{C}$. □



(a) SCG $\mathcal{G}_1^s$. (b) FT-MPDAG $\mathcal{C}_1^1$. (c) FT-MPDAG $\mathcal{C}_2^1$.

(d) SCG $\mathcal{G}_2^s$. (e) FT-MPDAG $\mathcal{C}_1^2$. (f) FT-MPDAG $\mathcal{C}_2^2$.

Figure 3: Visual illustration of Lemma 3, with the top FT-MPDAGs based on the top SCG $\mathcal{G}_1^s$ (unshielded collider) and the bottom FT-MPDAGs based on the bottom SCG $\mathcal{G}_2^s$ (shielded collider but with $S_{\mathbb{Z}}$ not a parent of $S_{\mathbb{Y}}$). Red illustrates the edges used to orient the edge $X_t - Y_t$, blue is the repetition of this edge by stationarity.

For a visual illustration of Lemma 3, consider the example given in Figure 3 consisting of two SCG $\mathcal{G}_1^s$ and $\mathcal{G}_2^s$ where $S_{\mathbb{X}} \leftrightarrows S_{\mathbb{Y}}$ satisfies condition of Lemma 3, and two compatible FT-MPDAGs for each. The SCG $\mathcal{G}_1^s$ contains an unshielded collider $S_{\mathbb{Z}} \to S_{\mathbb{X}} \leftrightarrows S_{\mathbb{Y}}$ and the SCG $\mathcal{G}_2^s$ contains a shielded collider $S_{\mathbb{Z}} \leftrightarrows S_{\mathbb{X}} \leftrightarrows S_{\mathbb{Y}}$ but with $S_{\mathbb{Z}} \notin Pa(S_{\mathbb{Y}}, \mathcal{G}_2^s)$. First, consider $\mathcal{G}_1^s$: (i) in $\mathcal{C}_1^1$, the edge between X_t and Y_t is oriented by UC-rule applied to the triplet $Z_t \to X_t \leftarrow Y_t$, where $Z_t \to X_t$ is oriented by background knowledge from $S_{\mathbb{Z}} \to S_{\mathbb{X}}$ in $\mathcal{G}_1^s$; and (ii) in $\mathcal{C}_2^1$, the orientation arises from Meek-rule 1 applied to the chain $Z_{t-1} \to X_t \to Y_t$ with $Z_{t-1} \to X_t$ oriented by time. Now, consider $\mathcal{G}_2^s$: (i) in $\mathcal{C}_1^2$, the edge between X_t and Y_t is oriented by UC-rule applied to the triplet $Z_{t-1} \to X_t \leftarrow Y_t$; and (ii) in $\mathcal{C}_2^1$, the orientation arises from Meek-rule 2 applied to the triple $Z_t \to X_t \to Y_t$, with $Y_t \to Z_t$ oriented by background knowledge from $S_{\mathbb{Y}} \to S_{\mathbb{Z}}$ in $\mathcal{G}_2^s$ and $Z_t \to X_t$

oriented to satisfy the constraint $S_{\mathbb{Z}} \in Pa(S_{\mathbb{X}}, \mathcal{G}^{\mathsf{s}}_2)$ since there is no $t' < t$ such that $Z_{t'} \to X_t$ in $\mathcal{C}$.

We have now explored all SCG configurations that ensure s-orientability, as emphasized by Theorem 1 which summarizes these results and establishes their completeness.

Theorem 1 (s-orientability of orientations in the FT-MPDAG from the SCG and faithful distribution)**.** *Let $\mathcal{G}^{\mathsf{s}}$ be an SCG and $\mathbb{P}^*$ the set of distributions compatible with $\mathcal{G}^{\mathsf{s}}$. Under Assumptions 1, 2 and 3, $\forall \mathcal{G} \in compatible(\mathcal{G}^{\mathsf{s}})$, status$(X_t, Y_t; \mathcal{G})$ is **not** s-orientable from* $(\mathcal{G}^{\mathsf{s}}, \mathbb{P}^*)$ if and only if*:*

1. *$S_{\mathbb{X}} \rightleftarrows S_{\mathbb{Y}} \in \mathcal{G}^{\mathsf{s}}$, and*
2. *both $S_{\mathbb{X}}$ and $S_{\mathbb{Y}}$ have self-loops simultaneously in $\mathcal{G}^{\mathsf{s}}$, and*
3. *$Pa(S_{\mathbb{X}}, \mathcal{G}^{\mathsf{s}}) = Pa(S_{\mathbb{Y}}, \mathcal{G}^{\mathsf{s}})$.*

Proof. Lemmas 1, 2 and 3 establish immediately that if at least one of the items of the theorem is not satisfied then status$(X_t, Y_t; \mathcal{G})$ is s-orientable.

In the following, we prove that if the three items of the theorem are satisfied then status$(X_t, Y_t; \mathcal{G})$ is not s-orientable. Let $\mathcal{G}^{\mathsf{s}}$ be an SCG, and let $S_{\mathbb{X}}$ and $S_{\mathbb{Y}}$ be two nodes of $\mathcal{G}^{\mathsf{s}}$ satisfying the three items of the theorem. We show that this implies the existence of an FT-MPDAG $\mathcal{C}$ compatible with $\mathcal{G}^{\mathsf{s}}$ and a faithful distribution $P \in \mathbb{P}^*$ such that, for every $\mathcal{G}$ compatible with $\mathcal{C}$, we have status$(X_t, Y_t; \mathcal{G}) = 1$, that is, the instantaneous edge $X_t - Y_t$ remains unoriented in $\mathcal{C}$. To construct such an FT-MPDAG $\mathcal{C}$, we begin by constructing an FT-DAG $\mathcal{G}$ as follows: (i) add the edge $X_t \to Y_t$ (or $X_t \leftarrow Y_t$, arbitrarily) (ii) for any nodes $S_{\mathbb{W}_1}, S_{\mathbb{W}_2}$ of $\mathcal{G}^{\mathsf{s}}$, if $S_{\mathbb{W}_1} \in Pa(S_{\mathbb{W}_2}, \mathcal{G}^{\mathsf{s}})$, we add the edge $W^1_{t-\gamma_{\max}} \to W^2_t$ to $\mathcal{G}$ (it implies in particular that $X_{t-\gamma_{\max}} \to X_t$, $Y_{t-\gamma_{\max}} \to Y_t$, $X_{t-\gamma_{\max}} \to Y_t$, and $Y_{t-\gamma_{\max}} \to X_t$ in $\mathcal{G}$ since items 1 and 2 of the theorem holds; and that for every $S_{\mathbb{Z}} \in Pa(S_{\mathbb{X}}, \mathcal{G}^{\mathsf{s}})$, $Z_{t-\gamma_{\max}} \to Y_t$ and $Z_{t-\gamma_{\max}} \to X_t$ in $\mathcal{G}$ since item 3 of the theorem holds so $S_{\mathbb{Z}} \in Pa(S_{\mathbb{Y}}, \mathcal{G}^{\mathsf{s}})$. By construction, it is immediate that the resulting FT-DAG $\mathcal{G}$ is compatible with $\mathcal{G}^{\mathsf{s}}$, and we finally define $\mathcal{C}$ as the FT-MPDAG obtained from $\mathcal{G}$ and incorporating only the background knowledge given by the SCG. The edge $X_t - Y_t$ must remain unoriented in $\mathcal{C}$ because all triples involving X_t and Y_t are shielded by construction, so neither the UC-Rule nor Meek's Rule 1 can be applied. Moreover, since no instantaneous edges connect to X_t or Y_t, none of the other Meek rules can be applied (as no cycle can occur between present and past, and by construction there is only one instantaneous edge : $X_t - Y_t$). We have thus shown that if the items of the theorem hold, it is always possible to construct an FT-MPDAG $\mathcal{C}$ in which the edge between X_t and Y_t is unoriented (status$(X_t, Y_t; \mathcal{G}) = 1$), that is status$(X_t, Y_t; \mathcal{G})$ is not s-orientable. □

Figure 4 illustrates the statement of the theorem. It shows that in any FT-MPDAG $\mathcal{C}$ compatible with $\mathcal{G}^{\mathsf{s}}$, the edge between Y_t and Z_t is orientable due to the directed edge $S_{\mathbb{Y}} \to S_{\mathbb{Z}}$ in $\mathcal{G}^{\mathsf{s}}$, while the edge between Z_t and W_t is orientable because of the unshielded collider $S_{\mathbb{W}} \rightleftarrows S_{\mathbb{Z}} \leftarrow S_{\mathbb{Y}}$ in $\mathcal{G}^{\mathsf{s}}$. In contrast, the edge between X_t and Y_t is not s-orientable, since there exists an FT-MPDAG in which this edge is unoriented, as illustrated in (b).

We emphasize that the theorem above is *sound and complete* in the following sense: if an edge between X_t and Y_t is s-orientable according to the theorem, then there cannot be an unoriented edge between X_t and Y_t in any FT-MPDAG compatible with the SCG; conversely, if the edge is not s-orientable according to the theorem, there exists at least one FT-MPDAG compatible with the SCG where the edge between X_t and Y_t is unoriented. In other words, in these cases we cannot guarantee orientability prior to applying a causal discovery algorithm such as tPC. However, this does not imply that, in cases of non-orientability, applying the tPC algorithm will *always* fail to orient the edge. It is possible that there exists a specific distribution for which the edge is orientable, and if tPC is applied to data generated from this distribution, the edge may in fact be correctly oriented. This is clarified by the following proposition.

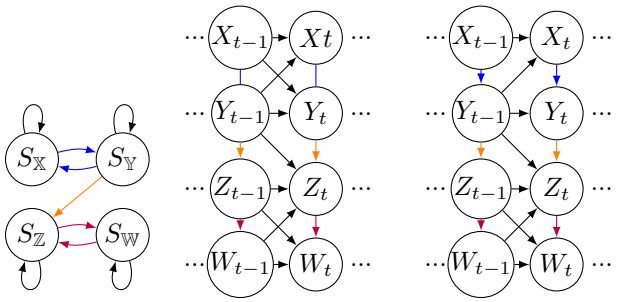
Proposition 1. *Let $\mathcal{G}^{\mathsf{s}}$ be an SCG. Under Assumptions 1, 2 and 3, $\exists \mathcal{G} \in compatible(\mathcal{G}^{\mathsf{s}})$ s.t. status$(X_t, Y_t; \mathcal{G})$ is orientable from $(\mathcal{G}^{\mathsf{s}}, \mathbb{P})$ where $\mathbb{P}$ is the set of distributions compatible with $\mathcal{G}$.*

Proof. Consider the trivial case without instantaneous relations which contains only directed edges oriented by time. □

Finally, we emphasize that non s-orientable cases are rare, constituting only a small fraction of all SCG configurations of a given size: among all possible SCGs of 5 nodes, less than 2% are not fully s-orientable (details in supplementary material).

4 IMPLICATIONS ON CAUSAL EFFECTS IDENTIFICATION

The previous section addressed the problem of identifying the orientation of an edge in an FT-MPDAG. This provides a qualitative understanding of the relationship between two variables, indicating whether $X_{t-\gamma}$ causes Y_t, whether the reverse holds, whether



(a) SCG $\mathcal{G}^{\mathbb{s}}$. (b) FT-MPDAG $\mathcal{C}_1$. (c) FT-MPDAG $\mathcal{C}_2$.

Figure 4: SCG $\mathcal{G}^{\mathbb{s}}$ with two examples of FT-MPDAG illustrating Theorem 1. Orange and purple edges are s-orientable and thus oriented in the FT-MPDAGs, while blue edge is non-s-orientable and may remain unoriented as in (b).

there is no direct causal relation between them, or whether the orientation cannot guaranteed to be determined from the SCG and a faithful distribution. In this section, we shift our focus to the quantitative implications of these results, namely how they inform the identification and estimation of causal effects. SCMs are particularly useful because they allow for the quantification of causal effects from observational data. There are many types of causal effects; in this section, we focus on two central ones: the total effect [Pearl, 1995] and the controlled direct effect [Pearl, 2001]. The *total effect* of $X_{t-\gamma}$ on Y_t captures the overall causal influence transmitted both directly and through mediating variables, whereas the *direct effect* isolates the portion of this influence that is not mediated by other variables. Formally, the total effect on Y_t of an intervention on $X_{t-\gamma}$ changing from x_0 to x_1 is defined as $\mathbb{E}[Y_t \mid \mathrm{do}(X_{t-\gamma} = x_1)] - \mathbb{E}[Y_t \mid \mathrm{do}(X_{t-\gamma} = x_0)]$, and the controlled direct effect is defined as $\mathbb{E}[Y_t \mid \mathrm{do}(X_{t-\gamma} = x_1), \mathrm{do}(\mathbb{S}_{Y_t,X_{t-\gamma},\mathcal{G}} = \mathbb{s}_{Y_t,X_{t-\gamma},\mathcal{G}})] - \mathbb{E}[Y_t \mid \mathrm{do}(X_{t-\gamma} = x_0), \mathrm{do}(\mathbb{S}_{Y_t,X_{t-\gamma},\mathcal{G}} = \mathbb{s}_{Y_t,X_{t-\gamma},\mathcal{G}})]$, where $\mathbb{S}_{Y_t,X_{t-\gamma},\mathcal{G}} = Pa(Y_t, \mathcal{G}) \setminus \{X_t\}$ and $\mathrm{do}(\cdot)$ denotes an intervention. For brevity, we will denote the total effect and the controlled direct effect, respectively, as $\Pr(y_t \mid \mathrm{do}(x_{t-\gamma}))$ and $\Pr(y_t \mid \mathrm{do}(x_{t-\gamma}), \mathrm{do}(\mathbb{s}_{Y_t,X_{t-\gamma},\mathcal{G}}))$.

A causal effect is said to be *identifiable* when it can be uniquely computed from any distribution compatible with an FT-DAG. In other words, identifying such a causal effect consists rewriting the causal effect using a do-free expression (only in terms of observed variables without any need to perform a real intervention). When the FT-DAG is specified and there are no unobserved confounders (Assumption 1), causal effects are always identifiable from the FT-DAG [Pearl, 2009]. In practice, however, it is often easier to obtain the SCG (for instance, from expert knowledge) than the full FT-DAG, since the SCG represents a higher level of abstraction that does not require knowledge of the precise causal relations between variables at each time step. From the SCG, causal effects can still be identified [Assaad et al., 2024, Ferreira and Assaad, 2024], but due to the presence of cycles, some causal effects may remain non-identifiable solely from the SCG. Moreover, the identification conditions in SCGs are intricate and typically require examining multiple paths.

For these reasons, one may consider using causal discovery techniques to estimate the FT-DAG. Given that the SCG is available, it can naturally be incorporated as background knowledge in the discovery process. However, as discussed earlier, constraint-based procedures do not recover the FT-DAG directly, but rather a partially directed graph, the FT-MPDAG. Because the FT-MPDAG may contain unoriented edges, causal effects are not always identifiable from it. Nevertheless, in some cases it has been shown that causal effects are identifiable from the FT-MPDAG [Perkovic, 2020, Flanagan, 2020]. This raises the following question: given the SCG, can we guarantee that the causal effect will be identifiable once a causal discovery step is performed? The following propositions provide an answer to this question.

Proposition 2 (Identifiability of total effects using s-orientability)**.** *Let $\mathcal{G}^{\mathbb{s}}$ be an SCG. Under Assumptions 1, 2 and 3, if for every $S_{\mathbb{Z}} \in Ne(S_{\mathbb{X}}, \mathcal{G}^{\mathbb{s}})$, $\forall \mathcal{G} \in compatible(\mathcal{G}^{\mathbb{s}})$, $status(Z_t, X_t; \mathcal{G})$ is s-orientable, then $P(Y_t \mid do(x_{t-\gamma}))$ is identifiable.*

Proof. The result follows directly from the backdoor criterion [Pearl, 2009], since under Assumption 1 the parents of the treatment form a valid adjustment set. Under the s-orientability condition, all edges connected to the treatment can be oriented, allowing recovery of its parents and identification of the total effect. □

Proposition 3 (Identifiability of controlled direct effects using s-orientability)**.** *Let $\mathcal{G}^{\mathbb{s}}$ be an SCG. Under Assumptions 1, 2 and 3, if for every $S_{\mathbb{Z}} \in Ne(S_{\mathbb{Y}}, \mathcal{G}^{\mathbb{s}})$, $\forall \mathcal{G} \in compatible(\mathcal{G}^{\mathbb{s}})$, $status(Z_t, Y_t; \mathcal{G})$ is s-orientable, then $\Pr(y_t \mid do(x_{t-\gamma}), do(\mathbb{s}_{Y_t,X_{t-\gamma},\mathcal{G}}))$ is identifiable.*

Proof. The result follows directly from Theorem 5.4 of Flanagan [2020], which states that a controlled direct effect is identifiable from the FT-MPDAG if and only if all edges adjacent to the outcome are oriented. Under the s-orientability condition, these edges can be oriented, allowing recovery of the relevant adjustment set and identification of the controlled direct effect. □

These results provide theoretical guarantees for causal effect identification, even in cases where identification theorems applied to the SCG alone [Assaad et al.,

2024, Ferreira and Assaad, 2024] do not succeed. Moreover, they offer simpler criteria that can be verified directly on the SCG, relying only on pairs of nodes rather than complex path-based conditions.

5 DISCUSSION

In this paper, we demonstrated that the background knowledge encoded in an SCG makes it possible to determine which edges are guaranteed to be oriented after applying a causal discovery algorithm to a faithful and causally sufficient distribution. At first glance, this may appear straightforward, for instance, an SCG without bidirected edges naturally enables the orientation of many edges that remain ambiguous in the output of causal discovery. However, the surprising and more insightful aspect of our results is that even bidirected edges in an SCG can provide valuable information. We showed that it is possible to guarantee the orientation of an edge at the micro level (e.g., between X_t and Y_t) even in cases where $S_{\mathbb{X}} \leftrightarrows S_{\mathbb{Y}}$ is in the SCG, provided that at least one of these macro vertices does not have a self-loop, or there exists a parent of $S_{\mathbb{X}}$ that is not a parent of $S_{\mathbb{Y}}$. Moreover, we established that these insights allow us to simplify existing identifiability conditions in the literature for quantifying causal effects, by leveraging SCGs in combination with access to a faithful distribution compatible with the SCG.

In this work, we assumed that the SCG is provided as background knowledge (e.g., by an expert). However, we note that it may also be obtained using an SCG discovery algorithm [Assaad et al., 2022b, Wahl et al., 2023, 2024, Ninad et al., 2025]. For instance, suppose that a coarse-grained SCG has already been learned beforehand, and that we now seek to perform a fine-grained causal discovery of the FT-MPDAG. In this case, the previously obtained SCG can be used in the same way as expert knowledge to assess the a priori orientability guarantees established by our results. The key point is that the procedure used to obtain an SCG does not affect our results, as long as the provided SCG is considered to represent the true underlying dynamic SCM.

A key limitation of this work concerns the strong assumptions generally required by the PC algorithm in causal discovery. Moreover, the stationarity assumption (Assumption 3) is crucial for our results, as it ensures (together with the definition of SCGs) the existence of certain edges in the FT-DAG that follow from the SCG. For instance, if $S_{\mathbb{X}} \rightleftarrows S_{\mathbb{Y}}$ is in the SCG, we assume that there necessarily exists an edge from some past or present instance of $\{Y_{t_0}, \cdots, Y_{t_{\max}}\}$ pointing to X_t and an edge from $\{X_{t_0}, \cdots, X_{t_{\max}}\}$ pointing to Y_t. Moreover, stationarity guarantees that orientations remain consistent when shifting the time lag, thereby ruling out situations where, for example, $X_{t-1} \to Y_{t-1}$ and $X_t \leftarrow Y_t$ coexist in the same FT-DAG, which would invalidate parts of our proofs. While this assumption is necessary for our theoretical results, it may appear restrictive in applications. It should be noted, however, that this is a standard assumption adopted by most causal discovery algorithms in time series settings [Peters et al., 2013, Malinsky and Spirtes, 2018, Runge et al., 2019, Runge, 2020, Assaad et al., 2022a]. In addition, a weaker version can be considered, where stationarity is assumed only on a subinterval $\mathcal{T} \subset [t_0, t_{\max}]$ rather than on the entire interval. This is feasible provided that multiple repetitions of the time series are available to perform causal discovery. Otherwise, full stationarity must be assumed, in which case the series can be partitioned into subsamples and causal discovery can be carried out on these subsampled segments.

Another strong assumption arises in the definition of the SCG: we assume that an edge $S_{\mathbb{X}} \to S_{\mathbb{Y}}$ exists in the SCG if and only if there exist times t' and t s.t. $X_{t'} \to Y_t$ in the FT-DAG. Since our proofs rely on this assumption, we do not allow cases where an effect is posited in the SCG but absent from the FT-DAG. Consider a concrete example: suppose that, a priori, there is an effect from some instance of $\{X_{t_0}, \cdots, X_{t_{\max}}\}$ to some instance of $\{Y_{t_0}, \cdots, Y_{t_{\max}}\}$, and that there also exists another instance where $\{Y_{t_0}, \cdots, Y_{t_{\max}}\}$ affects $\{X_{t_0}, \cdots, X_{t_{\max}}\}$, together with a self-loop on $S_{\mathbb{Y}}$. Such prior knowledge translates into the SCG depicted in Figure 2(d). Now, suppose that in reality the edge $S_{\mathbb{X}} \to S_{\mathbb{Y}}$ does not exist (i.e., the prior was incorrect), and instead we actually have $Y_t \to X_t$. In this case, causal discovery would identify the presence of an edge between X_t and Y_t, and by incorporating the prior knowledge, we would incorrectly orient it as $X_t \to Y_t$. Our results are therefore strong in the sense that they rely on having reliable prior knowledge encoded in the SCG, faithful to its definition. This highlights the importance of critically assessing the validity of SCG edges before reasoning with the results presented in this paper.

For future work, it would be interesting to extend these results to the FCI algorithm [Spirtes et al., 2001], an extension of the PC algorithm that does not rely on causal sufficiency. Moreover, while Propositions 2 and 3 provide sufficient conditions for identifying total and controlled direct effects, respectively, their completeness remains an open question. Investigating this aspect represents another promising direction for future research.

Acknowledgements

This work was supported by the CIPHOD project (ANR-23-CPJ1-0212-01) and by funding from the French government managed by the National Research Agency (ANR) under the France 2030 program (ANR-23-IACL-0007).

References

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 2022a.

Charles K. Assaad, Emilie Devijver, and Eric Gaussier. Entropy-based discovery of summary causal graphs in time series. *Entropy*, 24(8), 2022b. ISSN 1099-4300. doi: 10.3390/e24081156.

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.

Charles K. Assaad, Emilie Devijver, Eric Gaussier, Gregor Goessler, and Anouar Meynaoui. Identifiability of total effects from abstractions of time series causal graphs. In Negar Kiyavash and Joris M. Mooij, editors, *Proceedings of the Fortieth Conference on Uncertainty in Artificial Intelligence*, volume 244 of *Proceedings of Machine Learning Research*, pages 173–185. PMLR, 15–19 Jul 2024.

Ali Aït-Bachir, Charles K. Assaad, Christophe de Bignicourt, Emilie Devijver, Simon Ferreira, Eric Gaussier, Hosein Mohanna, and Lei Zan. Case studies of causal discovery from it monitoring time series, 2023. The History and Development of Search Methods for Causal Structure Workshop at the 39th Conference on Uncertainty in Artificial Intelligence.

Christine W. Bang and Vanessa Didelez. Do we become wiser with time? On causal equivalence with tiered background knowledge. In Robin J. Evans and Ilya Shpitser, editors, *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, pages 119–129. PMLR, 31 Jul–04 Aug 2023.

David Maxwell Chickering. Learning equivalence classes of bayesian-network structures. *J. Mach. Learn. Res.*, 2:445–498, March 2002. ISSN 1532-4435. doi: 10.1162/153244302760200696.

Diego Colombo and Marloes H. Maathuis. Order-independent constraint-based causal structure learning. *J. Mach. Learn. Res.*, 15(1):3741–3782, January 2014. ISSN 1532-4435.

Simon Ferreira and Charles K. Assaad. Identifiability of direct effects from summary causal graphs. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(18):20387–20394, Mar. 2024. doi: 10.1609/aaai.v38i18.30021.

Simon Ferreira and Charles K. Assaad. Identifying macro conditional independencies and macro total effects in summary causal graphs with latent confounding. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(25):26787–26795, Apr. 2025. doi: 10.1609/aaai.v39i25.34882.

Emily Flanagan. *Identification and Estimation of Direct Causal Effects*. PhD thesis, 2020.

Andreas Gerhardus and Jakob Runge. High-recall causal discovery for autocorrelated time series with latent confounders. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 12615–12625. Curran Associates, Inc., 2020.

Daniel Malinsky and Peter Spirtes. Causal structure learning from multivariate time series in settings with unmeasured confounding. In Thuc Duy Le, Kun Zhang, Emre Kıcıman, Aapo Hyvärinen, and Lin Liu, editors, *Proceedings of 2018 ACM SIGKDD Workshop on Causal Disocvery*, volume 92 of *Proceedings of Machine Learning Research*, pages 23–47. PMLR, 20 Aug 2018.

Christopher Meek. Causal inference and causal explanation with background knowledge. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, UAI'95, page 403–410, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc. ISBN 1558603859.

Urmi Ninad, Jonas Wahl, Andreas Gerhardus, and Jakob Runge. Causal discovery on vector-valued variables and consistency-guided aggregation, 2025.

Judea Pearl. Causal diagrams for empirical research. *Biometrika*, 82(4):669–688, 12 1995. ISSN 0006-3444. doi: 10.1093/biomet/82.4.669.

Judea Pearl. Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, UAI'01, page 411–420, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc. ISBN 1558608001.

Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, USA, 2nd edition, 2009. ISBN 052189560X.

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, Dominik Janzing, and Bernhard Schölkopf. Causal inference on time series using restricted structural equation models. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'13, page 154–162, Red Hook, NY, USA, 2013. Curran Associates Inc.

Jakob Runge. Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets. 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 1388–1397. PMLR, 03–06 Aug 2020.

Jakob Runge, Peer Nowack, Marlene Kretschmer, Seth Flaxman, and Dino Sejdinovic. Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, 5(11):eaau4996, 2019. doi: 10.1126/sciadv.aau4996.

Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, prediction, and search*. MIT press, 2001.

Thomas Verma and Judea Pearl. Equivalence and synthesis of causal models. In *Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence*, UAI '90, page 255–270, USA, 1990. Elsevier Science Inc. ISBN 0444892648.

Jonas Wahl, Urmi Ninad, and Jakob Runge. Vector causal inference between two groups of variables. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence*, AAAI'23/IAAI'23/EAAI'23. AAAI Press, 2023. ISBN 978-1-57735-880-0. doi: 10.1609/aaai.v37i10.26450.

Jonas Wahl, Urmi Ninad, and Jakob Runge. Foundations of causal discovery on groups of variables. *Journal of Causal Inference*, 12(1):20230041, 2024. doi: doi:10.1515/jci-2023-0041.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes]
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Not Applicable]
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [Not Applicable]

2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. [Yes]
   (b) Complete proofs of all theoretical results. [Yes]
   (c) Clear explanations of any assumptions. [Yes]

3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [Not Applicable]
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Not Applicable]
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [Not Applicable]
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [Not Applicable]

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. [Yes]
   (b) The license information of the assets, if applicable. [Not Applicable]
   (c) New assets either in the supplemental material or as a URL, if applicable. [Not Applicable]
   (d) Information about consent from data providers/curators. [Not Applicable]
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]

5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. [Not Applicable]
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not Applicable]
   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

Supplementary Materials

Empirical scarcity of non s-orientable cases

The structure of a non s-orientable pair in an SCG satisfying the assumptions is highly specific according to Theorem 1, and can therefore be expected to be rare. To gain an idea of its frequency, we enumerate all possible SCGs for a given number of nodes and check whether there exists at least one pair of nodes in the SCG that is non s-orientable (i.e., a pair of nodes satisfying Theorem 1). Since the number of possible SCGs with n nodes is huge ($2^{n \times n}$), we only report results up to $n = 5$. The results are summarized in Table 1.

n	# SCGs	# Not fully s-orientable	%
2	16	1	6.25
3	512	28	5.47
4	65,536	2,256	3.44
5	33,554,432	613,616	1.83

Table 1: Enumeration of all SCGs up to $n = 5$ nodes, showing the number of graphs that are not fully s-orientable.

These results already illustrate the rarity of such cases and show that, in the vast majority of theoretically possible SCGs, all edges are s-orientable using the background knowledge induced by the SCG. Moreover, we emphasize that these rare cases are non-s-identifiable from the SCG, but this does not imply that a causal discovery algorithm will fail to return an FT-DAG. Indeed, there is no a priori guarantee when considering only the SCG; however, in some cases (which may be numerous), it is still possible for an edge to be oriented in our specific setting. This is illustrated in Figure 4(c), where we observe a case in which a non s-orientable edge is nonetheless oriented.

Note, however, that while this observation suggests that non–fully s-orientable SCGs are relatively rare in a density sense over the space of all SCGs, one should keep in mind that such graphs may occur more frequently in practice, since real systems do not arise from a uniform sampling process. For instance, our computation includes SCGs with disconnected components, which are unlikely in realistic settings. This observation should therefore be understood only as a general indication rather than a practical estimate.