Causal reasoning in difference graphs

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

Editors: Biwei Huang and Mathias Drton

Abstract

Understanding causal mechanisms across different populations is essential for designing effective public health interventions. Recently, difference graphs have been introduced as a tool to visually represent causal variations between two distinct populations. While there has been progress in inferring these graphs from data through causal discovery methods, there remains a gap in systematically leveraging their potential to enhance causal reasoning. This paper addresses that gap by establishing conditions for identifying causal changes and effects using difference graphs. It specifically focuses on identifying total causal changes and total effects in a nonparametric setting, as well as direct causal changes and direct effects in a linear setting. In doing so, it provides a novel approach to causal reasoning that holds potential for various public health applications.

Keywords: Difference graphs, causal reasoning, total effect, direct effect, causal change.

1. Introduction

In epidemiology, it is critical to understand the causal mechanisms that impact health outcomes across diverse populations, which may differ based on geographic location, socioeconomic status, genetic predispositions, or environmental exposures. Recently, difference graphs have been introduced to represent differences between two Structural Causal Models (SCMs) (Pearl, 2009), with each model corresponding to a distinct population. These graphs offer a refined representation to visually compare and analyze the causal variations between these groups. By clearly delineating these causal differences, epidemiologists can tailor public health policies and interventions more effectively to address the unique needs of each population. Moreover, it has been highlighted by Malik et al. (2024) that such graphs are also interesting for root cause analysis (Assaad et al., 2023).

Recently, there has been a surge of interest in constructing difference graphs directly from data through methods rooted in causal discovery (Wang et al., 2018; Chen et al., 2023; Malik et al., 2024; Bystrova et al., 2024). The simplest approach involves independently discovering the SCM for each of the two populations and then computing the difference graph. However, a more effective strategy has emerged that involves directly constructing the difference graph, bypassing the two SCMs that define it. This approach reduces the need for stringent assumptions commonly associated with causal discovery (Spirtes et al., 2000; Glymour et al., 2019). Despite these advancements, an evident gap remains in the existing literature: there has been no systematic investigation into how difference graphs themselves can facilitate or enhance the process of causal reasoning. This oversight presents a unique opportunity to explore the potential of difference graphs as a tool not just for representation but also for driving identifiability of causal changes and causal effects.

To bridge this gap, this work establishes both sufficient and necessary conditions for identifying total causal changes and total effects in difference graphs within nonparametric models, under the assumption of no hidden confounding. Additionally, this work provides sufficient and necessary

© 2025 C.K. Assaad.

conditions for identifying direct causal changes and direct effects in difference graphs within linear models, under the assumption of no hidden confounding. Initially, both results are presented under the assumption that the two causal graphs, induced by the two SCMs defining the difference graph, share the same topological ordering. These results are then extended to accommodate cases where the graphs have different topological orderings.

In the remainder, Section 2 introduces difference graphs along with necessary tools and terminologies. Section 3 details the identifiability results for total effects. Section 4 focuses on the identifiability of direct effects. Section 5 presents related works and offers further discussion. Finally, Section 6 concludes the paper. In the following, uppercase letters represent variables and lowercase letters represent their values, while letters in blackboard bold denote sets.

2. Difference graphs, causal changes, and causal effects

This study utilizes the framework of structural causal models (SCMs) introduced by Pearl (2009). Formally, an SCM $\mathcal{M}$ is defined as a 4-tuple $(\mathbb{U}, \mathbb{V}, \mathbb{F}, P(\mathbb{U}))$, where $\mathbb{U}$ represents a set of exogenous (latent) variables and $\mathbb{V}$ denotes a set of endogenous (observed) variables. The set $\mathbb{F}$ consists of functions $\{f_i\}_{i=1}^{|\mathbb{V}|}$, where each endogenous variable is determined by a function $f_i \in \mathbb{F}$ that depends on its corresponding exogenous variable $U_i \in \mathbb{U}$ and a subset of endogenous variables (which are also called the direct causes or the parents). In a linear SCM, f_i is a linear equation where each coefficient is called path coefficient (Wright, 1920, 1921) and it designates a direct effect. The following assumption regarding exogenous variables is made throughout the paper.

Assumption 1 *All exogenous variables are mutually independent, i.e., no hidden confounding.*

Consistent with common practice in causal inference, it is also assumed that each SCM $\mathcal{M}$ induces a causal graph $\mathcal{G} = (\mathbb{V}, \mathbb{E})$ which is assumed to be a directed acyclic graph (DAG). For clarity, any causal graph that is a DAG will be referred to as a causal DAG. In this graph, $\mathbb{V}$ represents the set of vertices, each corresponding to a random variable, and $\mathbb{E}$ includes the directed edges, with each edge $X \to Y$ indicating a causal relationship between X and Y. For any graph $\mathcal{G} = (\mathbb{V}, \mathbb{E})$, standard terminology is adopted to describe relations to any vertex $Y \in \mathbb{V}$: $Pa(Y, \mathcal{G})$ denotes the set of vertices from which there is a direct edge to Y; $An(Y, \mathcal{G})$ refers to the ancestors of Y which is the collective set including Y and the parents of all vertices that can be recursively traced back from Y; $De(Y, \mathcal{G})$ refers to the descendants of Y which comprises all vertices X for which Y is found within $An(X, \mathcal{G})$. If a path contains the subpath $X \to W \leftarrow Y$, then W is a *collider* on the path. A path from X to Y is considered *active* given a vertex set $\mathbb{W}$, where neither X nor Y are in $\mathbb{W}$, if every non-collider on the path is not in $\mathbb{W}$, and every collider on the path has a descendant in $\mathbb{W}$. Otherwise, the path is *blocked* by $\mathbb{W}$. It is important to note that many SCMs can induce the same causal DAG and that a causal DAG can be compatible with many probability distributions.

Recently, difference graphs have been introduced as a tool to represent the differences between two SCMs, each corresponding to a distinct population, offering a refined method to visually compare causal variations between these groups.

Definition 1 (Difference graphs) *Consider an unordered pair of SCMs ($\mathcal{M}^1 = (\mathbb{V}, \mathbb{U}, \mathbb{F}^1, \Pr^1)$, $\mathcal{M}^2(\mathbb{V}, \mathbb{U}, \mathbb{F}^2, \Pr^2)$) that respectively induce the two causal DAGs $\mathcal{G}^1$ and $\mathcal{G}^2$ and that are respectively compatible with the two distributions $\Pr^1(\mathbb{v})$ and $\Pr^2(\mathbb{v})$. A* difference graph $\mathcal{D} = (\mathbb{V}, \mathbb{E}^{|1-2|})$ *between $\mathcal{M}^1$ and $\mathcal{M}^2$ is a directed graph where the set of vertices is identical of the set of endogenous variables in $\mathcal{M}^1$ and $\mathcal{M}^2$ and the set of edges is defined as follows:* $\mathbb{E}^{|1-2|} := \{X \to Y \mid$

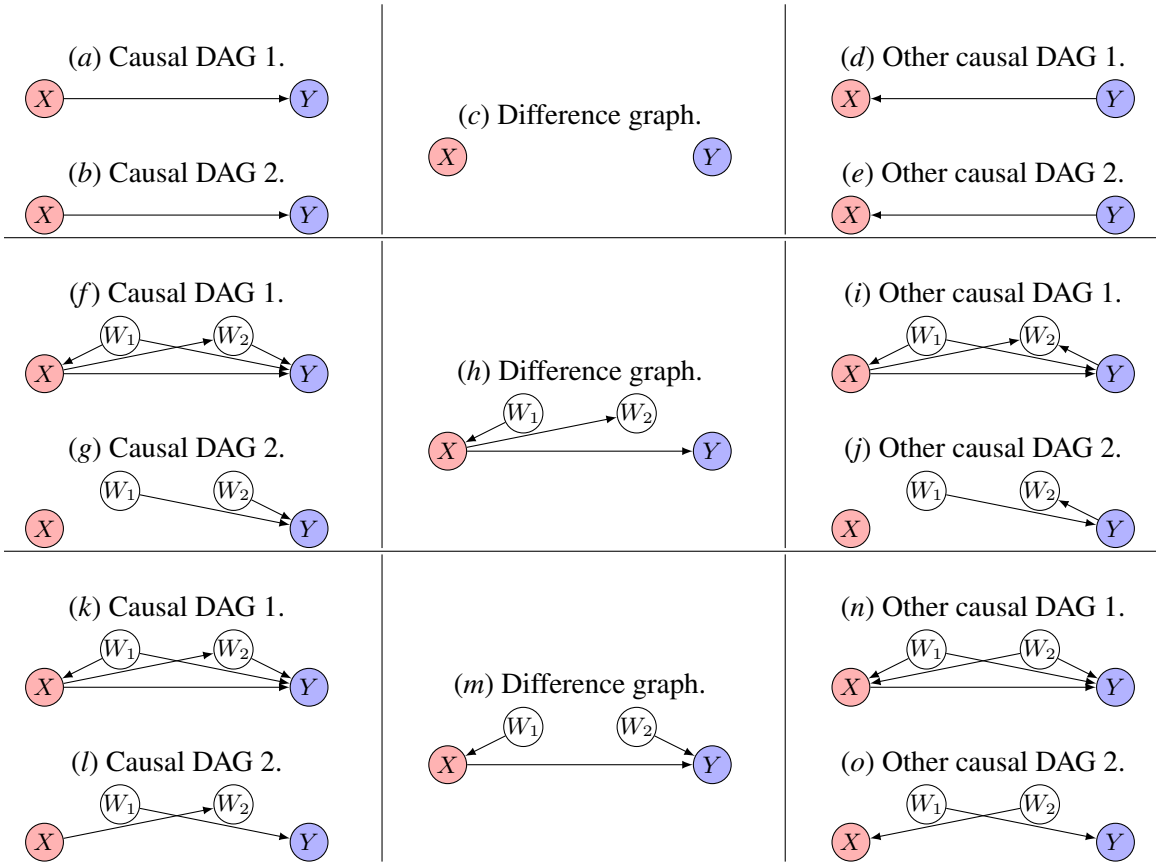

Figure 1: Three difference graphs ((c), (h), and (m)), each associated with two pairs of causal DAGs (one pair on the left and one on the right). The two causal DAGs in each pair share the same topological ordering, ensuring that Assumption 3 is satisfied. In each subfigure, red and blue vertices represent the cause and effect of interest. In (c), neither the total nor the direct effect are identifiable. In (h), only the total effect is identifiable, while in (m), only the direct effect is identifiable.

$\forall X, Y \in \mathbb{V}$ *if the direct effect of* X *on* Y *in* $\mathcal{M}^1$ *is different than the direct effect*[1] *of* X *on* Y *in* $\mathcal{M}^2$*, i.e., if there is a mechanism change (Tian and Pearl, 2001) at* Y *with respect to* X*.*

Note that multiple unordered pairs of SCMs can define the same difference graph. In Definition 1, difference graphs are defined using SCMs rather than causal DAGs because causal variations may not always be structurally evident by comparing the two causal DAGs, since changes are often parametric rather than structural. That said, Figures 1 and 2, feature six examples of difference graphs ((c), (h), (m) in Figure 1 and (c), (f), (k) in Figures 2) along with pairs of causal DAGs on both side of each difference graph that can also define the same difference graph (indicating structural differences[2] between the two populations). This selection was made intentionally in order to make the concept of difference graphs visually clearer and easier to understand. Nonetheless, it is important to remember that the theoretical results of this paper apply even when the difference

1. In a linear model, the direct effects are reduce to path coefficients (Wright, 1920, 1921) denoted as $\alpha_{x,y}$.
2. The term "structural differences" refers to "topological differences between the two causal DAGs".

between the two populations is not structurally apparent. However, all those results are grounded in the following assumption regarding the distributions that are compatible with the difference graph.

Assumption 2 $\Pr^1(\mathbb{v})$ *and* $\Pr^2(\mathbb{v})$ *are positive distributions.*

As shown in Figure 2, generally, there is no inherent necessity to impose similarity constraints on two causal DAGs and thus difference graphs can contain cycles. However, to make the results of this paper more comprehensible and consistent with recent literature on difference graphs, I initially assume that the two causal DAGs have the same topological ordering.

Assumption 3 $\mathcal{G}^1$ *and* $\mathcal{G}^2$ *share the same topological ordering.*

This assumption implies that the difference graph is also a DAG (Wang et al., 2018) and it can hold in many applications, for example, in biological pathways an upstream gene does not generally become a downstream gene in different conditions (Wang et al., 2018). However, in some domains, it is possible to imagine scienarios where causal relations can change direction under different conditions. For example, in epidemiology, obesity often leads to the development of type 2 diabetes in the general population, with no direct feedback from diabetes to obesity. However, in a subpopulation of individuals taking certain antipsychotic medications, the onset of type 2 diabetes can lead to significant weight gain due to the medication's side effects, without obesity being the original cause. Similarly, in nephrology, diabetes can lead to chronic kidney disease, as high blood sugar damages the kidneys in the general population. However, in patients with end-stage kidney disease, impaired kidney function can cause insulin resistance, leading to diabetes, without diabetes further worsening kidney disease. As a result, for each finding presented in this paper under Assumption 3, I will also provide an alternative result that extends to cases where Assumption 3 is not satisfied.

Recently, there has been increasing interest in discovering difference graphs directly from data, without first discovering the two SCMs that define them (Wang et al., 2018; Malik et al., 2024; Bystrova et al., 2024). However, a key question that has yet to be explored, but naturally arises, is whether these graphs can provide conditions for estimating causal changes between two populations using observational data. This work considers two types of causal changes: total causal change and direct causal change, also referred to as mechanism change (Tian and Pearl, 2001). The total causal change of X on Y is defined as the difference between the total effect of X on Y in $\mathcal{M}^1$ and the total effect of X on Y in $\mathcal{M}^2$. In a nonparametric setting, the total effects is a specific functional of the interventional distribution $\Pr(Y = y \mid \text{do}(X = x))$ (where the $do()$ operator represents the intervention) for different values of x. However, for simplicity, the total effect is often reffered to as $\Pr(Y = y \mid \text{do}(X = x))$ and by a slight abuse of notation, in the remainder, the total effect will be denoted as $\Pr(y \mid \text{do}(x))$. Similarly, the direct causal change is the difference between the direct effect of X on Y in $\mathcal{M}^1$ and the direct effect of X on Y in $\mathcal{M}^2$. In a nonparametric setting, the direct effect can be defined as a specific functional of $\Pr(y \mid do(x), do(\mathbb{S}_{x,y}))$ for different values of x, known as the controlled direct effect (Pearl, 2009), where $\mathbb{S}_{x,y}$ represents the set of all observed variables in the system except X and Y[3]. The controlled direct effect of X on Y is not uniquely defined as it can change with respect to the values of $\mathbb{S}_{x,y}$ (some variables in $\mathbb{S}_{x,y}$ are known as effect modifier). For instance, the direct effect of a pill on thrombosis is likely different for pregnant and non-pregnant women (Pearl, 2009). However, in a linear SCM, i.e, $\forall f_i \in \mathbb{F}$ f_i, is a linear function,

3. Another approach to defining a direct effect in a nonparametric setting is through the natural direct effect (Pearl, 2001), which relies on counterfactual notions rather than the do() operator. However, since it cannot be defined purely using interventional concepts, it is considered beyond the scope of this paper.

the direct effect, obtained through differentiation, corresponds to the familiar path coefficient $\alpha_{x,y}$ (Wright, 1920, 1921), which, unlike the non-parametric controlled direct effect, is uniquely defined. Therefore, in this work, whenever direct effect is discussed, it will refer specifically to linear SCMs.

Since the total causal change of X on Y is a functional of the total effect, $\Pr(y \mid \mathrm{do}(x))$, and the direct causal change is a functional of the direct effect, $\alpha_{x,y}$, the total causal change is identifiable if the total effect is identifiable, and likewise, the direct causal change is identifiable if the direct effect is identifiable[4]. Therefore, the focus of the remainder of this paper is to establish the identifiability of both the total and direct effects using difference graphs. I begin by providing a general definition of the identifiability of causal effects using difference graphs, which is adapted from the definition given by Pearl (2009) for the identifiability of causal effects using causal DAGs.

Definition 2 (Causal effect identifiability using difference graph) *Let X and Y be distinct vertices in a difference graph $\mathcal{D} = (\mathbb{V}, \mathbb{E}^{|1-2|})$. The causal effect Q of X on Y is identifiable in $\mathcal{D}$ if Q is uniquely computable from any observational distributions consistent with $\mathcal{D}$.*

Given a causal DAG, it is possible to identify $\Pr(y \mid \mathrm{do}(x))$ using the back-door criterion and to identify $\alpha_{x,y}$ using the single-door criterion which is only applicable in a linear setting.

Definition 3 (Back-door criterion, Pearl (1995)) *A set of variables $\mathbb{W}$ satisfies the back-door criterion relative to the total effect $\Pr(y \mid do(x))$ in a causal DAG $\mathcal{G}$ if $\forall W \in \mathbb{W}$, $W \notin De(X, \mathcal{G})$ and $\mathbb{W}$ blocks every path between X and Y which contains an arrow into X.*

Definition 4 (Single-door criterion, Spirtes et al. (1998); Pearl (1998, 2009)) *A set of variables $\mathbb{W}$ satisfies the single-door criterion relative to the direct effect $\alpha_{x,y}$ in a causal DAG $\mathcal{G}$ if $\forall W \in \mathbb{W}$, $W \notin De(Y, \mathcal{G})$ and $\mathbb{W}$ blocks every path between X and Y.*

Whenever a set $\mathbb{W}$ satisfies the back-door criterion[5] in a causal DAG, $\Pr(y \mid \mathrm{do}(x))$ is given by the adjsutment formula $\sum_{\mathbb{w}} \Pr(y \mid x, \mathbb{w}) \Pr(\mathbb{w})$. Whenever a set $\mathbb{W}$ satisfies the single-door criterion in a causal DAG, $\alpha_{x,y}$ is given by the regression coefficient of X on Y after $\mathbb{W}$ is partialled out.

Since this work focuses on difference graphs, the first objective is to identify a set that satisfies the back-door criterion across all causal DAGs compatible with a given difference graph when estimating the total effect—if such a set exists, the total effect is said to be identifiable by a *common back-door*. Similarly, the second objective is to find a set that satisfies the single-door criterion across all compatible causal DAGs when estimating the direct effect—if such a set exists, the direct effect is said to be identifiable by a *common single-door*. However, these goals cannot be acheived by applying the back-door and single-door criteria directly to difference graphs, even when the difference graph is acyclic. For example, in the difference graph shown in Figure 1 (m), the sets $\varnothing$, $\{W_1\}$, $\{W_2\}$, and $\{W_1, W_2\}$ all satisfy the back-door criterion for the total effect $\Pr(y \mid \mathrm{do}(x))$. But when examining the pair of causal DAGs in (k) and (l) that are compatible with this difference graph, it becomes clear that W_1 is necessary and W_2 must be excluded from any set that satisfies the back-door criterion, ruling out the empty set and $\{W_2\}$. Additionally, by looking at another pair of causal DAGs in (n) and (o), it becomes evident that in this case, $\{W_2\}$ is required for the back-door

4. In some cases, a causal change can be identified even when the corresponding causal effect is not, e.g., if there is no directed path from X to Y, the causal change is clearly zero, regardless of whether the causal effect is identifiable.
5. The back-door criterion is in general not complete. However, when X is a singleton (as considered in this paper), an adjustment set exists iff there is at least one set that satisfies the back-door criterion (Perković et al., 2018).

criterion to hold. This illustrates that, in this example, no single set satisfies the back-door criterion for every causal DAG compatible with the difference graph. Similarly, for the difference graph in Figure 1 (h), the sets $\varnothing$, $\{W_1\}$, $\{W_2\}$, and $\{W_1, W_2\}$ all satisfy the single-door criterion for the direct $\alpha_{x,y}$. However, the pair of causal DAGs shown in (f) and (g), both W_1 and W_2 are needed in any set that satisfies the single-door criterion, whereas in the pair of causal DAGs shown in (i) and (j), W_2 must not be included in any such set. This demonstrates that, in this example, no single set satisfies the single-door criterion for all causal DAGs compatible with the difference graph.

In addition to examining identifiability by a common back-door (for the total effect) or by a common single-door (for the direct effect), I also consider cases where the total or direct effect is identifiable because it is evident from the graph that the causal effect does not exist. For example, in a DAG where Y causes X, the total and direct effects of X on Y are both identifiable: $\Pr(y \mid \text{do}(x))$ reduces to $\Pr(y)$, and $\alpha_{x,y}$ is zero. In such cases, the total or direct effect are referred to as being identifiable *trivially*. Now that all necessary terminologies and notations have been established, the two problems that this work aims to address can be formally presented.

Problem 1 *Consider a difference graph $\mathcal{D}$ defined by two* unknown *SCMs that respectively induce two* unknown *causal DAGs. Is $\Pr(y \mid do(x))$ identifiable in $\mathcal{D}$ trivially or by a common back-door?*

Problem 2 *Consider a difference graph $\mathcal{D}$ defined by two* unknown *linear SCMs that respectively induce two* unknown *causal DAGs. Is $\alpha_{x,y}$ identifiable in $\mathcal{D}$ trivially or by a common single-door?*

3. Identification of total effects using difference graphs in a nonparametric setting

This section first presents conditions for identifying total effects using difference graphs in a nonparametric setting under Assumption 3 and then generalizes this result by relaxing the assumption.

Theorem 5 *Consider a difference graph $\mathcal{D}$ compatible with two different SCMs. Under Assumptions 1, 2 and 3, $\Pr(y \mid do(x))$ is identifiable in $\mathcal{D}$ trivially or by a common back-door iff:*

A.1 *$Y \in An(X, \mathcal{D})$; or*

A.2 *$X \in An(Y, \mathcal{D})$ and $\forall W \in \mathbb{V} \backslash \{X, Y\}$, $X \in An(W, \mathcal{D})$ or $W \in An(X, \mathcal{D})$.*

*Furthermore, if Condition **A.1** is satisfied then $\Pr(y \mid do(x)) = \Pr(y)$ and if Condition **A.2** is satisfied then $\Pr(y \mid do(x)) = \sum_{\mathbb{w}^{an}} \Pr(y \mid x, \mathbb{w}^{an}) \Pr(\mathbb{w}^{an})$, where $\mathbb{W}^{an} = An(X, \mathcal{D}) \backslash \{X\}$.*

Proof Lemmas 6 and 7 respectively prove that Conditions **A.1** and **A.2** are sufficient for identifiability and Lemmas 6 and 8 prove that they are also necessary. ■

Before delving into the four lemmas that form the foundation of the proof for Theorem 5, it is important to highlight the key implications of the theorem. The theorem provides conditions under which the total effect can be identified in difference graphs by a common back-door. The identification of these total effects hinges on the structural properties of the difference graph.

In particular, when there is no directed path between the variables of interest—either from X to Y or from Y to X—as seen in the difference graph in Figure 1 (c), the absence of such a connection creates ambiguity in the directionality of the total effect. This is illustrated by the pair of causal DAGs in (a) and (b), where the association between X and Y reflects the total effect of interest,

and by the pair in (d) and (e), where the association between X and Y does not correspond to the total effect being sought. Therefore, the total effect cannot be identified using a common back-door. This is an essential point, as it highlights that causal reasoning in difference graphs is counterintuitive, with nonexistent directed paths in the difference graph potentially leading to ambiguous causal interpretations. The following lemma builds on this observation and in addition it shows that Condition **A.1** of the theorem is sufficient.

Lemma 6 *Under Assumptions 1, 2 and 3, if there exists a directed path from Y to X then $\Pr(y \mid do(x)) = \Pr(y)$ and if there exists no directed path from X to Y or from Y to X in $\mathcal{D}$ then $\Pr(y \mid do(x))$ is not identifiable.*

Proof If there exists a directed path from Y to X then by Assumption 3, there cannot exists a directed path from X to Y, which means that X does not have an effect on Y, *i.e.*, $\Pr(y \mid \mathrm{do}(x)) = \Pr(y)$. If there is no directed path from X to Y or from Y to X in $\mathcal{D}$ then there exists a possible pair of causal DAGs $(\mathcal{G}^1, \mathcal{G}^2)$ defining $\mathcal{D}$ such that $X \to Y$ in $\mathcal{G}^1$ and $\mathcal{G}^2$. In this case $\Pr(y \mid \mathrm{do}(x)) \neq \Pr(y)$. At the same time there exists another possible pair of causal DAGs $(\mathcal{G}^{1'}, \mathcal{G}^{2'})$ defining $\mathcal{D}$ such that $X \leftarrow Y$ in $\mathcal{G}^{1'}$ and $\mathcal{G}^{2'}$. In this case, $\Pr(y \mid \mathrm{do}(x)) = \Pr(y)$. ■

Condition **A.2** of Theorem 5 can be applied to identify $\Pr(y \mid \mathrm{do}(x))$ in the difference graph shown in Figure 1 (h), as $W_1 \in An(X, \mathcal{D})$ and $X \in An(W_2, \mathcal{D})$. This implies that for any pair of causal DAGs compatible with the difference graph—such as the pairs depicted in Figures 1 (f) and (g) or Figures 1 (i) and (j)—the set $\{W_1\}$ satisfies the back-door criterion. The following lemma illustrates that Condition **A.2** is indeed sufficient for identifiability.

Lemma 7 *Under Assumptions 1, 2 and 3, if there exists a directed path from X to Y and for each vertex $W \in \mathbb{V} \backslash \{X, Y\}$ there exists a directed path from X to W or from W to X in $\mathcal{D}$, then $\Pr(y \mid do(x))$ is identifiable in $\mathcal{D}$ by a common back-door.*

Proof Consider any pair of causal DAGs ($\mathcal{G}^1$, $\mathcal{G}^2$) that define $\mathcal{D}$. Let $\mathbb{W}^{an} = An(X, \mathcal{D}) \backslash \{X, Y\}$ and let $\mathbb{W}^{de} = De(X, \mathcal{D}) \backslash \{X, Y\}$. For any $W \in \mathbb{W}^{an}$, W cannot be in $\mathbb{W}^{de}$, otherwise $\mathcal{D}$ would not be acyclic and Assumption 3 would not be satisfied. By the lemma, for any $W \in \mathbb{V} \backslash \{X, Y\}$ if $W \notin \mathbb{W}^{an}$ then $W \in \mathbb{W}^{de}$. It follows that $\forall W \in \mathbb{W}^{an}$, W has to be an ancestor of X in at least one of the two causal DAGs $\mathcal{G}^1$ and $\mathcal{G}^2$ and W is not a descendent of X in any of the two causal DAGs. Thus $\mathbb{W}^{an}$ has to block all paths from X to Y which contains an arrow into X and at the same time $\mathbb{W}^{an}$ cannot contain any descendant of X. Therefore, $\mathbb{W}^{an}$ satisfies the back-door criterion (Pearl, 1995) in $\mathcal{G}^1$ and in $\mathcal{G}^2$ and by Pearl (1995, Theorem 1) the total effect $\Pr(y \mid \mathrm{do}(x))$ in $\mathcal{M}^1$ and in $\mathcal{M}^2$ is given by $\sum_{\mathbb{w}^{an}} \Pr(y \mid x, \mathbb{w}^{an}) \Pr(\mathbb{w}^{an})$. ■

Theorem 5 indicates non-identifiability if Conditions **A.1** and **A.2** are not met. For instance, Theorem 5 shows that $\Pr(y \mid \mathrm{do}(x))$ is not identifiable in the difference graph in Figure 1 (m). This is evident by considering two possible pairs of causal DAGs compatible with the difference graph: Figures 1 (k) and (l), and Figures 1 (n) and (m). In the first pair, $\{W_1\}$ satisfies the back-door criterion, but $\{W_1, W_2\}$ does not. In contrast, in the second pair, only $\{W_1, W_2\}$ satisfies the back-door criterion. Lemma 6 provided one part of the proof for the necessity of these two conditions, and the following lemma completes the other half, thus concluding the proof of the theorem.

Lemma 8 *Under Assumptions 1, 2 and 3, if there exists a directed path from X to Y and there exists a vertex $W \in \mathbb{V}\backslash\{X, Y\}$ for which there is no directed path from X to W or from W to X in $\mathcal{D}$ then $\Pr(y \mid do(x))$ is not identifiable in $\mathcal{D}$ by a common back-door.*

Proof Suppose there is a directed path from X to Y in $\mathcal{D}$. Now, consider a vertex W for which there is no directed path from or to X in $\mathcal{D}$. It follows that in any pair of causal DAGs $\mathcal{G}^1$ and $\mathcal{G}^2$, W can be either independent of X or a descendant of X or a parent of X or an ancestor, but not a parent, of X. Let us focus on the second and third cases. In the second case, W cannot be included in any set that satisfies the back-door criterion in every causal DAG compatible with $\mathcal{D}$, since W is a descendant of X. In the third case, W cannot be a descendant of Y; otherwise, a cycle would form involving X, W, and Y. Therefore, in at least one causal DAG compatible with $\mathcal{D}$, W must be a parent of Y, meaning W must be included in any set that satisfies the back-door criterion in every causal DAG compatible with $\mathcal{D}$. This third case contradicts the second case, making it impossible to find a set that satisfies the back-door criterion in all compatible causal DAGs. ■

The remaining part of this section investigates the case where the two causal DAGs induced by the two SCMs that define the difference graph do not share the same topological ordering (i.e., Assumption 3 is not satisfied). In this case the difference graph can contain cycles.

It is clear that if Assumption 3 is not satisfied, the first part of Lemma 6 does not hold. An example of this can be seen in Figure 2 (c), where the difference graph contains a directed path from Y to X, yet the total effect cannot be identified due to the simultaneous existence of a directed path from X to Y. Additionally, it is evident that if Assumption 3 is not met, Lemma 7 as a whole will fail. Furthermore, even though the second part of Lemma 6 and Lemma 8 remain valid individually, together they do not provide necessary conditions. However, as shown in the following theorem, minor modifications to these lemmas would allow acheiving correctness. The proof of this new theorem will simply highlight how the proofs of the previous lemmas should be modified.

Theorem 9 *Consider a difference graph $\mathcal{D}$ compatible with two different SCMs. Under Assumptions 1 and 2, $\Pr(y \mid do(x))$ is identifiable in $\mathcal{D}$ trivially or by a common back-door iff*

B.1 *Condition **A.1** is satisfied and $X \notin An(Y, \mathcal{D})$; or*

B.2 *Condition **A.2** is satisfied and $An(X, \mathcal{D}) \cap De(X, \mathcal{D}) = \{X\}$.*

*Furthermore, if Condition **B.1** is satisfied then $\Pr(y \mid do(x)) = \Pr(y)$ and if Condition **B.2** is satisfied then $\Pr(y \mid do(x)) = \sum_{\mathbb{W}^{an}} \Pr(y \mid x, \mathbb{W}^{an}) \Pr(\mathbb{W}^{an})$, where $\mathbb{W}^{an} = An(X, \mathcal{D})\backslash\{X\}$.*

Proof The proof is very similar to the proof of Theorem 5 with sime minor difference. First, notice that replacing Assumption 3 with the condition that there is no directed path from X to Y in the proof of Lemma 6 would maintain correctness. Second notice that replacing Assumption 3 with the condition that there is no cycle containing X and another vertex in the proof of Lemma 7 would maintain correctness. Finally, to extend Lemma 8, I add the condition that the total effect is not identifiable by a common back-door if a cycle exists on X. Combining this with the second part of Lemma 6 establishes that the conditions are necessary. This can be easily shown because if there is a cycle in $\mathcal{D}$ involving X and another vertex W (where $W \neq Y$), then there must exist a causal DAG in which W is a parent of X, as well as another causal DAG in which W is a descendant of X. In both of these graphs, W can also be a parent of Y. This implies that in the first DAG,

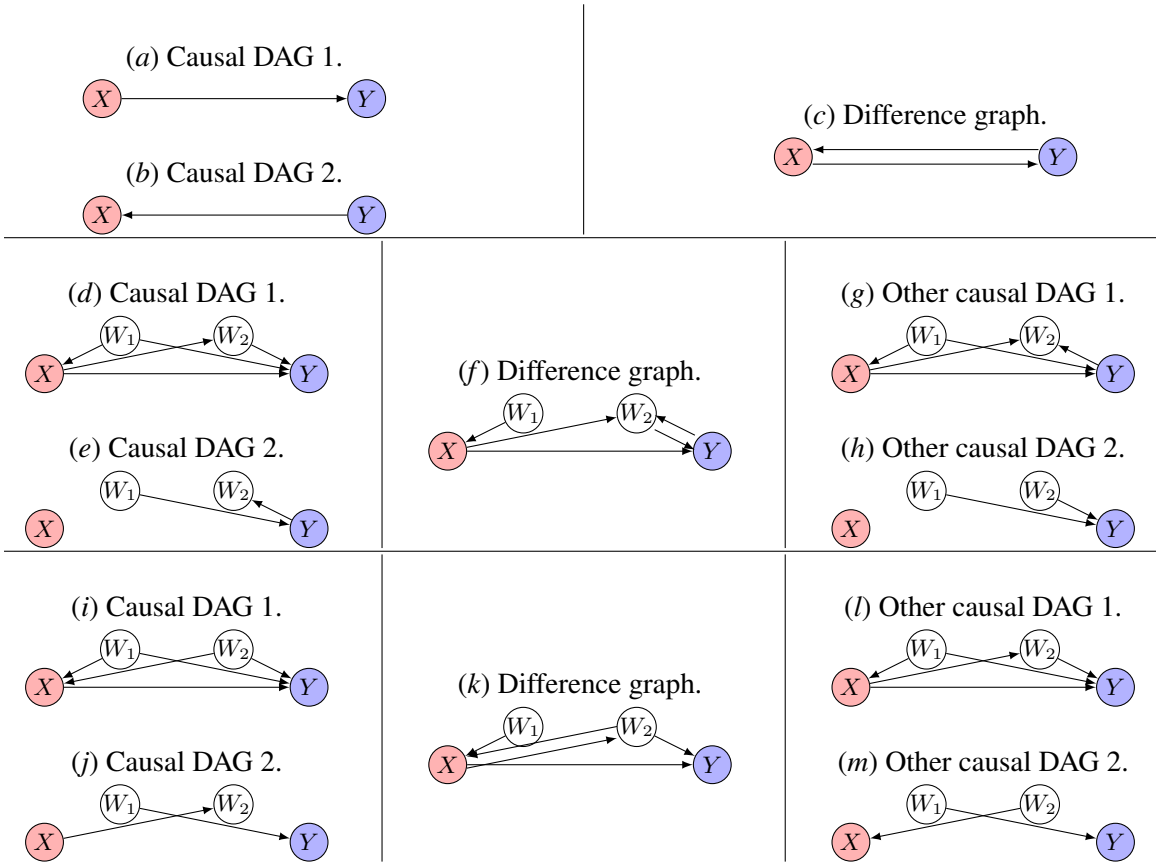

Figure 2: Three difference graphs ((c), (f), and (k)), each associated with two pairs of causal DAGs (one pair on the left and one on the right). The two causal DAGs in each pair do not share the same topological ordering, therefore Assumption 3 is not satisfied. In each subfigure, the red and blue vertices represent the cause and effect of interest, respectively. In the first difference graph (c), neither the total nor the direct effect is identifiable. In the second graph (f), only the total effect is identifiable, while in the third graph (k), only the direct effect is identifiable.

W must be included in any set satisfying the back-door criterion, while in the second DAG, W cannot be included in such a set. Therefore, no set can satisfy the back-door criterion for all causal DAGs compatible with $\mathcal{D}$. Moreover, if the cycle only involves X and Y, the total effect is also not identifiable for the same reason the total effect is not identifiable when there is no directed path from X to Y or from Y to X. ■

Theorem 9 shows that $\Pr(y \mid \mathrm{do}(x))$ is not identifiable in the difference graphs depicted in Figures 2 (c) and (k), but it is identifiable in the difference graph shown in Figure 2 (f).

Small simulation study for total causal changes. The adjustment set provided by Theorems 5 and 9 is not the only possible adjustment set that could be selected using the back-door criterion if the true pair of DAGs were known. However, as shown in Figure 3 (a), in practice, the set used in Theorems 5 and 9 performs as well as the set consisting of the parents of the exposure and outperforms the minimal set derived from the back-door criterion.

4. Mediation analysis using difference graphs in a linear setting

This section presents conditions for identifying the direct effect using a difference graph in a linear setting and as in the previous section, it starts by given conditions under Assumption 3 and then generalizes this result by relaxing the assumption.

Theorem 10 *Consider a difference graph $\mathcal{D}$ compatible with two different linear SCMs. Under Assumptions 1, 2 and 3, $\alpha_{x,y}$ is identifiable in $\mathcal{D}$ trivially or by a common single-door iff:*

C.1 *$Y \in An(X, \mathcal{D})$; or*

C.2 *$X \in An(Y, \mathcal{D})$ and $\forall W \in \mathbb{V} \backslash \{X, Y\}$, $Y \in An(W, \mathcal{D})$ or $W \in An(Y, \mathcal{D})$.*

*Furthermore, if Condition **C.1** is satisfied then $\alpha_{x,y} = 0$ and if Condition **C.2** is satisfied then $\alpha_{x,y} = r_{YX.\mathbb{W}^{an}}$ where $r_{YX.\mathbb{W}^{an}}$ is the regression coefficient of X on Y after $\mathbb{W}^{an}$ is partialled out.*

Proof Lemmas 11 and 12 respectively prove that Conditions **C.1** and **C.2** are sufficient for identifiability and Lemmas 11 and 13 prove that they are also necessary. ■

Theorem 10 establishes that the direct effect $\alpha_{x,y}$ is identifiable by a common single-door in the difference graph presented in Figure 1 (m), but it is not identifiable by a common single-door in the difference graphs shown in Figures 1 (c) and (h). The same reasoning applied in the previous section for the non-identifiability of the total effect in Figure 1 (c) also applies here for the direct effect. For the case of Figure 1 (m), it becomes clear why the direct effect is not identifiable when considering the two pairs of causal DAGs compatible with the difference graph, shown in Figures (f) and (g), and Figures (i) and (j). In the first pair, $\{W_1, W_2\}$ is the only set that satisfies the single-door criterion. However, in the second pair, W_2 cannot be included in any set that satisfies the single-door criterion. I now move on to present the lemmas required for the proof of the theorem, noting that their proofs follow the same reasoning as the lemmas introduced in the previous section.

Lemma 11 *Under Assumptions 1, 2 and 3, if there exists a directed path from Y to X then $\alpha_{x,y} = 0$ and if there exists no directed path from X to Y or from Y to X in $\mathcal{D}$ then $\alpha_{x,y}$ is not identifiable.*

Proof The proof follows directly from the proof of Lemma 6. ■

Lemma 12 *Under Assumptions 1, 2 and 3, if there exists a directed path from X to Y and for each vertex $W \in \mathbb{V} \backslash \{X, Y\}$ there exists a directed path from Y to W or from W to Y in $\mathcal{D}$, then $\alpha_{x,y}$ is identifiable in $\mathcal{D}$ by a common single-door.*

Proof Consider any pair of causal DAGs $(\mathcal{G}^1, \mathcal{G}^2)$ that define $\mathcal{D}$. Let $\mathbb{W}^{an} = An(Y, \mathcal{D}) \backslash \{X, Y\}$ and let $\mathbb{W}^{de} = De(Y, \mathcal{D}) \backslash \{X, Y\}$. For any $W \in \mathbb{W}^{an}$, W cannot be in $\mathbb{W}^{de}$, otherwise $\mathcal{D}$ would not be acyclic and Assumption 3 would not be satisfied. By the lemma, for any $W \in \mathbb{V} \backslash \{X, Y\}$ if $W \notin \mathbb{W}^{an}$ then $W \in \mathbb{W}^{de}$. It follows that $\forall W \in \mathbb{W}^{an}$, W has to be an ancestor of Y in at least one of the two causal DAGs $\mathcal{G}^1$ and $\mathcal{G}^2$ and W is not a descendent of Y in any of the two causal DAGs. Thus $\mathbb{W}^{an}$ has to block all paths between X and Y and at the same time $\mathbb{W}^{an}$ cannot contain any descendant of Y. Therefore, $\mathbb{W}^{an}$ satisfies the single-door criterion (Spirtes et al., 1998; Pearl, 1998) in $\mathcal{G}^1$ and in $\mathcal{G}^2$ and by Pearl (1998, Theorem 6) the direct effect $\alpha_{x,y}$ in $\mathcal{M}^1$ and in $\mathcal{M}^2$ is given by $r_{Y,X.\mathbb{W}^{an}}$. ■

Lemma 13 *Under Assumptions 1, 2 and 3, if there exists a directed path from X to Y and there exists a vertex $W \in \mathbb{V} \backslash \{X, Y\}$ for which there is no directed path from Y to W or from W to Y then $\alpha_{x,y}$ is not identifiable in $\mathcal{D}$ by a common single-door.*

Proof Suppose there is a directed path from X to Y. Now, consider a vertex W for which there is no directed path from or to Y in $\mathcal{D}$. It follows that in any pair of causal DAGs $\mathcal{G}^1$ and $\mathcal{G}^2$, W can be either independent of Y or a descendant of Y or a parent of Y or an ancestor, but not a parent, of Y. Let us focus on the second and third cases. In the second case, W cannot be included in any set that satisfies the single-door criterion in every causal DAG compatible with $\mathcal{D}$, since W is a descendant of Y. In the third case, it is possible to imagine that thre exists a causal DAG compatible with $\mathcal{D}$ where W is either a parent of a child of X. Therefore, in this causal DAG, W must be included in any set that satisfies the single-door criterion in every causal DAG compatible with $\mathcal{D}$. This third case contradicts the second case, making it impossible to find a set that satisfies the single-door criterion in all compatible causal DAGs. ■

The following theorem extends Theorem 10 to the case where the two causal DAGs, derived from the two SCMs that form the difference graph, do not have the same topological ordering.

Theorem 14 *Consider a difference graph $\mathcal{D}$ compatible with two different linear SCMs. Under Assumptions 1 and 2, $\alpha_{x,y}$ is identifiable in $\mathcal{D}$ trivially or by a common single-door iff:*

D.1 *Condition **C.1** is satisfied and $X \notin An(Y, \mathcal{D})$; or*

D.2 *Condition **C.2** is satisfied and $An(Y, \mathcal{D}) \cap De(Y, \mathcal{D}) = \{Y\}$.*

*Furthermore, if Condition **D.1** is satisfied then $\alpha_{x,y} = 0$ and if Condition **D.2** is satisfied then $\alpha_{x,y} = r_{YX.\mathbb{W}^{an}}$ where $r_{YX.\mathbb{W}^{an}}$ is the regression coefficient of X on Y after $\mathbb{W}^{an}$ is partialled out.*

Proof The proof closely follows the structure of the proof for Theorem 10, with a few key modifications. First, note that replacing Assumption 3 with the condition that there is no directed path from X to Y in the proof of Lemma 11 preserves its correctness. Second, replacing Assumption 3 with the condition that no cycle contains both Y and another vertex in the proof of Lemma 12 also maintains its validity. To extend Lemma 13, I introduce the condition that the direct effect cannot be identified using a common single-door criterion if there is a cycle involving Y. Combining this new condition with the second part of Lemma 11 demonstrates that these conditions of the Theorem are necessary. This can be shown as follows: if a cycle exists in $\mathcal{D}$ involving Y and another vertex W (where $W \neq X$), then there must be at least one causal DAG where W is a parent of Y and another causal DAG where W is a descendant of Y. In both of these DAGs, W can also act as either a parent or descendant of X. This implies that in the first DAG, W must be included in any set that satisfies the single-door criterion, while in the second DAG, W cannot be included in such a set. As a result, no single set can satisfy the single-door criterion across all causal DAGs compatible with $\mathcal{D}$. Furthermore, if the cycle only involves X and Y, the direct effect is also not identifiable for the same reason it is not identifiable when there is no directed path between X and Y. ■

Theorem 14 indicates that the direct effect $\alpha_{x,y}$ is not identifiable in the difference graphs shown in Figures 2 (c) and (f), but it is identifiable in the difference graph depicted in Figure 2 (k).

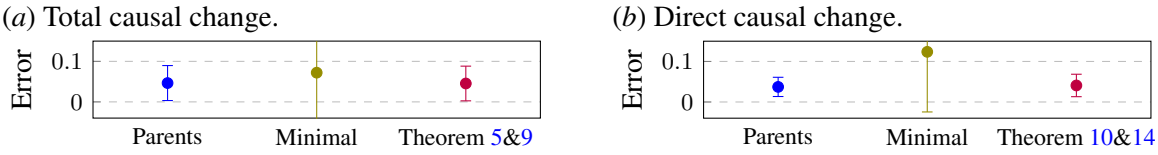

Figure 3: Mean absolute error and standard deviation for estimating the total and direct causal changes of X on Y. Subfigure (a) compares errors in total causal change estimation using (i) the set of parents of X, (ii) a minimal set satisfying the back-door criterion, and (iii) the set derived from Theorem 5 and 9 using a difference graph. Subfigure (b) evaluates direct causal change estimation using (i) the set of parents of Y, (ii) a minimal set satisfying the single-door criterion, and (iii) the set from Theorem 10 and 14. Results are reported over 100 pairs of dataset, consisting of 1000 observations and 10 variables, with each 10 pairs sharing the same difference graph for which the causal effect of interest is identifiable by our theorems. Each dataset is simulated using a linear SCM, where coefficients are randomly selected between 0 and 1, with Gaussian noise.

Small simulation study for direct causal changes. The adjustment set provided by Theorems 10 and 14 is not the only possible adjustment set that could be selected using the single-door criterion if the true pair of DAGs were known. So similarly, to Section 3, in Figure 3 (b), the set used in Theorems 10 and 14 is compared with the set consisting of the parents of the outcome and the minimal set derived from the single-door criterion. In this case, the set used in Theorems 10 and 14 performs slightly less well than the set of parents and better than the minimal set.

5. Related works and discussion

There has been extensive research on causal reasoning using graphs that represent a class of graphs. For example, Maathuis and Colombo (2013); Perkovic et al. (2016); Perkovic (2020); Jaber et al. (2022); Wang et al. (2023) extended causal reasoning to partially oriented graphs such as CPDAGs and PAGs. PAGs are similar to CPDAGs but allow for hidden confounding. Both CPDAGs and PAGs represent multiple causal DAGs that share the same skeleton, some common edge orientations, and the same topological ordering. Moreover, extensive research has been conducted for recovering these structures from observational data under specific assumptions (Spirtes et al., 2000; Glymour et al., 2019). From a different perspective, Iwasaki and Simon (1994); Chalupka et al. (2016); Anand et al. (2023); Assaad et al. (2023); Ferreira and Assaad (2024); Assaad et al. (2024) have examined causal reasoning using cluster graphs, where each vertex corresponds to a cluster of low-level variables, and each edge denotes a causal relationship between two clusters. A cluster graph can represent multiple causal DAGs but does not necessarily share the same skeleton as the true causal DAG and is not necessarily acyclic. Like CPDAGs and PAGs, cluster graphs can also be discovered from data under certain assumptions (Wahl et al., 2024). In addition, other works, such as Richardson (1997); Forré and Mooij (2020), have specifically focused on cyclic directed graphs, where vertices do not represent clusters. More closely related to difference graphs are selection diagrams (Bareinboim and Pearl, 2012) which can also represent differences between populations. There is a substantial body of literature on their use in causal reasoning (Bareinboim and Pearl, 2012) and methods for discovering them from data (Li et al., 2023). However, these two graphi-

cal representations serve different purposes. Selection diagrams provide more detailed information about stable causal relations and changes in noise, whereas difference graphs are more informative regarding changes in direct effects. In summary, to date, there has been no investigation into causal reasoning using difference graphs. The existing work in the literature cannot be directly applied to difference graphs because the absence of an edge in these graphs has a different interpretation than in other types of graphs.

This work addresses this gap and allow analysts to focus solely on the differences between two populations, simplifying reasoning, while still achieving identification for each domain individually. I emphasize that in this work, I assume that the difference graph is given and correct. The difference graph may be provided either by domain experts or by an algorithmic discovery method. In many domains, experts cannot construct a difference graph without knowing the true causal DAGs for both populations. However, in certain cases, this is feasible. For example, consider two cell populations: Population A (untreated) and Population B (treated with a drug that inhibits a specific protein). While the complete causal DAGs of these populations may be complex and unknown, domain knowledge about the drug's direct mechanism of action allows researchers to construct a difference graph representing the expected changes in causal structure. When domain experts cannot directly specify a difference graph, they may rely on algorithmic approaches to discover it from data. Some algorithms discover an equivalence class of graphs (Wang et al., 2018; Bystrova et al., 2024), while others aim to recover the exact difference graph (Chen et al., 2023). If the exact difference graph is known, all theorems intorduced in this paper, can be applied directly. For instance, consider the difference graph shown in (Chen et al., 2023, Figure 4 (a)), discovered from a real ovarian cancer dataset. Based on Theorem 5, the total effect of BIRC3 on PRKAR2B is not identifiable, whereas according to Theorem 10, the direct effect of BIRC3 on PRKAR2B is identifiable, assuming linearity. In contrast, when working with Markov equivalence classes of difference graphs, the applicability of the results of this paper depends on whether the conditions of theorems hold across all graphs in the equivalence class. For instance, consider a scenario with four variables: X, Y, W_1, and W_2. If every difference graph within a given Markov equivalence class includes the edges $X \leftarrow W_1$, $X \rightarrow W_2$, and $X \rightarrow Y$, then the conditions of Theorem 5 are satisfied across the entire equivalence class, ensuring that the total effect of X on Y is identifiable. Similarly, for the identification of the direct effect, suppose that in every difference graph within a given Markov equivalence class, the edges $W_1 \leftarrow Y$, $W_2 \rightarrow Y$, and $X \rightarrow Y$ are present. This guarantees that the conditions of Theorem 10 hold, allowing for the identification of $\alpha_{x,y}$.

6. Conclusion

This paper is the first to explore causal reasoning through the lens of difference graphs. While difference graphs may not be as intuitive as classical causal DAGs for identifying causal effects, this work establishes conditions for identifying total effects (and total causal changes) in a nonparametric setting and direct effects (and direct causal changes) in a linear setting, directly from the difference graph. For future works, it would be valuable to extend the results of this paper to direct effects in a nonparametric setting to path-specific effects. Additionaly, if it will be shown that difference graphs with hidden confounding are useful and can be learned or provided by experts, then it will be valuable to investigate the identifiability conditions in the presence of hidden confounding.

Acknowledgments

I thank Elias Bareinboim from Columbia University for the insightful discussions on difference graphs and selection diagrams. This work was supported by the CIPHOD project (ANR-23-CPJ1-0212-01).

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. doi: 10.1609/aaai.v37i10.26435.

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 *Proceedings of the Fourtieth Conference on Uncertainty in Artificial Intelligence*, Proceedings of Machine Learning Research. PMLR, 2024.

Elias Bareinboim and Judea Pearl. Transportability of causal effects: Completeness results. *Proceedings of the AAAI Conference on Artificial Intelligence*, 26(1):698–704, Sep. 2012. doi: 10.1609/aaai.v26i1.8232.

Daria Bystrova, Emilie Devijver, Vardan Manucharian, Julie Mondet, and Pascal Mossuz. Difference graph over two populations: Implicit difference inference algorithm. In *9th Causal Inference Workshop at UAI 2024*, 2024.

Krzysztof Chalupka, Frederick Eberhardt, and Pietro Perona. Multi-level cause-effect systems. In Arthur Gretton and Christian C. Robert, editors, *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pages 361–369, Cadiz, Spain, 09–11 May 2016. PMLR.

Tianyu Chen, Kevin Bello, Bryon Aragam, and Pradeep Ravikumar. iscan: Identifying causal mechanism shifts among nonlinear additive noise models. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 44671–44706. Curran Associates, Inc., 2023.

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.

Patrick Forré and Joris M. Mooij. Causal calculus in the presence of cycles, latent confounders and selection bias. In Ryan P. Adams and Vibhav Gogate, editors, *Proceedings of The 35th Uncertainty in Artificial Intelligence Conference*, volume 115 of *Proceedings of Machine Learning Research*, pages 71–80. PMLR, 22–25 Jul 2020.

Clark Glymour, Kun Zhang, and Peter Spirtes. Review of causal discovery methods based on graphical models. *Frontiers in Genetics*, 10, 2019. ISSN 1664-8021. doi: 10.3389/fgene.2019.00524.

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

Amin Jaber, Adele Ribeiro, Jiji Zhang, and Elias Bareinboim. Causal identification under markov equivalence: Calculus, algorithm, and completeness. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 3679–3690. Curran Associates, Inc., 2022.

Adam Li, Amin Jaber, and Elias Bareinboim. Causal discovery from observational and interventional data across multiple environments. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 16942–16956. Curran Associates, Inc., 2023.

Marloes Maathuis and Diego Colombo. A generalized backdoor criterion. *The Annals of Statistics*, 43, 07 2013. doi: 10.1214/14-AOS1295.

Vineet Malik, Kevin Bello, Asish Ghoshal, and Jean Honorio. Identifying causal changes between linear structural equation models. In *The 40th Conference on Uncertainty in Artificial Intelligence*, 2024.

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. Graphs, causality, and structural equation models. *Sociological Methods & Research*, 27(2):226–284, 1998. doi: 10.1177/0049124198027002004.

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.

Emilija Perkovic, Johannes Textor, Markus Kalisch, and Marloes H. Maathuis. Complete graphical characterization and construction of adjustment sets in markov equivalence classes of ancestral graphs. *J. Mach. Learn. Res.*, 18:220:1–220:62, 2016.

Emilija Perković, Johannes Textor, Markus Kalisch, and Marloes H. Maathuis. Complete graphical characterization and construction of adjustment sets in markov equivalence classes of ancestral graphs. *Journal of Machine Learning Research*, 18(220):1–62, 2018.

Thomas Richardson. A characterization of markov equivalence for directed cyclic graphs. *International Journal of Approximate Reasoning*, 17(2):107–162, 1997. ISSN 0888-613X. doi: https://doi.org/10.1016/S0888-613X(97)00020-0. Uncertainty in AI (UAI'96) Conference.

Peter Spirtes, Thomas Richardson, Christopher Meek, Richard Scheines, and Clark Glymour. Using path diagrams as a structural equation modeling tool. *Sociological Methods & Research*, 27(2): 182–225, 1998. doi: 10.1177/0049124198027002003.

Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search*. MIT press, 2nd edition, 2000.

Jin Tian and Judea Pearl. Causal discovery from changes. In *Conference on Uncertainty in Artificial Intelligence*, 2001.

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.

Tian-Zuo Wang, Tian Qin, and Zhi-Hua Zhou. Estimating possible causal effects with latent variables via adjustment. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 36308–36335. PMLR, 23–29 Jul 2023.

Yuhao Wang, Chandler Squires, Anastasiya Belyaeva, and Caroline Uhler. Direct estimation of differences in causal graphs. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 3774–3785, Red Hook, NY, USA, 2018. Curran Associates Inc.

Sewall Wright. The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. *Proceedings of the National Academy of Sciences of the United States of America*, 6(6):320–332, 1920. ISSN 00278424.

Sewall Wright. Correlation and causation. *Journal of agricultural research*, 20(7):557–585, 1921.