On the Identification of Causal Effects

Jin Tian

Department of Computer Science

226 Atanasoff Hall

Iowa State University

Ames, IA 50011

jtian@cs.iastate.edu

Judea Pearl

Cognitive Systems Laboratory

Computer Science Department

University of California

Los Angeles, CA 90024

judea@cs.ucla.edu

Abstract

This paper concerns the assessment of the effects of actions or policies from a combination of: (i) nonexperimental data, and (ii) causal assumptions. The assumptions are encoded in the form of a directed acyclic graph, also called "causal graph", in which some variables are presumed to be unobserved. The paper establishes new criteria for deciding whether the assumptions encoded in the graph are sufficient for assessing the strength of causal effects and, if the answer is positive, computational procedures are provided for expressing causal effects in terms of the underlying joint distribution.

1 Introduction

This paper explores the feasibility of inferring cause effect relationships from various combinations of data and theoretical assumptions. The assumptions considered will be represented in the form of an acyclic causal diagram containing unmeasured variables [Pearl, 1995, Pearl, 2000] in which arrows represent the potential existence of direct causal relationships between the corresponding variables. Our main task will be to decide whether the assumptions represented in any given diagram are sufficient for assessing the strength of causal effects from nonexperimental data and, if sufficiency is proven, to express the target causal effect in terms of estimable quantities.

It is well known that, in the absence of unmeasured confounders, all causal effects are *identifiable*, that is, the joint response of any set Y of variables to intervention on a set T of action variables, denoted $P_t(y)$,[1] can be estimated consistently from nonexperimental data [Robins, 1986, Spirtes *et al.*, 1993, Pearl, 1993]. If some confounders are not measured, then the question of identifiability arises, and whether the desired quantity can be estimated depends critically on the precise locations (in the diagram) of those confounders vis a vis the sets T and Y. Sufficient graphical conditions for ensuring the identification of $P_t(y)$ were established by several authors [Spirtes *et al.*, 1993, Pearl, 1993, Pearl, 1995] and are summarized in [Pearl, 2000, Chapters 3 and 4]. For example, a criterion called "back-door" permits one to determine whether a given causal effect $P_t(y)$ can be obtained by "adjustment", that is, whether a set C of covariates exists such that

$$P_t(y) = \sum_c P(y|c, t)P(c) \qquad (1)$$

When there exists no set of covariates that is sufficient for adjustment, causal effects can sometimes be estimated by invoking multi-stage adjustments, through a criterion called "front door" [Pearl, 1995]. More generally, identifiability can be decided using *do*-calculus derivations [Pearl, 1995], that is, a sequence of syntactic transformations capable of reducing expressions of the type $P_t(y)$ to subscript-free expressions. Using *do*-calculus as a guide, [Galles and Pearl, 1995] devised a graphical criterion for identifying $P_x(y)$ (where X and Y are singletons) that combines and expands the "front-door" and "back-door" criteria (see [Pearl, 2000, pp. 114-8]).[2] [Pearl and Robins, 1995] further derived a graphical condition under which it is possible to identify $P_t(y)$ where T consists of an arbitrary set of variables. This permits one to predict the effect of time varying treatments from longitudinal data, in the presence of unmeasured confounders, some of which are affected by previous treatments. This criterion was further extended by [Robins, 1997] and [Kuroki and Miyakawa, 1999].

This paper develops new graphical identification criteria that generalize and simplify existing criteria in several ways. In Sections 3-5, we study the

---

[1][Pearl, 1995, Pearl, 2000] used the notation $P(y|set(t))$, $P(y|do(t))$, or $P(y|\hat{t})$ for the post-intervention distribution, while [Lauritzen, 2000] used $P(y||t)$.

[2][Galles and Pearl, 1995] claimed their graphical criterion to embrace all cases where identification is verifiable by *do*-calculus. We show in this paper (Section 4.7) that their criterion is *not* complete in this sense.

identifiability problem in a simpler type of models called semi-Markovian models in which each unobserved variable is a root node with exactly two observed children. Section 3 concerns the identification of $P_x(v)$, where X is a singleton and V is the set of all variables excluding X. It asserts that $P_x(v)$ is identifiable if and only if there is no consecutive sequence of confounding arcs between X and X's immediate successors in the diagram.[3] When interest lies in the effect of X on a subset S of outcome variables, not on the entire set V, it is possible, however, that $P_x(s)$ would be identifiable even though $P_x(v)$ is not. Section 4 first gives a sufficient criterion for identifying $P_x(s)$, which is an extension of the criterion for identifying $P_x(v)$. It says that $P_x(s)$ is identifiable if there is no consecutive sequence of confounding arcs between X and X's children in the subgraph composed of the ancestors of S. Other than this requirement, the diagram may have an arbitrary structure, including any number of confounding arcs between X and S. This simple criterion is shown to cover all criteria reported in the literature (with X singleton), including the "back-door", "front-door", and those developed by [Galles and Pearl, 1995]. However, the criterion is still not necessary for identifying $P_x(s)$. Section 4 further devises a procedure for the identification and computation of $P_x(s)$, based on systematic removal of certain nonessential nodes from G. This procedure is shown to be more powerful than the one devised by [Galles and Pearl, 1995] ([Pearl, 2000, pp. 114-8]). Section 5 deals with the identification of general causal effects, $P_t(s)$, where T and S are arbitrary subsets of variables, representing multiple interventions and multiple outcomes, such as those encountered in the management of time varying treatments. This section established criteria that extend those of [Pearl and Robins, 1995] and [Kuroki and Miyakawa, 1999], and also provides criteria for the identification of *direct effects*, that is, the effect of one variable on another when all other variables are held fixed (Section 5.4). Finally, in Section 6, we show that causal effects in a model with arbitrary sets of unobserved variables can be identified by converting the model into a semi-Markovian model with the same identifiability properties. Section 7 concludes the paper.

---

[3]A variable Z is an "immediate successor" (or a "child") of X if there exists an arrow $X \rightarrow Z$ in the diagram.

2 Notation, Definitions, and Problem Formulation

The use of causal models for encoding distributional and causal assumptions is now fairly standard (see, for example, [Pearl, 1988, Spirtes *et al.*, 1993, Pearl, 1993, Jordan, 1998, Greenland *et al.*, 1999, Lauritzen, 2000, Pearl, 2000]). The simplest such model, called *Markovian*, consists of a directed acyclic graph (DAG) over a set $V = \{V_1, \ldots, V_n\}$ of vertices, representing variables of interest, and a set E of directed edges, or arrows, that connect these vertices. The interpretation of such a graph has two components, probabilistic and causal. The probabilistic interpretation views the arrows as representing probabilistic dependencies among the corresponding variables, and the missing arrows as representing conditional independence assertions: Each variable is independent of all its non-descendants given its direct parents in the graph.[4] These assumptions amount to asserting that the joint probability function $P(v) = P(v_1, \ldots, v_n)$ factorizes according to the product

$$P(v) = \prod_i P(v_i|pa_i) \tag{2}$$

where Pa_i denotes the set of parents of variable V_i in the graph.[5]

The causal interpretation views the arrows as representing causal influences between the corresponding variables. In this interpretation, the factorization of (2) still holds, but the factors are further assumed to represent autonomous data-generation processes, that is, each conditional probability $P(v_i|pa_i)$ represents a stochastic process by which the values of V_i are chosen in response to the values pa_i (previously chosen for V_i's parents), and the stochastic variation of this assignment is assumed independent of the variations in all other assignments. Moreover, each assignment process remains invariant to possible changes in the assignment processes that govern other variables in the system. This modularity assumption enables us to predict the effects of interventions, whenever interventions are described as specific

---

[4]We use family relationships such as "parents," "children," "ancestors," and "descendants," to describe the obvious graphical relationships. For example, we say that V_i is a parent of V_j if there is an arrow from node V_i to V_j, $V_i \rightarrow V_j$.

[5]We use uppercase letters to represent variables or sets of variables, and use corresponding lowercase letters to represent their values (instantiations). For example, pa_i represents an instantiation of Pa_i.

modifications of some factors in the product of (2). The simplest such intervention involves fixing a set T of variables to some constants $T = t$, which yields the post-intervention distribution

$$P_t(v) = \begin{cases} \prod_{\{i|V_i \notin T\}} P(v_i|pa_i) & \text{for all } v \text{ consistent with } T = t. \\ 0 & \text{for all } v \text{ inconsistent with } T = t. \end{cases} \quad (3)$$

Eq. (3) represents a truncated factorization of (2), with factors corresponding to the manipulated variables removed. This truncation follows immediately from (2) since, assuming modularity, the post-intervention probabilities $P(v_i|pa_i)$ corresponding to variables in T are either 1 or 0, while those corresponding to unmanipulated variables remain unaltered.[6] If T stands for a set of treatment variables and Y for an outcome variable in $V \backslash T$, then Eq. (3) permits us to calculate the probability $P_t(y)$ that event $Y = y$ would occur if treatment condition $T = t$ were enforced uniformly over the population. This quantity, often called the "causal effect" of T on Y, is what we normally assess in a controlled experiment with T randomized, in which the distribution of Y is estimated for each level t of T.

We see from Eq. (3) that the model needed for predicting the effect of interventions requires the specification of three elements

$$M = \langle V, G, P(v_i|pa_i) \rangle$$

where (i) $V = \{V_1, \ldots, V_n\}$ is a set of variables, (ii) G is a directed acyclic graph with nodes corresponding to the elements of V, and (iii) $P(v_i|pa_i)$, $i = 1, \ldots, n$, is the conditional probability of variable V_i given its parents in G. Since $P(v_i|pa_i)$ is estimable from nonexperimental data whenever V is observed, we see that, given the causal graph G, all causal effects are estimable from the data as well.[7]

Our ability to estimate $P_t(v)$ from nonexperimental data is severely curtailed when some variables in a Markovian model are unobserved, or, equivalently, if two or more variables in V are affected by unobserved confounders; the presence of such confounders would not permit the decomposition in (2). Letting $V = \{V_1, \ldots, V_n\}$ and $U = \{U_1, \ldots, U_{n'}\}$ stand for the sets

---

[6]Eq. (3) was named "Manipulation Theorem" in [Spirtes *et al.*, 1993], and is also implicit in Robins' (1987) G-computation formula.

[7]It is in fact enough that the parents of each variable in T be observed [Pearl, 2000, p. 78].

of observed and unobserved variables, respectively, the observed probability distribution, $P(v)$, becomes a mixture of products:

$$P(v) = \sum_u \prod_{\{i|V_i \in V\}} P(v_i|pa_{v_i}) \prod_{\{i|U_i \in U\}} P(u_i|pa_{u_i}) \qquad (4)$$

where Pa_{v_i} and Pa_{u_i} stand for the sets of parents of V_i and U_i respectively, and the summation ranges over all the U variables. The post-intervention distribution,[8] likewise, will be given as a mixture of truncated products

$$P_t(v) = \left\{ \begin{array}{ll} \sum_u \prod_{\{i|V_i \notin T\}} P(v_i|pa_{v_i}) \prod_i P(u_i|pa_{u_i}) & v \text{ consistent with } t. \\ 0 & v \text{ inconsistent with } t. \end{array} \right. \qquad (5)$$

And, the question of identifiability arises, i.e., whether it is possible to express $P_t(v)$ as a function of the observed distribution $P(v)$. Clearly, given a causal model M and any two sets T and S in V, $P_t(s)$ can be determined unambiguously using (5). The question of identifiability is whether a given causal effect $P_t(s)$ can be determined uniquely from the distribution $P(v)$ of the observed variables, and is thus independent of the unknown quantities, $P(v_i|pa_{v_i})$ and $P(u_i|pa_{u_i})$, that involve elements of U.

Definition 1 [*Causal-Effect Identifiability*]
The causal effect *of a set of variables T on a disjoint set of variables S is said to be* identifiable *from a graph G if the quantity $P_t(s)$ can be computed uniquely from any positive probability of the observed variables—that is, if $P_t^{M_1}(s) = P_t^{M_2}(s)$ for every pair of models M_1 and M_2 with $P^{M_1}(v) = P^{M_2}(v) > 0$ and $G(M_1) = G(M_2) = G$.*

In other words, given the causal graph G, the quantity $P_t(s)$ can be determined from the observed distribution $P(v)$ alone; the details of M are irrelevant.

If, in a Markovian model with unobserved variables, each unobserved variable is a root node with exactly two observed children, then the corresponding model is called a *semi-Markovian* model. Semi-Markovian models have relatively simple structures, and in Sections 3-5 we will study the identifiability problem in semi-Markovian models only. In Section 6, we show that

---

[8]We only consider interventions on observed variables.

causal effects in a Markovian model with arbitrary sets of unobserved variables can be identified by first converting the model into a semi-Markovian model while keeping the identifiability properties.

In a semi-Markovian model, the observed probability distribution $P(v)$ in Eq.(4) can be written as

$$P(v) = \sum_u \prod_i P(v_i|pa_i, u^i) \prod_i P(u_i) \tag{6}$$

where Pa_i and U^i stand for the sets of the observed and unobserved parents of V_i respectively. The post-intervention distribution is then given by

$$P_t(v) = \begin{cases} \sum_u \prod_{\{i|V_i \notin T\}} P(v_i|pa_i, u^i) \prod_i P(u_i) & v \text{ consistent with } t. \\ 0 & v \text{ inconsistent with } t. \end{cases} \tag{7}$$

It is convenient to represent a semi-Markovian model with a causal graph G that does not show the elements of U explicitly but, instead, represents the confounding effects of U variables using bidirected edges. A bidirected edge between nodes V_i and V_j in G represents divergent edges $V_i \leftarrow U_k \rightarrow V_j$ (see Figure 3 for an example graph). The presence of such bidirected edge in G represents unmeasured factors (or confounders) that may influence two variables in V; we assume that substantive knowledge permits us to decide if such confounders can be ruled out from the model.

3 Identification of $P_x(v)$

Let X be a singleton variable. In this section we study the problem of identifying the causal effects of X on $V \setminus \{X\}$, (namely, on all other variables in V), a quantity denoted by $P_x(v)$.

3.1 The easiest case

Theorem 1 *If there is no bidirected edge connected to X, then $P_x(v)$ is identifiable and is given by*

$$P_x(v) = P(v|x, pa_x)P(pa_x) \tag{8}$$

Proof: Since there is no bidirected edge connected to X, we have that the term $P(x|pa_x, u^x) = P(x|pa_x)$ in Eq. (6) can be moved ahead of the summation, giving

$$
\begin{aligned}
P(v) &= P(x|pa_x) \sum_u \prod_{\{i|V_i \neq X\}} P(v_i|pa_i, u^i) P(u) \\
&= P(x|pa_x) P_x(v).
\end{aligned}
\tag{9}
$$

Hence,

$$
P_x(v) = P(v)/P(x|pa_x) = P(v|x, pa_x) P(pa_x)
\tag{10}
$$

$\square$

Theorem 1 also follows from Theorem 3.2.5 of [Pearl, 2000] which states that for any disjoint sets S and T in a Markovian model M, if the parents of T are measured, then $P_t(s)$ is identifiable. Indeed, when the parents of X are measured, there would be no bidirected edge entering X in the semi-Markovian representation of M and the identification of $P_x(v)$ is insured.

3.2 A more interesting case

The case where there is no bidirected edge connected to any child of X is also easy to handle. As an example, consider the graph given in Figure 1. We have

$$
P(v) = P(z|x) \sum_u P(x|u) P(y|z, u) P(u),
\tag{11}
$$

$$
P_x(v) = P(z|x) \sum_u P(y|z, u) P(u).
\tag{12}
$$

From Eq. (11), we have

$$
\sum_u P(x|u) P(y|z, u) P(u) = P(v)/P(z|x),
\tag{13}
$$

hence,

$$
\sum_u P(y|z, u) P(u) = \sum_x \sum_u P(x|u) P(y|z, u) P(u) = \sum_x P(v)/P(z|x).
\tag{14}
$$

8
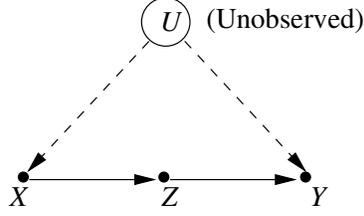
Figure 1:

Substituting Eq. (14) into Eq. (12), we obtain

$$P_x(y, z) = P(z|x) \sum_{x'} P(x', y, z)/P(z|x') = P(z|x) \sum_{x'} P(y|x', z)P(x'). \quad (15)$$

This derivation can be generalized to the case where X has several children. Letting Ch_x denote the set of X's children, we have the following theorem.

Theorem 2 *If there is no bidirected edge connected to any child of X, then $P_x(v)$ is identifiable and is given by*

$$P_x(v) = \left(\prod_{\{i|V_i \in Ch_x\}} P(v_i|pa_i) \right) \sum_x \frac{P(v)}{\prod_{\{i|V_i \in Ch_x\}} P(v_i|pa_i)} \quad (16)$$

Proof: Let $S = V \setminus (Ch_x \cup \{X\})$. Since there is no bidirected edge connected to any child of X, the factors corresponding to the variables in Ch_x can be moved ahead of the summation in Eqs. (6) and (7), and we have

$$P(v) = \left(\prod_{\{i|V_i \in Ch_x\}} P(v_i|pa_i) \right) \sum_u P(x|pa_x, u^x) \prod_{\{i|V_i \in S\}} P(v_i|pa_i, u^i)P(u), \quad (17)$$

and

$$P_x(v) = \left(\prod_{\{i|V_i \in Ch_x\}} P(v_i|pa_i) \right) \sum_u \prod_{\{i|V_i \in S\}} P(v_i|pa_i, u^i)P(u). \quad (18)$$

The variable X does not appear in the factors of $\prod_{\{i|V_i \in S\}} P(v_i|pa_i, u^i)$, hence we augment $\prod_{\{i|V_i \in S\}} P(v_i|pa_i, u^i)$ with the term $\sum_x P(x|pa_x, u^x) = 1$, and

9
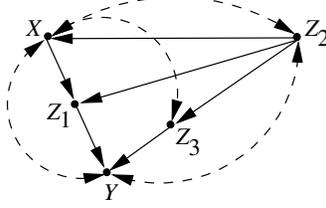
Figure 2:

write

$$\sum_u \prod_{\{i|V_i \in S\}} P(v_i|pa_i, u^i)P(u) = \sum_x \sum_u P(x|pa_x, u^x) \prod_{\{i|V_i \in S\}} P(v_i|pa_i, u^i)P(u)$$

$$= \sum_x \frac{P(v)}{\prod_{\{i|V_i \in Ch_x\}} P(v_i|pa_i)}. \quad \text{(by (17))} \qquad (19)$$

Substituting this expression into Eq. (18) leads to Eq. (16). □

The usefulness of Theorem 2 can be demonstrated in the model of Figure 2. Although the diagram is quite complicated, Theorem 2 is applicable, and readily gives

$$\begin{aligned} P_x(z_1, z_2, z_3, y) &= P(z_1|x, z_2) \sum_{x'} \frac{P(x', z_1, z_2, z_3, y)}{P(z_1|x', z_2)} \\ &= P(z_1|x, z_2) \sum_{x'} P(y, z_3|x', z_1, z_2)P(x', z_2). \end{aligned} \qquad (20)$$

Note that this expression remains valid when we add bidirected edges between Z_3 and Y and between Z_3 and Z_2.

3.3   The general case

When there are bidirected edges connected to the children of X, it may still be possible to identify $P_x(v)$. To illustrate, consider the graph in Figure 3, for which we have

$$P(v) = \sum_{u_1} P(x|u_1)P(z_2|z_1, u_1)P(u_1) \sum_{u_2} P(z_1|x, u_2)P(y|x, z_1, z_2, u_2)P(u_2),$$

$$(21)$$

10
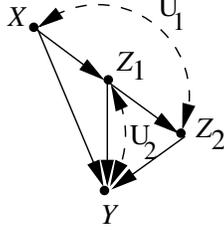
Figure 3:

and

$$P_x(v) = \sum_{u_1} P(z_2|z_1, u_1)P(u_1) \sum_{u_2} P(z_1|x, u_2)P(y|x, z_1, z_2, u_2)P(u_2). \quad (22)$$

Let

$$Q_1 = \sum_{u_1} P(x|u_1)P(z_2|z_1, u_1)P(u_1), \quad (23)$$

$$(24)$$

and

$$Q_2 = \sum_{u_2} P(z_1|x, u_2)P(y|x, z_1, z_2, u_2)P(u_2). \quad (25)$$

Eq. (21) can then be written as

$$P(v) = Q_1 \cdot Q_2, \quad (26)$$

and Eq. (22) as

$$P_x(v) = Q_2 \sum_x Q_1. \quad (27)$$

Thus, if Q_1 and Q_2 can be computed from $P(v)$, then $P_x(v)$ is identifiable and given by Eq. (27). In fact, it is enough to show that Q_1 can be computed from $P(v)$ (i.e., identifiable); Q_2 would then be given by $P(v)/Q_1$. To show that Q_1 can indeed be obtained from $P(v)$, we sum both sides of Eq. (21) over y, and get

$$P(x, z_1, z_2) = Q_1 \cdot \sum_{u_2} P(z_1|x, u_2)P(u_2). \quad (28)$$

11

Summing both sides of (28) over z_2, we get

$$P(x, z_1) = P(x) \sum_{u_2} P(z_1|x, u_2)P(u_2), \qquad (29)$$

hence,

$$\sum_{u_2} P(z_1|x, u_2)P(u_2) = P(z_1|x). \qquad (30)$$

From Eqs. (30) and (28),

$$Q_1 = P(x, z_1, z_2)/P(z_1|x) = P(z_2|x, z_1)P(x), \qquad (31)$$

and from Eq. (26),

$$Q_2 = P(v)/Q_1 = P(y|x, z_1, z_2)P(z_1|x). \qquad (32)$$

Finally, from Eq. (27), we obtain

$$P_x(v) = P(y|x, z_1, z_2)P(z_1|x) \sum_{x'} P(z_2|x', z_1)P(x'). \qquad (33)$$

From the preceding example, we see that because the two bidirected arcs in Figure 3 do not share a common node, the set of factors (of $P(v)$) containing U_1 is disjoint of those containing U_2, and $P(v)$ can be decomposed into a product of two terms, each being a summation of products. This decomposition, to be treated next, plays an important role in the general identifiability problem.

3.3.1 C-component

Let a path composed entirely of bidirected edges be called a *bidirected path*. The set of variables V can be partitioned into disjoint groups by assigning two variables to the same group if and only if they are connected by a bidirected path. Assume that V is thus partitioned into k groups $S_1, \ldots, S_k$, and denote by N_j the set of U variables that are parents of those variables in S_j. Clearly, the sets $N_1, \ldots, N_k$ form a partition of U. Define

$$Q_j = \sum_{n_j} \prod_{\{i|V_i \in S_j\}} P(v_i|pa_i, u^i)P(n_j), \ j = 1, \ldots, k. \qquad (34)$$

12

The disjointness of $N_1, \ldots, N_k$ implies that $P(v)$ can be decomposed into a product of Q_j's:

$$P(v) = \sum_u \prod_i P(v_i | pa_i, u^i) \prod_i P(u_i) = \prod_{j=1}^{k} Q_j. \tag{35}$$

We will call each S_j a *c-component* (abbreviating "confounded component") of V in G or a c-component of G, and Q_j the *c-factor* corresponding to the c-component S_j. The product expressed in (35) will be called the *Q-decomposition* of $P(v)$. For example, in the model of Figure 3, V is partitioned into the c-components $S_1 = \{X, Z_2\}$ and $S_2 = \{Z_1, Y\}$, the corresponding c-factors are given in Eq.s (23) and (25), and $P(v)$ is decomposed into a product of c-factors as in (26).

Let $Pa(S)$ denote the union of a set S and the set of parents of S, that is, $Pa(S) = S \cup (\cup_{V_i \in S} Pa_i)$. We see that Q_j is a function of $Pa(S_j)$. Moreover, each Q_j can be interpreted as the post-intervention distribution of the variables in S_j, under an intervention that sets all other variables to constants, or

$$Q_j = P_{v \setminus s_j}(s_j). \tag{36}$$

The importance of the c-factors stems from that all c-factors are identifiable, as shown in the following lemma.

Lemma 1 *Let a topological order over V be $V_1 < \ldots < V_n$, and let $V^{(i)} = \{V_1, \ldots, V_i\}$, $i = 1, \ldots, n$, and $V^{(0)} = \emptyset$. For any set C, let G_C denote the subgraph of G composed only of variables in C. Then*
(i) Each c-factor Q_j, $j = 1, \ldots, k$, is identifiable and is given by

$$Q_j = \prod_{\{i | V_i \in S_j\}} P(v_i | v^{(i-1)}). \tag{37}$$

(ii) Each factor $P(v_i | v^{(i-1)})$ can be expressed as

$$P(v_i | v^{(i-1)}) = P(v_i | pa(T_i) \setminus \{v_i\}), \tag{38}$$

where T_i is the c-component of $G_{V^{(i)}}$ that contains V_i.

13

Proof: We prove (i) and (ii) simultaneously by induction on the number of variables n.

Base: $n = 1$; we have one c-component $Q_1 = P(v_1)$, which is identifiable and is given by Eq. (37), and Eq. (38) is satisfied.

Hypothesis: When there are n variables, all c-factors are identifiable and are given by Eq. (37), and Eq. (38) holds for all $V_i \in V$.

Induction step: When there are $n + 1$ variables in V, assuming that V is partitioned into c-components $S_1, \ldots, S_l, S'$, with corresponding c-factors $Q_1, \ldots, Q_l, Q'$, and that $V_{n+1} \in S'$, we have

$$P(v) = Q' \prod_i Q_i. \tag{39}$$

Summing both sides of (39) over v_{n+1} leads to

$$P(v^{(n)}) = (\sum_{v_{n+1}} Q') \prod_i Q_i. \tag{40}$$

It is clear that each $S_i, i = 1, \ldots, l$, is a c-component of $G_{V^{(n)}}$. By the induction hypothesis, each $Q_i, i = 1, \ldots, l$, is identifiable and is given by Eq. (37). From Eq. (39), Q' is identifiable as well, and is given by

$$Q' = \frac{P(v)}{\prod_i Q_i} = \prod_{\{i|V_i \in S'\}} P(v_i|v^{(i-1)}), \tag{41}$$

which is clear from Eq. (37) and the chain decomposition $P(v) = \prod_i P(v_i|v^{(i-1)})$.

By the induction hypothesis, Eq. (38) holds for i from 1 to n. Next we prove that it holds for V_{n+1}, which is in the c-component S' of G. In Eq. (41), Q' is a function of $Pa(S')$, and each term $P(v_i|v^{(i-1)})$, $V_i \in S'$ and $V_i \neq V_{n+1}$, is a function of $Pa(T_i)$ by Eq. (38), where T_i is a c-component of the graph $G_{V^{(i)}}$ that contains V_i and therefore is a subset of S'. Hence we obtain that $P(v_{n+1}|v^{(n)})$ is a function only of $Pa(S')$ and is independent of $C = V \setminus Pa(S')$, which leads to

$$\begin{aligned} P(v_{n+1}|pa(S') \setminus \{v_{n+1}\}) &= \sum_c P(v_{n+1}|v^{(n)})P(c|pa(S') \setminus \{v_{n+1}\}) \\ &= P(v_{n+1}|v^{(n)}) \sum_c P(c|pa(S') \setminus \{v_{n+1}\}) \\ &= P(v_{n+1}|v^{(n)}) \end{aligned} \tag{42}$$
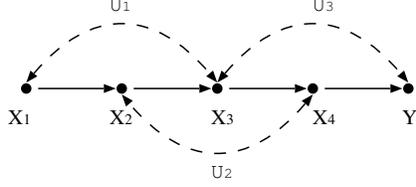
Figure 4:

□

The proposition (ii) in Lemma 1 can also be proved by using d-separation criterion [Pearl, 1988] to show that V_i is independent of $V^{(i)} \setminus Pa(T_i)$ given $Pa(T_i) \setminus \{V_i\}$.

We show the use of Lemma 1 by an example shown in Figure 4, which has two c-components $S_1 = \{X_2, X_4\}$ and $S_2 = \{X_1, X_3, Y\}$. $P(v)$ decomposes into

$$P(x_1, x_2, x_3, x_4, y) = Q_1 Q_2, \tag{43}$$

where

$$Q_1 = \sum_{u_2} P(x_2|x_1, u_2)P(x_4|x_3, u_2)P(u_2), \tag{44}$$

$$Q_2 = \sum_{u_1, u_3} P(x_1|u_1)P(x_3|x_2, u_1, u_3)P(y|x_4, u_3)P(u_1)P(u_3). \tag{45}$$

By Lemma 1, both Q_1 and Q_2 are identifiable. The only admissible order of variables is $X_1 < X_2 < X_3 < X_4 < Y$, and Eq. (37) gives

$$Q_1 = P(x_4|x_1, x_2, x_3)P(x_2|x_1), \tag{46}$$

$$Q_2 = P(y|x_1, x_2, x_3, x_4)P(x_3|x_1, x_2)P(x_1). \tag{47}$$

We can also check that the expressions obtained in Eq.s (31) and (32) for Figure 3 satisfy Lemma 1.

3.3.2   The identification criterion for $P_x(v)$

The Q-decomposition of $P(v)$ (Eq. (35)) combined with Lemma 1 has important implications on the general identifiability problem, and in this section we show how to use this property to identify $P_x(v)$.

15

Let X belong to the c-component S^X with corresponding c-factor Q^X. Let Q^X_x denote the c-factor Q^X with the term $P(x|pa_x, u^x)$ removed, that is,

$$Q^X_x = \sum_{n^X} \prod_{\{i|V_i \neq X, V_i \in S^X\}} P(v_i|pa_i, u^i) P(n^X), \qquad (48)$$

where n^X is the set of U variables that are parents of S^X. We have

$$P(v) = Q^X \prod_i Q_i, \qquad (49)$$

and

$$P_x(v) = Q^X_x \prod_i Q_i. \qquad (50)$$

Since all Q_i's are identifiable by Lemma 1, $P_x(v)$ is identifiable if and only if Q^X_x is identifiable, and we have the following theorem.

Theorem 3 $P_x(v)$ *is identifiable if and only if there is no bidirected path connecting X to any of its children. When $P_x(v)$ is identifiable, it is given by*

$$P_x(v) = \frac{P(v)}{Q^X} \sum_x Q^X, \qquad (51)$$

where Q^X is the c-factor corresponding to the c-component S^X that contains X.

Proof: (**if**) If there is no bidirected path connecting X to any of its children, then none of X's children is in S^X. Under this condition, removing the term $P(x|pa_x, u^x)$ from Q^X is equivalent to summing Q^X over X, and we can write

$$Q^X_x = \sum_x Q^X. \qquad (52)$$

Hence from Eq.s (50) and (49), we obtain

$$P_x(v) = \left(\sum_x Q^X\right) \prod_i Q_i = \left(\sum_x Q^X\right) \frac{P(v)}{Q^X}, \qquad (53)$$

which proves the identifiability of $P_x(v)$.

(**only if**) Sketch: Assuming that there is a bidirected path connecting X to a child of X, one can construct two models (by specifying all conditional

16

probabilities) such that $P(v)$ has the same values in both models while $P_x(v)$ takes different values. The proof is lengthy and is given in Appendix A. □

We demonstrate the use of Theorem 3 by identifying $P_{x_1}(x_2, x_3, x_4, y)$ in Figure 4. The graph has two c-components $S_1 = \{X_2, X_4\}$ and $S_2 = \{X_1, X_3, Y\}$, with corresponding c-factors given in (46) and (47). Since X_1 is in S_2 and its child X_2 is not in S_2, Theorem 3 ensures that $P_{x_1}(x_2, x_3, x_4, y)$ is identifiable and is given by

$$P_{x_1}(x_2, x_3, x_4, y) = Q_1 \sum_{x_1} Q_2$$
$$= P(x_4|x_1, x_2, x_3)P(x_2|x_1) \sum_{x_1'} P(y|x_1', x_2, x_3, x_4)P(x_3|x_1', x_2)P(x_1'). \quad (54)$$

More examples where Theorem 3 is applicable can be found in Figure 3.8 of [Pearl, 2000], some of which required complicated do-calculus derivations.

4 Identification of $P_x(s)$

Let X be a singleton variable and $S \subseteq V$ be a set of variables. In this section, we study the problem of identifying $P_x(s)$. Clearly, whenever $P_x(v)$ is identifiable, so is $P_x(s)$. However, there are obvious cases where $P_x(v)$ is not identifiable and still $P_x(s)$ is identifiable for some subsets S of V. The simplest such example can be seen in Figure 5. Here, variable Z can be ignored in the computation of $P_x(y)$, giving $P_x(y) = P(y|x)$ and $P_x(z) = P(z)$, while (by Theorem 3) $P_x(y, z)$ is not identifiable. This example suggests that a criterion similar to that of Theorem 3, applicable in some subgraphs of G, would establish the identifiability of $P_x(s)$. We will show indeed that $P_x(s)$ is identified when a systematic removal of certain nonessential nodes from G will lead to an identification criterion based on Theorem 3. First we give a criterion for identifying $P_x(s)$ which is a simple extension of Theorem 3.

4.1 A criterion for identifying $P_x(s)$

For any set $C \subseteq V$, let $An(C)$ denote the union of C and the set of ancestors of the variables in C. The nonancestors of S are nonessential for identifying $P_x(s)$ and we have the following lemma.
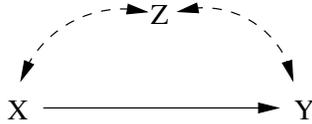
Figure 5:

Lemma 2 $P_x(s)$ *is identifiable if and only if* $P_x(s)$ *is identifiable in the subgraph* $G_{An(S)}$.

Proof: See Appendix B.   □

From Lemma 2, a direct extension of Theorem 3 leads to the following criterion.

Theorem 4 $P_x(s)$ *is identifiable if there is no bidirected path connecting* X *to any of its children in* $G_{An(S)}$.

When the condition in Theorem 4 is satisfied, we can compute $P_x(an(S))$ by applying Theorem 4 in $G_{An(S)}$, and $P_x(s)$ can be obtained by marginalizing over $P_x(an(S))$.

   This simple criterion can classify correctly all the examples treated in the literature with X singleton, including those contrived by [Galles and Pearl, 1995]. In fact, for X and S being singletons, it is shown in [Tian and Pearl, 2002a] that if there is a bidirected path connecting X to one of its children such that every node on the path is in $An(S)$, then none of the "back-door", "front-door", and [Galles and Pearl, 1995] criteria is applicable. However, this criterion is not necessary for identifying $P_x(s)$. In the next section, we give an example in which $P_x(s)$ is identifiable but Theorem 4 is not applicable, and the process of computing $P_x(s)$ will give us hints on how to improve the criterion.

4.2   An example

To illustrate the general process of computing $P_x(s)$ making use of the factorization of $P(v)$ into c-factors, we work out an example in this section. First we introduce a new notation. For any set $C \subseteq V$, define the quantity $Q[C](v)$ to denote the following function

$$Q[C](v) = P_{v \setminus c}(c) = \sum_u \prod_{\{i | V_i \in C\}} P(v_i | pa_i, u^i) P(u). \qquad (55)$$

18
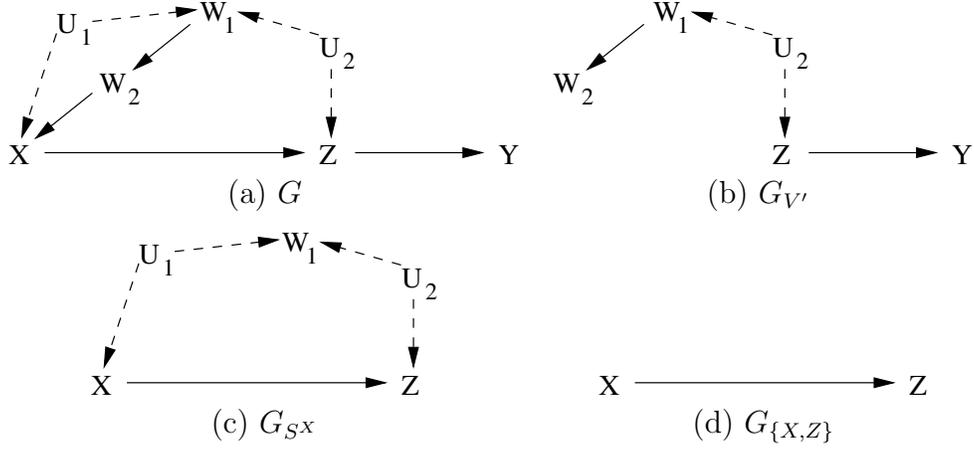
Figure 6:

In particular, we have $Q[V](v) = P(v)$. And we set $Q[\emptyset](v) = 1$ since $\sum_u P(u) = 1$. For convenience, we will often write $Q[C](v)$ as $Q[C]$. $Q[C]$ is a generalization of the Q_i's defined in Section 3.3.1, where the set C was restricted to be a c-component of G. Using this notation the Q-decomposition Eq. (35) becomes

$$Q[V] = \prod_i Q[S_i], \tag{56}$$

where S_i's are the c-components of G. Lemma 1 says that all $Q[S_i]$'s are computable from $Q[V]$.

Consider the problem of identifying $P_x(y)$ in Figure 6(a). Theorem 4 is not applicable, but we will show that $P_x(y)$ is identifiable. Let $V = \{X, Z, Y, W_1, W_2\}$ and $V' = \{Z, Y, W_1, W_2\}$. V is partitioned into three c-components: $S^X = \{X, Z, W_1\}$, $\{W_2\}$, and $\{Y\}$. $P(v)$ can be decomposed into

$$P(v) = P(w_2|w_1)P(y|z)Q[S^X], \tag{57}$$

where

$$Q[S^X] = \sum_{u_1, u_2} P(x|w_2, u_1)P(w_1|u_1, u_2)P(z|x, u_2)P(u_1)P(u_2) \tag{58}$$

$$= P(v)/(P(w_2|w_1)P(y|z)) = P(z, x|w_2, w_1)P(w_1). \tag{59}$$

19

$P_x(v')$ is decomposed into

$$P_x(v') = Q[V'] = P(w_2|w_1)P(y|z)\sum_{u_1,u_2} P(w_1|u_1,u_2)P(z|x,u_2)P(u_1)P(u_2).$$
(60)

We want to compute $P_x(y)$:

$$P_x(y) = \sum_{z,w_1,w_2} P_x(v')$$

$$= \sum_{z,w_1,w_2} Q[V']$$

$$= \sum_{z,w_1} P(y|z)\sum_{u_1,u_2} P(w_1|u_1,u_2)P(z|x,u_2)P(u_1)P(u_2) \quad (\sum_{w_2} P(w_2|w_1) = 1)$$

$$= \sum_z P(y|z)\sum_{u_1,u_2} P(z|x,u_2)P(u_1)P(u_2) \quad (\sum_{w_1} P(w_1|u_1,u_2) = 1)$$

$$= \sum_z P(y|z)Q[\{Z\}].$$
(61)

Note that the key reason for the factors of W_1 and W_2 to be summed out is that $Q[V']$ factorizes according to the subgraph $G_{V'}$ and that W_1 and W_2 are not ancestors of Y in $G_{V'}$ (see Figure 6(b)). The problem of computing $P_x(y)$ is then reduced to computing $Q[\{Z\}]$, which may be computed from $Q[S^X]$. Again, noticing that W_1 is not an ancestor of Z in G_{S^X} (see Figure 6(c)), we sum W_1 over Eq. (58):

$$\sum_{w_1} Q[S^X] = Q[\{X,Z\}]$$
(62)

$$= \sum_{u_1,u_2} P(x|w_2,u_1)P(z|x,u_2)P(u_1)P(u_2)$$
(63)

$$= (\sum_{u_1} P(x|w_2,u_1)P(u_1))(\sum_{u_2} P(z|x,u_2)P(u_2))$$
(64)

$$= Q[\{X\}]Q[\{Z\}]$$
(65)

To compute $Q[\{X\}]$ and $Q[\{Z\}]$, summing Z over Eq. (64), we obtain

$$Q[\{X\}] = \sum_{z,w_1} Q[S^X] = \sum_{w_1} P(x|w_2,w_1)P(w_1),$$
(66)

and from Eq. (65)

$$Q[\{Z\}] = \frac{\sum_{w_1} Q[S^X]}{Q[\{X\}]} = \frac{\sum_{w_1} P(z, x|w_2, w_1)P(w_1)}{\sum_{w_1} P(x|w_2, w_1)P(w_1)}. \tag{67}$$

Finally, substituting the expression for $Q[\{Z\}]$ (67) into Eq. (61), we obtain

$$P_x(y) = \sum_z P(y|z) \frac{\sum_{w_1} P(z, x|w_2, w_1)P(w_1)}{\sum_{w_1} P(x|w_2, w_1)P(w_1)}. \tag{68}$$

From this example, we see that the quantity $Q[C]$ we defined in Eq. (55) plays an important role in identifying $P_x(y)$. The ingredients that allowed us to compute $P_x(y)$ were (i) our ability to sum out some factors from $Q[V']$ as in Eqs. (61), due to the fact that W_1 and W_2 are not ancestors of Y in $G_{V'}$; (ii) our ability to compute $Q[\{X\}]$ and $Q[\{Z\}]$ from $Q[\{X, Z\}]$, which is due to the decomposition of $Q[\{X, Z\}]$ into the product of $Q[\{X\}]$ and $Q[\{Z\}]$ (Eq. (65)) because in the graph $G_{\{X,Z\}}$ (Figure 6(d)), $\{X, Z\}$ is partitioned into two c-components $\{X\}$ and $\{Z\}$. Next, we generalize these ideas and present two lemmas about $Q[C]$ which will facilitate the computing of $P_x(s)$ in general.

4.3 Lemmas

The next lemma provides a condition under which summing $Q[C]$ over some variables is equivalent to removing the corresponding factors. It also provides a condition under which we can compute $Q[W]$ from $Q[C]$, where W is a subset of C, by simply summing $Q[C]$ over the remaining variables (in $C \backslash W$), like ordinary marginalization in probability theory. A set $A \subseteq V$ is called an *ancestral set* if it contains its own ancestors ($A = An(A)$).

Lemma 3 *Let $W \subseteq C \subseteq V$, and $W' = C \setminus W$. If W is an ancestral set in the subgraph G_C ($An(W)_{G_C} = W$), or equivalently, if none of the parents of W is in W' ($Pa(W) \cap W' = \emptyset$), then*

$$\sum_{W'} Q[C] = Q[W]. \tag{69}$$

Proof: First we show that the two conditions are equivalent. If W is an ancestral set in G_C, then obviously none of the parents of W is in W'. On

21

the other hand, if the parents of W are not in W', then $Pa(W)_{G_C} = W$, and therefore $An(W)_{G_C} = W$.

From the definition of $Q[C]$ (see Eq. (55)), $Q[C]$ decomposes according to the subgraph G_C. Summing both sides of (55) over W', the set of nonancestors of W in G_C, then leads to Eq. (69). □

Next, we generalize Eq. (35) (i.e., Eq. (56)) and Lemma 1 to proper subgraphs of G and obtain the following lemma.

Lemma 4 (Generalized Q-decomposition) *Let $H \subseteq V$, and assume that H is partitioned into c-components $H_1, \ldots, H_l$ in the subgraph G_H. Then we have*

(i) $Q[H]$ decomposes as

$$Q[H] = \prod_i Q[H_i]. \tag{70}$$

(ii) Let k be the number of variables in H, and let a topological order of the variables in H be $V_{h_1} < \cdots < V_{h_k}$ in G_H. Let $H^{(i)} = \{V_{h_1}, \ldots, V_{h_i}\}$ be the set of variables in H ordered before V_{h_i} (including V_{h_i}), $i = 1, \ldots, k$, and $H^{(0)} = \emptyset$. Then each $Q[H_j]$, $j = 1, \ldots, l$, is computable from $Q[H]$ and is given by

$$Q[H_j] = \prod_{\{i | V_{h_i} \in H_j\}} \frac{Q[H^{(i)}]}{Q[H^{(i-1)}]}, \tag{71}$$

where each $Q[H^{(i)}]$, $i = 0, 1, \ldots, k$, is given by

$$Q[H^{(i)}] = \sum_{h \backslash h^{(i)}} Q[H]. \tag{72}$$

(iii) Each $Q[H^{(i)}]/Q[H^{(i-1)}]$ is a function only of $Pa(T_i)$, where T_i is the c-component of the subgraph $G_{H^{(i)}}$ that contains V_{h_i}.

Proof: (i) The decomposition of $Q[H]$ into Eq. (70) follows directly from the definition of c-component.

(ii)&(iii) Eq. (72) follows from Lemma 3 since each $H^{(i)}$ is an ancestral set. We prove (ii) and (iii) simultaneously by induction on k.

Base: $k = 1$. There is one c-component $Q[H_1] = Q[H] = Q[H^{(1)}]$ which satisfies Eq. (71) because $Q[\emptyset] = 1$, and $Q[H_1]$ is a function of $Pa(H_1)$.

Hypothesis: When there are k variables in H, all $Q[H_i]$'s are computable from $Q[H]$ and are given by Eq. (71), and (iii) holds for i from 1 to k.

Induction step: When there are $k + 1$ variables in H, assuming that the c-components of G_H are $H_1, \ldots, H_m, H'$, and that $V_{h_{k+1}} \in H'$, we have

$$Q[H] = Q[H^{(k+1)}] = Q[H'] \prod_i Q[H_i]. \tag{73}$$

Summing both sides of (73) over $V_{h_{k+1}}$ leads to

$$\sum_{v_{h_{k+1}}} Q[H] = Q[H^{(k)}] = \Big(\sum_{v_{h_{k+1}}} Q[H'] \Big) \prod_i Q[H_i], \tag{74}$$

where we have used Lemma 3. It is clear that each H_i, $i = 1, \ldots, m$, is a c-component of the subgraph $G_{H^{(k)}}$. Then by the induction hypothesis, each $Q[H_i], i = 1, \ldots, m$, is computable from $Q[H^{(k)}] = \sum_{v_{h_{k+1}}} Q[H]$ and is given by Eq. (71), where each $Q[H^{(i)}]$, $i = 0, 1, \ldots, k$, is given by

$$Q[H^{(i)}] = \sum_{h^{(k)} \backslash h^{(i)}} Q[H^{(k)}] = \sum_{h \backslash h^{(i)}} Q[H]. \tag{75}$$

From Eq. (73), $Q[H']$ is computable as well, and is given by

$$Q[H'] = \frac{Q[H^{(k+1)}]}{\prod_i Q[H_i]} = \prod_{\{i|V_{h_i} \in H'\}} \frac{Q[H^{(i)}]}{Q[H^{(i-1)}]}, \tag{76}$$

which is clear from Eq. (71) and the chain decomposition $Q[H^{(k+1)}] = \prod_{i=1}^{k+1} \frac{Q[H^{(i)}]}{Q[H^{(i-1)}]}$.

By the induction hypothesis, (iii) holds for i from 1 to k. Next we prove that it holds for $Q[H^{(k+1)}]/Q[H^{(k)}]$. The c-component of G that contains $V_{h_{k+1}}$ is H'. In Eq. (76), $Q[H']$ is a function of $Pa(H')$, and each term $Q[H^{(i)}]/Q[H^{(i-1)}]$, $V_{h_i} \in H'$ and $V_{h_i} \neq V_{h_{k+1}}$, is a function of $Pa(T_i)$, where T_i is a c-component of the graph $G_{H^{(i)}}$ that contains V_{h_i} and therefore is a subset of H'. Hence we obtain that $Q[H^{(k+1)}]/Q[H^{(k)}]$ is a function only of $Pa(H')$. □

The use of Lemma 4 can be shown with the example studied in Section 4.2, where the subgraph $G_{\{X,Z\}}$ (Figure 6(d)) is partitioned into two c-components $\{X\}$ and $\{Z\}$, and therefore $Q[\{X\}]$ and $Q[\{Z\}]$ are both

computable from $Q[\{X, Z\}]$. We can check that Eqs. (66) and (67) satisfy (71).

The proposition (iii) in Lemma 4 may imply a set of functional constraints to the distribution $P(v)$ whenever $Q[H]$ is computable from $P(v)$. For example, $Q[\{Z\}]$ is a function of $Pa(\{Z\}) = \{X, Z\}$, therefore the right hand side of Eq. (67) is independent of the value of w_2, which is a constraint to $P(v)$. A procedure that systematically finds functional constraints imposed on the observed distributions in causal models with unobserved variables is given in [Tian and Pearl, 2002b]. Next, we present a procedure for computing $P_x(s)$ based on Lemmas 1, 3, and 4.

4.4   Computing $P_x(s)$

Let V be partitioned into c-components $S^X, S_1, \ldots, S_k$, where $X \in S^X$, and let $V' = V \setminus \{X\}$. We have

$$P(v) = Q[V] = Q[S^X] \prod_i Q[S_i], \tag{77}$$

and

$$P_x(v') = Q[V'] = Q[S^X \setminus \{X\}] \prod_i Q[S_i]. \tag{78}$$

We want to compute

$$P_x(s) = \sum_{V' \setminus S} P_x(v') = \sum_{V' \setminus S} Q[V']. \tag{79}$$

Let $D = An(S)_{G_{V'}}$. By Lemma 3, Eq. (79) becomes

$$P_x(s) = \sum_{D \setminus S} \sum_{V' \setminus D} Q[V'] = \sum_{D \setminus S} Q[D]. \tag{80}$$

Let $D^X = D \cap S^X$, and $D_i = D \cap S_i, i = 1, \ldots, k$. From Eq. (78), $Q[D]$ can be written as

$$Q[D] = Q[D^X] \prod_i Q[D_i] \tag{81}$$

D_i is an ancestral set in G_{S_i} from its definition, hence by Lemma 3,

$$Q[D_i] = \sum_{S_i \setminus D_i} Q[S_i], \ i = 1, \ldots, k. \tag{82}$$

24

However, D^X may not be an ancestral set in G_{S^X} (although it is an ancestral set in $G_{S^X \setminus \{X\}}$), because X could be an ancestor of D^X. Combining Eqs. 80–82, we obtain

$$P_x(s) = \sum_{D \setminus S} Q[D^X] \prod_i \sum_{S_i \setminus D_i} Q[S_i]. \tag{83}$$

Assume that in the graph G_{D^X}, D^X is partitioned into c-components $D_1^X, \ldots, D_l^X$. Then $Q[D^X] = \prod_j Q[D_j^X]$, and we obtain

$$P_x(s) = \sum_{D \setminus S} \prod_j Q[D_j^X] \prod_i \sum_{S_i \setminus D_i} Q[S_i]. \tag{84}$$

Since all the c-factors $Q[S_i]$'s are identifiable, we obtain that $P_x(s)$ is identifiable if all $Q[D_j^X]$'s are identifiable.

Since $D_j^X \subset S^X$, $Q[D_j^X]$ is identifiable if it is computable from $Q[S^X]$. Next, we study the conditions for $Q[D_j^X]$ to be computable from $Q[S^X]$. Let $F = An(D_j^X)_{G_{S^X}}$.

- If $F = D_j^X$, that is, if D_j^X is an ancestral set in G_{S^X}, then by Lemma 3, $Q[D_j^X]$ can be computed as

$$Q[D_j^X] = \sum_{S^X \setminus D_j^X} Q[S^X]. \tag{85}$$

- If $F = S^X$, we are unable to determine whether $Q[D_j^X]$ is computable from $Q[S^X]$ at this moment.

- Assume that $D_j^X \subset F \subset S^X$. By Lemma 3, we have

$$Q[F] = \sum_{S^X \setminus F} Q[S^X]. \tag{86}$$

  Assume that in the graph G_F, D_j^X is contained in a c-component H (the variables in D_j^X are connected by bidirected paths among themselves hence belong to one same c-component). By Lemma 4, $Q[H]$ can be computed from $Q[F]$ and thus is identifiable. We obtain that the problem of whether $Q[D_j^X]$ is computable from $Q[S^X]$ is reduced to that whether $Q[D_j^X]$ is computable from $Q[H]$.

25

Function Identify(C, T, Q)
INPUT: $C \subseteq T \subseteq V$, $Q = Q[T]$. Assuming G_T is composed of one single c-component.
OUTPUT: Expression for $Q[C]$ in terms of Q or FAIL to determine.

Let $A = An(C)_{G_T}$.

- IF $A = C$, output $Q[C] = \sum_{T \setminus C} Q$.

- IF $A = T$, output FAIL.

- IF $C \subset A \subset T$

    1. Assume that in G_A, C is contained in a c-component T'.
    2. Compute $Q[T']$ from $Q[A] = \sum_{T \setminus A} Q$ by Lemma 4.
    3. Output Identify$(C, T', Q[T'])$.

Figure 7: A function determining if $Q[C]$ is computable from $Q[T]$.

The preceding analysis gives a recursive procedure for determining whether $Q[D_j^X]$ is computable from $Q[S^X]$; at each step, we either find an expression for $Q[D_j^X]$, find the problem indeterminable, or reduce the problem to a simpler one in the sense that $H \subset S^X$. In general, for $C \subseteq T \subseteq V$, a recursive algorithm for determining if $Q[C]$ is computable from $Q[T]$ is presented in Figure 7.

In summary, an algorithm for computing $P_x(s)$ is given in Figure 8. The procedure consists of three basic phases. In phase-1, we compute the expressions for all c-factors and find (graphically) the sets D_j^X from the graph G. In phase-2, we attempt to compute $Q[D_j^X]$'s from $Q[S^X]$ by calling the function Identify$(D_j^X, S^X, Q[S^X])$ given in Figure 7. In phase-3, if all $Q[D_j^X]$'s are computable, we output the expression for $P_x(s)$ given in Eq. (84).

From the preceding analysis, we see that the problem of identifying $P_x(s)$ is reduced to that of computing $Q[C]$ from $Q[T]$ for some sets $C \subset T \subseteq V$, for which we give an algorithm in Figure 7. Now the open problem is: Is $Q[C]$ computable from $Q[T]$ if (i) G_C has only one c-component (C itself), (ii) G_T has only one c-component (T itself), and (iii) in G_T, all variables in $T \setminus C$ are ancestors of C $(An(C)_{G_T} = T)$?

Algorithm 1 (Computing $P_x(s)$)
INPUT: a set $S \subset V$.
OUTPUT: the expression for $P_x(s)$ or fail to determine.
Phase-1:

1. *Find the c-components of G: $S^X, S_1, \ldots, S_k$, where $X \in S^X$.*

2. *Compute the c-factors $Q[S^X], Q[S_1], \ldots, Q[S_k]$ by Lemma 1.*

3. *Let $D = An(S)_{G_{V \setminus \{X\}}}$, $D^X = D \cap S^X$.*

4. *Let the c-components of G_{D^X} be $D_j^X, j = 1, \ldots, l$.*

Phase-2:
For each set D_j^X:
  Compute $Q[D_j^X]$ from $Q[S^X]$ by calling the function Identify($D_j^X, S^X, Q[S^X]$) given in Figure 7. If the function returns FAIL, then stop and output FAIL.

Phase-3:
Output $P_x(s) = \sum_{D \setminus S} \prod_j Q[D_j^X] \prod_i \sum_{S_i \setminus D} Q[S_i].$

Figure 8: An algorithm for computing $P_x(s)$

4.5 Useful graphical criteria

We have given a procedure for determining the identifiability of $P_x(s)$ and finding its expression (when identifiable) in Figure 8. Next, we give some graphical criteria based on Algorithm 1 which can be used for quickly judging the identifiability of $P_x(s)$ by looking at the causal graph G.

The idea lies in systematically removing certain nonessential nodes from G till Theorem 4 is applicable (or no more nodes can be removed). First, Lemma 2 can be used to remove nonancestors of S from G. Next, we show that all variables that are not in the same c-components as X can be removed. To prove this conclusion, we present a utility lemma first. Let $A \subseteq B \subseteq V$. We use $Q[A]_{G_B}$ to denote the function $Q[A] = \sum_u \prod_{\{i|V_i \in A\}} P(v_i|pa'_i, u^i)P(u)$ where $PA'_i = PA_i \cap B$. The difference between $Q[A]_{G_B}$ and $Q[A] = Q[A]_{G_V}$ is that some parents of A in G are removed in G_B.

Lemma 5 *Let $A \subseteq B \subseteq V$. $Q[A]$ is computable from $Q[B]$ if and only if $Q[A]_{G_B}$ is computable from $Q[B]_{G_B}$.*

Proof: See Appendix C. □

Using Lemma 5, we obtain the following lemma which reduces the identifiability problem to some subgraph of G.

Lemma 6 *Assume that X is in the c-component S^X, and let $D^X = An(S)_{G_{V \setminus \{X\}}} \cap S^X$. Then $P_x(s)$ is identifiable if in the graph G_{S^X}, $P_x(D^X)$ is identifiable.*

Proof: From Eq. (83), $P_x(s)$ is identifiable if $Q[D^X]$ is identifiable. By Lemma 5, $Q[D^X]$ is identifiable if $Q[D^X]_{G_{S^X}}$ is identifiable. Let $E^X = (S^X \setminus D^X) \setminus \{X\}$. In G_{S^X}, we have

$$P_x(D^X) = \sum_{E^X} P_x(S^X \setminus \{X\}) = \sum_{E^X} Q[S^X \setminus \{X\}]_{G_{S^X}} = Q[D^X]_{G_{S^X}}, \quad (87)$$

where we used Lemma 3 in the last step. Hence we obtain that $P_x(s)$ is identifiable if $P_x(D^X)$ is identifiable in G_{S^X}. □

Lemma 2 and 6 reduce the original problems of deciding the identifiability of $P_x(s)$ in G to (usually simpler) problems of identifying the causal effect of X on a different set of variables in some subgraphs of G. If the latter problem is not recognized to be identifiable (via Theorem 4), we can of course repeat the process and attempt to reduce it further, using Lemma 2 and 6

alternatively.[9] Such recursive application of Lemma 2 and 6 is illustrated in the next example.

4.6 An Example

Consider the problem of identifying $P_x(y)$ in Figure 9(a). By Lemma 2, $P_x(y)$ is identifiable in Figure 9(a) if it is identifiable in Figure 9(b), then by Lemma 6, if it is identifiable in Figure 9(c). After applying Lemma 2 and 6 again (see Figure 9(d) and (e)), the problem is finally reduced to whether $P_x(y)$ is identifiable in Figure 9(f), which is obviously true, and we conclude that $P_x(y)$ is identifiable in Figure 9(a).

We now demonstrate the use of Algorithm 1 by computing $P_x(y)$ in Figure 9(a).

Phase-1:

1. The whole graph is one c-component.

2. $D^X = D = An(\{Y\})_{G_{V \setminus \{X\}}} = \{Y\}$.

3. We want to compute $P_x(y) = Q[\{Y\}]$.

Phase-2:

1. Compute $Q[\{Y\}]$ by calling the function Identify$(\{Y\}, V, P(v))$ in Figure 7. Let $A_1 = An(\{Y\})_G = \{X, Y, W_1, W_2, W_3, W_4\}$. We have $\{Y\} \subset A_1 \subset V$. The graph G_{A_1} (Figure 9(b)) has two c-components: $T_1 = \{X, Y, W_1, W_2, W_3\}$ and $\{W_4\}$, and we have

$$Q[A_1] = \sum_{w_5} P(v) = P(a_1) = Q[T_1]Q[\{W_4\}]. \tag{88}$$

A topological sort over A_1 is: $W_3 < W_4 < W_1 < W_2 < X < Y$. By Lemma 4, we obtain

$$Q[\{W_4\}] = \frac{Q[\{W_4, W_3\}]}{Q[\{W_3\}]} = \frac{\sum_{w_1, w_2, x, y} P(a_1)}{\sum_{w_4, w_1, w_2, x, y} P(a_1)} = P(w_4 | w_3), \tag{89}$$

and from (88),

$$Q[T_1] = P(a_1)/P(w_4|w_3) = P(x, y, w_1, w_2 | w_3, w_4)P(w_3)$$
$$= P(x, y | w_1, w_2, w_3, w_4)P(w_1, w_2, w_3). \tag{90}$$

---

[9]Note that some causal effects identified by Algorithm 1 may not be identified by repeatly using Lemma 2 and 6 which are meant for quick judgement only.

(a) G

(b) G_{A_1}

(c) G_{T_1}

(d) G_{A_2}

(e) G_{T_2}

(f) G_{A_3}
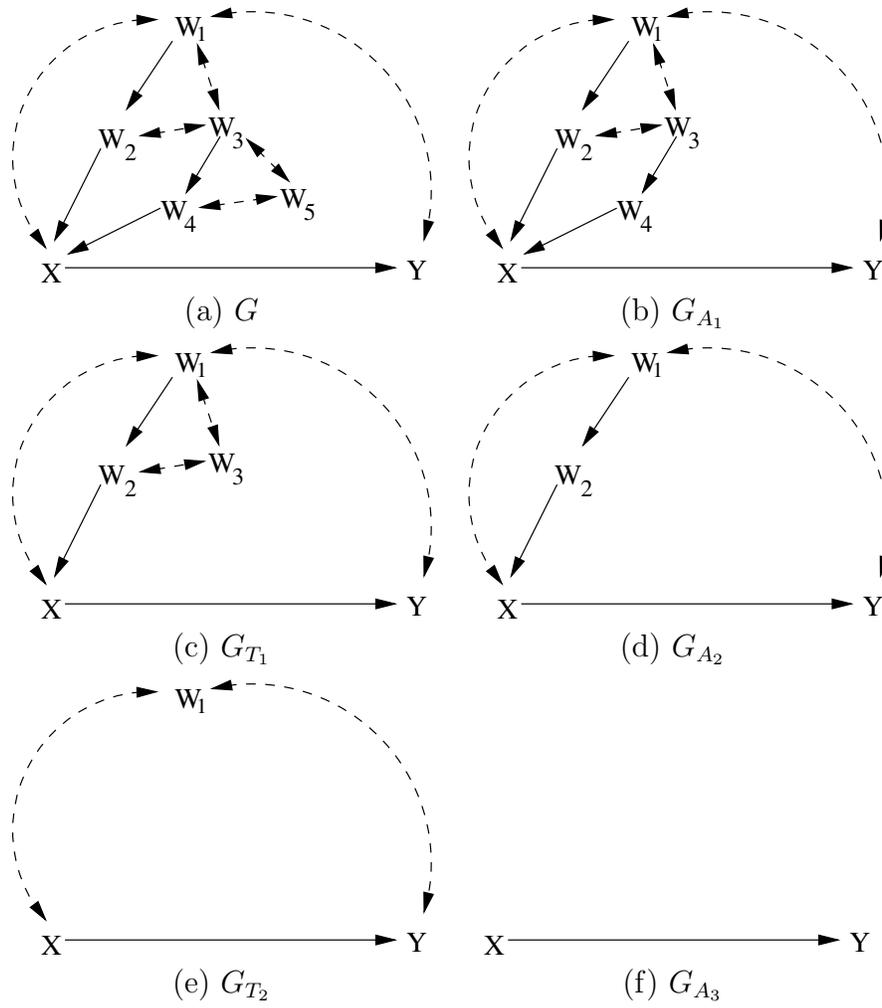
Figure 9:

2. Call the function Identify($\{Y\}, T_1, Q[T_1]$). Let $A_2 = An(\{Y\})_{G_{T_1}} = \{X, Y, W_1, W_2\}$ (see Figure 9(c)). We have $\{Y\} \subset A_2 \subset T_1$. The graph G_{A_2} (Figure 9(d)) has two c-components: $T_2 = \{X, Y, W_1\}$ and $\{W_2\}$, and we have

$$Q[A_2] = \sum_{w_3} Q[T_1] = Q[T_2]Q[\{W_2\}]. \tag{91}$$

A topological sort over A_2 is: $W_1 < W_2 < X < Y$. By Lemma 4, we obtain

$$Q[\{W_2\}] = \frac{Q[\{W_2, W_1\}]}{Q[\{W_1\}]} = \frac{\sum_{x,y} Q[A_2]}{\sum_{w_2,x,y} Q[A_2]} = P(w_2|w_1), \tag{92}$$

and from (91) and (90),

$$Q[T_2] = \sum_{w_3} Q[T_1]/P(w_2|w_1)$$

$$= \sum_{w_3} P(x, y|w_1, w_2, w_3, w_4)P(w_3|w_1, w_2)P(w_1). \tag{93}$$

3. Call the function Identify($\{Y\}, T_2, Q[T_2]$). Let $A_3 = An(\{Y\})_{G_{T_2}} = \{X, Y\}$ (see Figure 9(e)). We have $\{Y\} \subset A_3 \subset T_2$. The graph G_{A_3} (Figure 9(f)) has two c-components: $\{X\}$ and $\{Y\}$, and we have

$$Q[A_3] = \sum_{w_1} Q[T_2] = Q[\{X\}]Q[\{Y\}]. \tag{94}$$

The only admissible order over A_3 is: $X < Y$. By Lemma 4, we obtain

$$Q[\{X\}] = \sum_{y} \sum_{w_1} Q[T_2] = \sum_{w_1,w_3} P(x|w_1, w_2, w_3, w_4)P(w_3|w_1, w_2)P(w_1), \tag{95}$$

and

$$Q[\{Y\}] = (\sum_{w_1} Q[T_2])/Q[\{X\}]$$

$$= \frac{\sum_{w_1,w_3} P(x, y|w_1, w_2, w_3, w_4)P(w_3|w_1, w_2)P(w_1)}{\sum_{w_1,w_3} P(x|w_1, w_2, w_3, w_4)P(w_3|w_1, w_2)P(w_1)}. \tag{96}$$

Phase-3:
Finally, we obtain

$$P_x(y) = Q[\{Y\}] = \frac{\sum_{w_1,w_3} P(x, y|w_1, w_2, w_3, w_4)P(w_3|w_1, w_2)P(w_1)}{\sum_{w_1,w_3} P(x|w_1, w_2, w_3, w_4)P(w_3|w_1, w_2)P(w_1)}. \tag{97}$$

4.7 Galles&Pearl's graphical criterion vs. *do*-calculus

[Galles and Pearl, 1995] claimed that their graphical criterion will embrace all cases where identification is verifiable by *do*-calculus. Here we show that their criterion is not complete in this sense. Consider the problem of identifying $P_x(z)$ in Figure 6(a). Neither "back-door" nor "front-door" criterion is applicable. The graphical criterion in [Galles and Pearl, 1995] also fails because there is no set which can block all back-door paths from X to Z. However we have that $P_x(z) = Q[\{Z\}]$ is identifiable and is given in Eq. (67). $P_x(z)$ can also be computed by *do*-calculus as

$$P(z|\hat{x}) = P(z|\hat{x}, \hat{w_1}) \tag{98}$$
$$= P(z|x, \hat{w_1}) \tag{99}$$
$$= P(z|x, w_2, \hat{w_1}) \tag{100}$$
$$= \frac{P(z, x, w_2|\hat{w_1})}{P(x, w_2|\hat{w_1})} \tag{101}$$
$$= \frac{\sum_{w_1} P(z, x|w_2, w_1)P(w_1)}{\sum_{w_1} P(x|w_2, w_1)P(w_1)} \tag{102}$$

Hence we see that the graphical criterion in [Galles and Pearl, 1995] is *not* complete with respect to *do*-calculus. [Galles and Pearl, 1995] may have failed to consider the possibility of removing a hat by transforming Eq. (100) to (101).

5 Identification of $P_t(s)$

So far, we have assumed that intervention is applied to a single variable X. In this section we study the problem of identifying $P_t(s)$ where S and T are arbitrary (disjoint) subsets of V. We will show that, as for identifying $P_x(s)$, the problem of identifying $P_t(s)$ is also reduced to identifying $Q[C]$ from $Q[C']$ for some sets $C \subset C'$, and we give a procedure for computing $P_t(s)$.

5.1 Computing $P_t(s)$

Let $T' = V \setminus T$, we want to compute

$$P_t(s) = \sum_{T' \setminus S} P_t(t') = \sum_{T' \setminus S} Q[T']. \tag{103}$$

Let $D = An(S)_{G_{T'}}$. Then by Lemma 3,

$$P_t(s) = \sum_{D \setminus S} \sum_{T' \setminus D} Q[T'] = \sum_{D \setminus S} Q[D]. \tag{104}$$

Assume that the subgraph G_D is partitioned into c-components $D_1, \ldots, D_l$. Then $Q[D]$ can be decomposed into products of $Q[D_i]$'s, and Eq. (104) can be rewritten as

$$P_t(s) = \sum_{D \setminus S} \prod_i Q[D_i]. \tag{105}$$

We obtain that $P_t(s)$ is identifiable if all $Q[D_i]$'s are identifiable. Let G be partitioned into c-components $S_1, \ldots, S_k$. Then any D_i is a subset of certain S_j since if the variables in D_i are connected by a bidirected path in a subgraph of G then they must be connected by a bidirected path in G. Assuming $D_i \subseteq S_j$, whether $Q[D_i]$ is identifiable can be determined by using the function Identify$(D_i, S_j, Q[S_j])$ given in Figure 7.

In summary, an algorithm for computing $P_t(s)$ is given in Figure 10. The procedure consists of three basic phases. In phase-1, we compute the expressions for all c-factors and find (graphically) the set of D_i's from the graph G. In phase-2, we attempt to compute $Q[D_i]$'s by calling the function Identify$(D_i, S_j, Q[S_j])$ given in Figure 7. In phase-3, if all $Q[D_i]$'s are identifiable, we output the expression for $P_t(s)$ given in Eq. (105).

5.2 Useful graphical criteria

Next, we give some graphical criteria for quick judgement of the identifiability of $P_t(s)$ by looking at the causal graph G. First we give some graphical conditions for identifying $P_t(v) = P_t(v \setminus t)$, the causal effect of T on all other variables in V. The following criterion is a corollary of Lemma 1.

Algorithm 2 (Computing $P_t(s)$)
INPUT: two disjoint sets $S, T \subset V$.
OUTPUT: the expression for $P_t(s)$ or fail to determine.
Phase-1:

1. *Find the c-components of G: $S_1, \ldots, S_k$.*

2. *Compute the c-factors $Q[S_1], \ldots, Q[S_k]$ by Lemma 1.*

3. *Let $D = An(S)_{G_{V \setminus T}}$.*

4. *Let the c-components of G_D be D_i, $i = 1, \ldots, l$.*

Phase-2:
For each set D_i such that $D_i \subseteq S_j$:
  Compute $Q[D_i]$ from $Q[S_j]$ by calling the function $Identify(D_i, S_j, Q[S_j])$
in Figure 7. If the function returns FAIL, then stop and output FAIL.

Phase-3:
Output $P_t(s) = \sum_{D \setminus S} \prod_i Q[D_i]$.

Figure 10: An algorithm for computing $P_t(s)$

Theorem 5 *If there is no bidirected edge connecting variables in a set T to variables not in T, then $P_t(v)$ is identifiable. Let a topological order over V be $V_1 < \ldots < V_n$, and let $V^{(i)} = \{V_1, \ldots, V_i\}$, $i = 1, \ldots, n$, and $V^{(0)} = \emptyset$. Then $P_t(v)$ is given by*

$$P_t(v \setminus t) = \prod_{\{i|V_i \in V \setminus T\}} P(v_i|pa(C_i) \setminus \{v_i\}), \tag{106}$$

where C_i is the c-component of $G_{V^{(i)}}$ that contains V_i.

In general, let $T' = V \setminus T$, let V be partitioned into c-components $S_1, \ldots, S_k$, and let $T_i = T \cap S_i, T_i' = T' \cap S_i, i = 1, \ldots, k$. We have

$$P_t(t') = \prod_i Q[T_i']. \tag{107}$$

Hence $P_t(t')$ is identifiable if and only if each $Q[T_i']$ is computable from $Q[S_i]$. On the other hand, we have

$$P_{t_j}(v \setminus t_j) = Q[T_j'] \prod_{i \neq j} Q[S_i]. \tag{108}$$

Hence $P_{t_j}(v \setminus t_j)$ is identifiable if and only if $Q[T_j']$ is computable from $Q[S_j]$. And we obtain the following lemma.

Lemma 7 *Let V be partitioned into c-components $S_1, \ldots, S_k$, and let $T_i = T \cap S_i$, $i = 1, \ldots, k$. $P_t(v)$ is identifiable if and only if each $P_{t_i}(v), i = 1, \ldots, k$, is identifiable.*

In the subgraph G_{S_j},

$$P(s_j) = Q[S_j]_{G_{S_j}}, \text{ and } P_{t_j}(s_j \setminus t_j) = Q[T_j']_{G_{S_j}}. \tag{109}$$

Hence by Lemma 5, $Q[C_j]$ is computable from $Q[S_j]$ if and only if $P_{t_j}(s_j \setminus t_j)$ is identifiable in G_{S_j}, which gives the following lemma.

Lemma 8 *Let S_i be a c-component of G, and $T_i \subseteq S_i$. $P_{t_i}(v)$ is identifiable if and only if $P_{t_i}(s_i)$ is identifiable in the graph G_{S_i}.*

One simple condition for $Q[T_i']$ to be computable from $Q[S_i]$ is that T_i' is an ancestral set in G_{S_i}, or T_i contains its own descendants in G_{S_i}. Under this condition, by Lemma 3,

$$Q[T_i'] = \sum_{T_i} Q[S_i]. \tag{110}$$

And we obtain the following theorem.

Theorem 6 *Let S_i be a c-component of G, and $T_i \subseteq S_i$. If the children of variables in T_i are either in T_i or outside of S_i (i.e. T_i contains its own descendants in G_{S_i}), then $P_{t_i}(v)$ is identifiable, and is given by*

$$P_{t_i}(v \setminus t_i) = \frac{P(v)}{Q[S_i]} \sum_{T_i} Q[S_i]. \tag{111}$$

Next, we give some graphical conditions for quick judgment of the identifiability of $P_t(s)$.

Lemma 9 *Let V be partitioned into c-components $S_1, \ldots, S_k$. Let $T_i = T \cap S_i, D_i = An(S)_{G_{V \setminus T}} \cap S_i, i = 1, \ldots, k$. Then $P_t(s)$ is identifiable if every $P_{t_i}(d_i)$ is identifiable in G_{S_i} for $i = 1, \ldots, k$.*

Proof: From Eq. (105), $P_t(s)$ is identifiable if each $Q[D_i]$ is identifiable. By Lemma 5, $Q[D_i]$ is computable from $Q[S_i]$ if $Q[D_i]_{G_{S_i}}$ is computable from $Q[S_i]_{G_{S_i}}$. Let $T_i' = S_i \setminus T_i$. In G_{S_i}, we have

$$P_{t_i}(d_i) = \sum_{T_i' \setminus D_i} P_{t_i}(t_i') = \sum_{T_i' \setminus D_i} Q[T_i']_{G_{S_i}} = Q[D_i]_{G_{S_i}}, \tag{112}$$

where we used Lemma 3 in the last step. Hence we obtain that $P_t(s)$ is identifiable if each $P_{t_i}(d_i)$ is identifiable in G_{S_i}. $\qquad \square$

Lemma 10 *Let $T_1 = T \cap An(S)$. $P_t(s)$ is identifiable if and only if $P_{t_1}(s)$ is identifiable in $G_{An(S)}$.*

Proof: It is well-known that $P_t(s) = P_{t_1}(s)$. The rest of the proof is the same as that for Lemma 2 (see Appendix B). $\qquad \square$

Lemma 9 and 10 reduce the original problems of deciding the identifiability of $P_t(s)$ in G to some (usually simpler) identifiability problems in subgraphs of G. They can be repeatedly applied to further reduce the problems, till inapplicable or till those problems are recognized to be identifiable (for example, via Theorem 4 or 6).
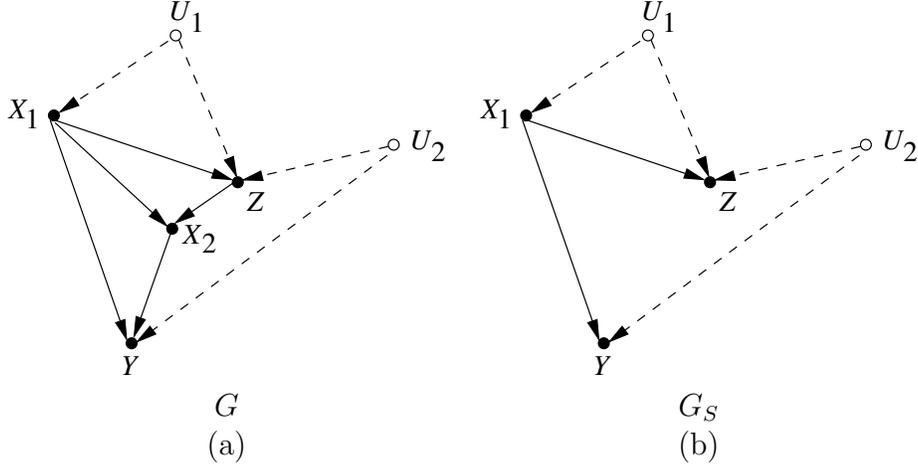
Figure 11:

5.3 Examples

Next, we study some examples, to illustrate the use of Algorithm 2 and the graphical criteria in Section 5.2.

Consider the problem of identifying $P_{x_1x_2}(y)$ in Figure 11(a), which was studied in [Pearl and Robins, 1995]. G has two c-components $S = \{X_1, Z, Y\}$ and $\{X_2\}$, and X_1 and X_2 are in different c-components. Letting $C = V \setminus \{X_1, X_2\} = \{Y, Z\}$, then $An(\{Y\})_{G_C} = \{Y\} \subset S$. By Lemma 9 we have that $P_{x_1x_2}(y)$ is identifiable if $P_{x_1}(y)$ is identifiable in the subgraph G_S (Figure 11(b)). Since the latter is true by Theorem 4, we conclude that $P_{x_1x_2}(y)$ is identifiable. Next we compute $P_{x_1x_2}(y)$. We have

$$P(v) = P(x_2|x_1, z)Q[S], \tag{113}$$

from which we obtain

$$Q[S] = P(v)/P(x_2|x_1, z) = P(y|x_1, x_2, z)P(x_1, z). \tag{114}$$

$P_{x_1x_2}(y)$ is computed as

$$P_{x_1x_2}(y) = \sum_z Q[\{Y, Z\}] = Q[\{Y\}], \tag{115}$$

which can be computed by calling Identify($\{Y\}, S, Q[S]$) in Figure 7. Let $A = An(\{Y\})_{G_S} = \{X_1, Y\}$. We have $\{Y\} \subset A \subset S$. The graph G_A has two

37
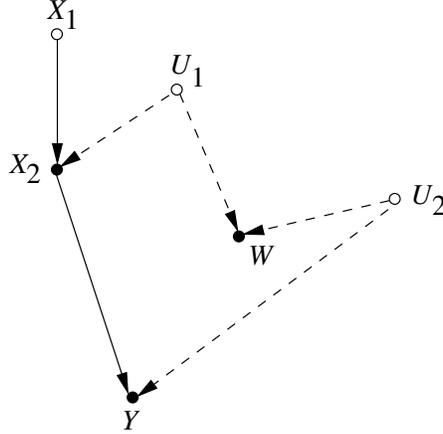
Figure 12:

c-components: $\{X_1\}$ and $\{Y\}$, and we have

$$Q[A] = \sum_z Q[S] = Q[\{X_1\}]Q[\{Y\}]. \tag{116}$$

The only admissible order over A is: $X_1 < Y$. By Lemma 4, we obtain

$$Q[\{X_1\}] = \sum_y \sum_z Q[S] = P(x_1), \tag{117}$$

and

$$Q[\{Y\}] = \sum_z Q[S]/Q[\{X_1\}] = \sum_z P(y|x_1, x_2, z)P(z|x_1). \tag{118}$$

Finally, we obtain

$$P_{x_1x_2}(y) = Q[\{Y\}] = \sum_z P(y|x_1, x_2, z)P(z|x_1), \tag{119}$$

which coincides with Eq. (4.3) of [Pearl, 2000, page 122].

Consider the problem of identifying $P_{x_1x_2}(y)$ in Figure 12, which was studied in [Pearl and Robins, 1995]. G has two c-components $S = \{X_2, W, Y\}$ and $\{X_1\}$, and X_1 and X_2 are in different c-components. Letting $C = V \setminus \{X_1, X_2\} = \{Y, W\}$, then $An(\{Y\})_{G_C} = \{Y\} \subset S$. By Lemma 9,
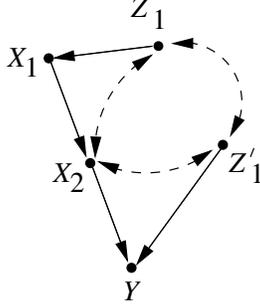
Figure 13:

$P_{x_1 x_2}(y)$ is identifiable if $P_{x_2}(y)$ is identifiable in G_S. It is clear that $P_{x_2}(y)$ is identifiable (by Theorem 4), hence $P_{x_1 x_2}(y)$ is identifiable.

Consider the problem of identifying $P_{x_1 x_2}(y)$ in Figure 13, which was studied in [Pearl and Robins, 1995]. G has three c-components $\{X_1\}, \{Y\}$, and $S = \{X_2, Z_1, Z_1'\}$, and X_1 and X_2 are in different c-components. By Lemma 7, $P_{x_1 x_2}(v)$ is identifiable if both $P_{x_1}(v)$ and $P_{x_2}(v)$ are identifiable, which is true by Theorem 3. Therefore $P_{x_1 x_2}(v)$ is identifiable. Next we compute $P_{x_1 x_2}(v)$. We have

$$P(v) = P(x_1|z_1)P(y|x_2, z_1')Q[S], \tag{120}$$

from which we obtain

$$Q[S] = P(v)/(P(x_1|z_1)P(y|x_2, z_1')) = P(x_2, z_1'|x_1, z_1)P(z_1). \tag{121}$$

$P_{x_1 x_2}(v)$ is computed as

$$\begin{aligned} P_{x_1 x_2}(y, z_1, z_1') &= P(y|x_2, z_1')Q[\{Z_1, Z_1'\}] \\ &= P(y|x_2, z_1') \sum_{x_2} Q[S] \\ &= P(y|x_2, z_1')P(z_1'|x_1, z_1)P(z_1) \\ &= P(y|x_2, z_1')P(z_1', z_1). \end{aligned} \tag{122}$$

Next, consider the problem of identifying $P_{x_1 x_2}(y)$ in Figure 14, which was studied in [Kuroki and Miyakawa, 1999]. X_1 and X_2 are in the same c-component $S = \{X_1, X_2, Y\}$, and their children other than X_2 itself are not in S, hence Theorem 6 is applicable and $P_{x_1 x_2}(v)$ is identifiable. We have

$$P(v) = P(z_1|x_1)P(z_2|x_1, x_2)Q[S], \tag{123}$$

39
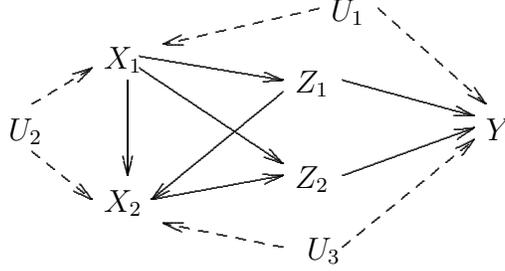
Figure 14: from [Kuroki and Miyakawa, 1999]

from which we obtain

$$Q[S] = P(v)/\big(P(z_1|x_1)P(z_2|x_1, x_2)\big)$$
$$= P(y|x_1, x_2, z_1, z_2)P(x_2|x_1, z_1)P(x_1). \tag{124}$$

From Theorem 6, we have

$$P_{x_1x_2}(y, z_1, z_2) = P(z_1|x_1)P(z_2|x_1, x_2) \sum_{x_1,x_2} Q[S]$$
$$= P(z_1|x_1)P(z_2|x_1, x_2) \sum_{x_1',x_2'} P(y|x_1', x_2', z_1, z_2)P(x_2'|x_1', z_1)P(x_1').$$
$$\tag{125}$$

We further obtain

$$P_{x_1x_2}(y) = \sum_{z_1,z_2} P(z_1|x_1)P(z_2|x_1, x_2) \sum_{x_1',x_2'} P(y|x_1', x_2', z_1, z_2)P(x_2'|x_1', z_1)P(x_1'),$$
$$\tag{126}$$

which coincides with Eq. (3.12) of [Kuroki and Miyakawa, 1999].

Consider the problem of identifying $P_{x_1x_2}(y)$ in Figure 15(a), which was studied in [Kuroki and Miyakawa, 1999]. X_1 and X_2 are in the same c-component $S = \{X_1, X_2, Y\}$. By Lemma 8, $P_{x_1x_2}(v)$ is identifiable if $P_{x_1x_2}(y)$ is identifiable in G_S (Figure 15(b)). Let $A = An(\{Y\})_{G_S} = \{X_1, Y\}$. By Lemma 10, $P_{x_1x_2}(y)$ is identifiable in G_S if $P_{x_1}(y)$ is identifiable in the subgraph G_A (Figure 15(c)). Since $P_{x_1}(y)$ is obviously identifiable in G_A, we conclude that $P_{x_1x_2}(v)$ is identifiable. We have

$$P(v) = P(z_2|x_1, x_2)Q[S], \tag{127}$$

40
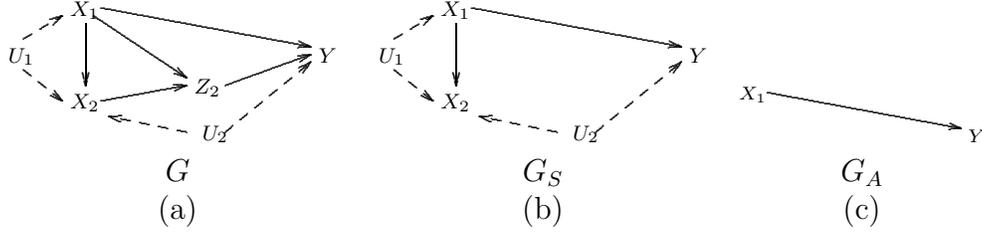
Figure 15: from [Kuroki and Miyakawa, 1999]

from which we obtain

$$Q[S] = P(v)/P(z_2|x_1, x_2) = P(y|z_2, x_1, x_2)P(x_1, x_2). \qquad (128)$$

$P_{x_1 x_2}(v)$ is computed as

$$P_{x_1,x_2}(z_2, y) = P(z_2|x_1, x_2)Q[\{Y\}]. \qquad (129)$$

$Q[\{Y\}]$ can be computed by calling Identify$(\{Y\}, S, Q[S])$ in Figure 7. We have

$$Q[A] = \sum_{x_2} Q[S] = Q[\{X_1\}]Q[\{Y\}], \qquad (130)$$

from which we obtain

$$Q[\{X_1\}] = \sum_{y} Q[A] = P(x_1), \qquad (131)$$

and

$$Q[\{Y\}] = \sum_{x_2} Q[S]/Q[\{X_1\}] = \sum_{x_2} P(y|z_2, x_1, x_2)P(x_2|x_1). \qquad (132)$$

Finally, substituting (132) into (129), we obtain

$$P_{x_1,x_2}(z_2, y) = P(z_2|x_1, x_2) \sum_{x_2'} P(y|z_2, x_1, x_2')P(x_2'|x_1), \qquad (133)$$

and

$$P_{x_1,x_2}(y) = \sum_{z_2} P(z_2|x_1, x_2) \sum_{x_2'} P(y|z_2, x_1, x_2')P(x_2'|x_1), \qquad (134)$$

41

(a) G (b) G_S (c) $G_{S'}$
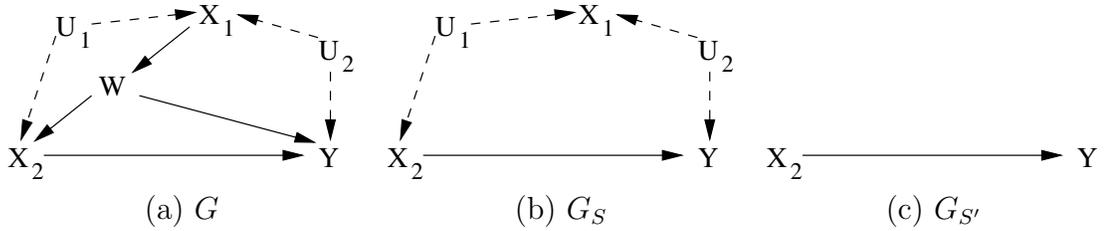
Figure 16:

which coincides with Eq. (3.21) of [Kuroki and Miyakawa, 1999].

In the examples studied so far, in Figure 11(a), 12, and 13, $P_{x_1 x_2}(y)$ can be identified using the criteria given in [Pearl and Robins, 1995]. In Figure 14 and 15(a), $P_{x_1 x_2}(y)$ can be identified by the extended front-door criterion and the mixed-door criterion given in [Kuroki and Miyakawa, 1999] respectively. Next we give an example shown in Figure 16(a), for which $P_{x_1 x_2}(w, y)$ is identifiable, but none of the criteria in [Pearl and Robins, 1995] and [Kuroki and Miyakawa, 1999] is applicable. X_1 and X_2 are in the same c-component $S = \{X_1, X_2, Y\}$. By Lemma 8, $P_{x_1 x_2}(v)$ is identifiable if $P_{x_1 x_2}(y)$ is identifiable in G_S (Figure 16(b)). The latter is obviously true, hence we conclude that $P_{x_1 x_2}(w, y)$ is identifiable. (Formally, let $S' = An(\{Y\})_{G_S} = \{X_2, Y\}$; by Lemma 10, $P_{x_1 x_2}(y)$ is identifiable in G_S if $P_{x_2}(y)$ is identifiable in the subgraph $G_{S'}$ (Figure 15(c)), which is obvious.)

5.4   Identification of Direct Effects $P_{pa_y}(y)$

Let Y be a single variable and let $V_Y = V \setminus \{Y\}$ be the set of all other variables. A special case of the identifiability problem is to identify the *direct effect* $P_{v_y}(y)$. We have

$$P_{v_y}(y) = P_{pa_y}(y) = Q[\{Y\}]. \tag{135}$$

Let Y be in the c-component S^Y. In general, the identifiability of $P_{pa_y}(y)$ can be determined by using the function Identify($\{Y\}, S^Y, Q[S^Y]$) in Figure 7. In this section we give some graphical criteria for determining whether $P_{pa_y}(y)$ is identifiable.

Theorem 7 *If Y is not connected to bidirected links, then $P_{pa_y}(y)$ is identifiable, and is given by*

$$P_{pa_y}(y) = P(y|pa_y). \tag{136}$$

42

Theorem 7 is obvious. The use of Theorem 7 can be shown by identifying the direct effect on Y in Figure 13. Theorem 7 says that $P_{x_2,z_1'}(y)$ is identifiable and is equal to $P(y|x_2, z_1')$.

Theorem 8 *Let Y be in the c-component S^Y. If there is no bidirected path connecting Y and any of its parents (i.e., Y is not in the same c-components with any of its parents), then $P_{pa_y}(y)$ is identifiable, and is given by*

$$P_{pa_y}(y) = \sum_{S^Y \setminus \{Y\}} Q[S^Y]. \tag{137}$$

Proof: Since none of the variables in $S^Y \setminus \{Y\}$ is an ancestor of Y in the subgraph G_{S^Y}, by Lemma 3, $Q[\{Y\}] = \sum_{S^Y \setminus \{Y\}} Q[S^Y]$. $\qquad \square$
We demonstrate the use of Theorem 8 by identifying the direct effect on Y in Figure 14. Y is in the c-component $S = \{X_1, X_2, Y\}$, and $Q[S]$ is given in Eq. (124). By Theorem 8, $P_{z_1,z_2}(y)$ is identifiable and is given by

$$P_{z_1,z_2}(y) = \sum_{x_1,x_2} P(y|x_1, x_2, z_1, z_2)P(x_2|x_1, z_1)P(x_1). \tag{138}$$

Lemma 11 *The direct effect on Y is identifiable if and only if the direct effect on Y is identifiable in $G_{An(\{Y\})}$.*

Lemma 11 follows from Lemma 10.

Lemma 12 *Let Y be in the c-component S^Y. The direct effect on Y is identifiable if and only if the direct effect on Y is identifiable in G_{S^Y}.*

Proof: By Lemma 5, $Q[\{Y\}]$ is computable from $Q[S^Y]$ if and only if $Q[\{Y\}]_{G_{S^Y}}$ is computable from $Q[S^Y]_{G_{S^Y}}$. $\qquad \square$

Lemma 11 and 12 can be applied alternatively to remove nodes from a graph, until it is clear that the direct effect on Y is identifiable or until neither lemmas is applicable. This leads to the following criterion.

Theorem 9 *The direct effect on Y is identifiable if there exists no subgraph G_S of G satisfying all of the following: (i) $Y \in S$; (ii) G_S has only one c-component, S itself; (iii) All variables in S are ancestors of Y in G_S.*

The graph in Figure 17 satisfies conditions (i)-(iii), and for general graphs of such a type, we are unable to determine the identifiability of the direct effect on Y.
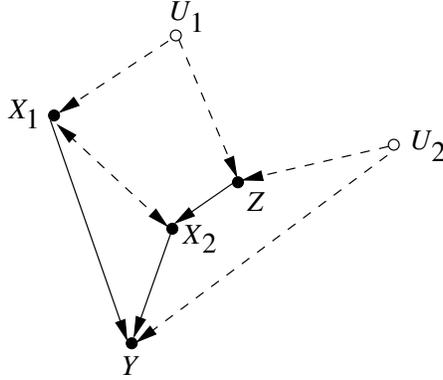
Figure 17:

6 Beyond Semi-Markovian Models

In Sections 3-5 we have studied the identifiability problem in semi-Markovian models. Our method is based on the Q-decomposition of $P(v)$ and Lemmas 1, 3, and 4. In a Markovian model with arbitrary sets of unobserved variables, $P(v)$ in Eq. (4) may also be decomposed into a product of summations as

$$P(v) = \prod_i Q[S_i], \tag{139}$$

where S_i's form a partition of V and, similar to Eq. (55), each $Q[S_i]$ is given by

$$Q[S_i] = \sum_u \prod_{\{i|V_i \in S_i\}} P(v_i|pa_{v_i}) \prod_{\{i|U_i \in U\}} P(u_i|pa_{u_i}). \tag{140}$$

The graphical conditions for this decomposition to be feasible are more complicated than that in Section 3.3.1 and are given in [Tian and Pearl, 2002b], which also showed that properties as given in Lemmas 1, 3, and 4 hold as well. Therefore, we can use the same method developed in Sections 3-5 to identify causal effects in a Markovian model with arbitrary sets of unobserved variables. [Tian and Pearl, 2002b] also suggests that, instead of working directly with a complicated model with arbitrary sets of unobserved variables, we may work with its semi-Markovian *projection* [Verma, 1993].

44

Definition 2 (Projection) *The projection of a DAG G over $V \cup U$ on the set V, denoted by $PJ(G,V)$, is a DAG over V with bidirected edges constructed as follows:*

1. *Add each variable in V as a node of $PJ(G,V)$.*

2. *For each pair of variables $X, Y \in V$, if there is an edge between them in G, add the edge to $PJ(G,V)$.*

3. *For each pair of variables $X, Y \in V$, if there exists a directed path from X to Y in G such that every internal node on the path is in U, add edge $X \to Y$ to $PJ(G,V)$ (if it does not exist yet).*

4. *For each pair of variables $X, Y \in V$, if there exists a divergent path between X and Y in G such that every internal node on the path is in U ($X \dashleftarrow U_i \dashrightarrow Y$), add a bidirected edge $X \dashleftarrow\!-\!-\!-\!\dashrightarrow Y$ to $PJ(G,V)$.*

It is shown in [Tian and Pearl, 2002b] that G and $PJ(G,V)$ have the same topological relations over V and the same partition of V into c-components. Based on the results in [Tian and Pearl, 2002b], we conclude that if $P_t(s)$ is identified in $PJ(G,V)$ (using the methods in Sections 3-5), then it is identified in G with the same expression.

In summary, to identify a causal effect $P_t(s)$ in a model with arbitrary sets of unobserved variables, we first construct the projection graph $PJ(G,V)$, then attempt to compute $P_t(s)$ in $PJ(G,V)$; if $P_t(s)$ is computable in $PJ(G,V)$, then $P_t(s)$ is identifiable in G with the same expression.

7   Conclusion

This paper develops graphical criteria that permit one to decide, by merely inspecting a causal diagram, whether the effect of a given action or policy can be determined from passive observations, namely, from observations that involve no experimental manipulations. The criteria developed simplify, generalize, and unify those reported in the literature, and are based on a general decomposition scheme, called Q-decomposition, whereby a causal graph is decomposed into C-components (subgraphs) among which there exists no path of consecutive spurious dependencies.

We have shown that the effect of a singleton action on all other variables in the system can be predicted if and only if the action variable is not in the

same C-component as any of its direct successors. Extensions were further developed to cases where the action affects several variables at once, and where attention is focused on a subset of the response variables.

These results have wide and immediate applications in the health and social sciences, where investigators are often required to elucidate cause-effect relationships (e.g., the effect of treatments on diseases) from observational studies of populations under natural conditions. They also have applications in artificial intelligence systems where agents, equipped with incomplete models of environment, are required to control their environment with no prior manipulative training. The results developed in this paper enable researchers and agents to decide whether the observations available are sufficient for controlling one's environment, whether additional observations are required, or whether the assumptions underlying the model need be refined.

References

[Galles and Pearl, 1995] D. Galles and J. Pearl. Testing identifiability of causal effects. In P. Besnard and S. Hanks, editors, *Uncertainty in Artificial Intelligence 11*, pages 185–195. Morgan Kaufmann, San Francisco, 1995.

[Greenland *et al.*, 1999] S. Greenland, J. Pearl, and J.M Robins. Causal diagrams for epidemiologic research. *Epidemiology*, 10:37–48, 1999.

[Jordan, 1998] M.I. Jordan. *Learning in Graphical Models*. Series D: Behavioural and Social Sciences – Vol. 89. Kluwer Academic Publishers, Dordrecht, 1998.

[Kuroki and Miyakawa, 1999] M. Kuroki and M. Miyakawa. Identifiability criteria for causal effects of joint interventions. *Journal of the Japan Statistical Society*, 29(2):105–117, 1999.

[Lauritzen, 2000] S. Lauritzen. Graphical models for causal inference. In O.E. Barndorff-Nielsen, D. Cox, and C. Kluppelberg, editors, *Complex Stochastic Systems*, chapter 2, pages 67–112. Chapman and Hall/CRC Press, London/Boca Raton, 2000.

[Pearl and Robins, 1995] J. Pearl and J.M. Robins. Probabilistic evaluation of sequential plans from causal models with hidden variables. In P. Besnard

and S. Hanks, editors, *Uncertainty in Artificial Intelligence 11*, pages 444–453. Morgan Kaufmann, San Francisco, 1995.

[Pearl, 1988] J. Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Mateo, CA, 1988.

[Pearl, 1993] J. Pearl. Comment: Graphical models, causality, and intervention. *Statistical Science*, 8:266–269, 1993.

[Pearl, 1995] J. Pearl. Causal diagrams for empirical research. *Biometrika*, 82:669–710, December 1995.

[Pearl, 2000] J. Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, NY, 2000.

[Robins, 1986] J.M. Robins. A new approach to causal inference in mortality studies with a sustained exposure period – applications to control of the healthy workers survivor effect. *Mathematical Modeling*, 7:1393–1512, 1986.

[Robins, 1987] J.M. Robins. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *Journal of Chronic Diseases*, 40(Suppl 2):139S–161S, 1987.

[Robins, 1997] J.M. Robins. Causal inference from complex longitudinal data. In *Latent Variable Modeling with Applications to Causality*, pages 69–117. Springer-Verlag, New York, 1997.

[Spirtes *et al.*, 1993] P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search*. Springer-Verlag, New York, 1993.

[Tian and Pearl, 2002a] J. Tian and J. Pearl. A general identification condition for causal effects. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI)*, pages 567–573, Menlo Park, CA, 2002. AAAI Press/The MIT Press.

[Tian and Pearl, 2002b] J. Tian and J. Pearl. On the testable implications of causal models with hidden variables. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI)*, 2002.

[Verma, 1993] T. S. Verma. Graphical aspects of causal models. Technical Report R-191, Computer Science Department, University of California, Los Angles, 1993.

A Proof for Theorem 3

In this appendix we prove the necessity part of the criterion given in Theorem 3 for the identifiability of $P_x(v)$. To facilitate the proof, first we prove the following lemma.

Lemma 13 *Let $S, T \subseteq V$ be two disjoint sets of variables. If $P_t(s)$ is not identifiable in G, then $P_t(s)$ is not identifiable in the graph resulted from adding a directed or bidirected edge to G. Equivalently, if $P_t(s)$ is identifiable in G, then $P_t(s)$ is still identifiable in the graph resulted from removing a directed or bidirected edge from G.*

Proof: If $P_t(s)$ is not identifiable in G, then there exist two models with the same causal graph G, M_1 and M_2, such that

$$P^{M_1}(v) = P^{M_2}(v) > 0, \text{ and } P_t^{M_1}(s) \neq P_t^{M_2}(s), \tag{141}$$

where

$$P^{M_k}(v) = \sum_u \prod_i P^{M_k}(v_i | pa_i, u^i) P^{M_k}(u), \quad k = 1, 2. \tag{142}$$

For a graph G' with extra edges added to G, we can always construct new models in such a way that the added edges are ineffective.

(i) Let G' be the graph identical to G except with an extra edge $Y \longrightarrow V_j$. $P(v)$ decomposes as

$$P(v) = \sum_u P(v_j | pa_j, y, u^j) \prod_{i \neq j} P(v_i | pa_i, u^i) P(u). \tag{143}$$

We construct two models M_1' and M_2' with the causal graph G' as

$$P^{M_k'}(v_i | pa_i, u^i) = P^{M_k}(v_i | pa_i, u^i), \ i \neq j, \ k = 1, 2, \tag{144}$$

$$P^{M_k'}(v_j | pa_j, y, u^j) = P^{M_k}(v_j | pa_j, u^j), \ k = 1, 2, \tag{145}$$

$$P^{M_k'}(u) = P^{M_k}(u), \ k = 1, 2. \tag{146}$$

Clearly, if the pair (M_1, M_2) satisfies (141), so would the pair (M_1', M_2'). Hence $P_t(s)$ is not identifiable in G'.
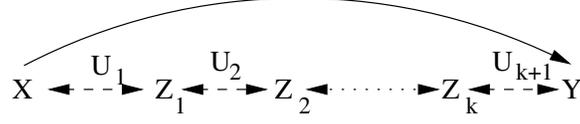
Figure 18:

(ii) Let G' be the graph identical to G except with an extra edge $V_l \longleftrightarrow V_j$. $P(v)$ decomposes as

$$P(v) = \sum_{u'} P(u') \sum_{u} P(v_j|pa_j, u^j, u')P(v_l|pa_l, u^l, u') \prod_{i \neq j, i \neq l} P(v_i|pa_i, u^i)P(u),$$
(147)

where U' represents the new unobserved variable. We construct two models M'_1 and M'_2 with the causal graph G' as

$$P^{M'_k}(v_i|pa_i, u^i) = P^{M_k}(v_i|pa_i, u^i), \ i \neq j, \ i \neq l, \quad k = 1, 2, \tag{148}$$

$$P^{M'_k}(v_i|pa_i, u^i, u') = P^{M_k}(v_i|pa_i, u^i), \ i = j, l, \quad k = 1, 2, \tag{149}$$

$$P^{M'_k}(u) = P^{M_k}(u), \quad k = 1, 2. \tag{150}$$

Again, if the pair (M_1, M_2) satisfies (141), so would the pair (M'_1, M'_2). Hence $P_t(s)$ is not identifiable in G'. $\qquad \square$

Next we prove the necessity part of Theorem 3.

Theorem *If there is a bidirected path connecting X to any of its children in G, then $P_x(v)$ is not identifiable.*

Proof: Let Y be a child of X and assume that there is a bidirected path connecting X and Y with variables $Z_1, \ldots, Z_k$ on the path (see Figure 18). We will prove that, for any $k \geq 1$, $P_x(y, z_1, \ldots, z_k)$ is not identifiable in the graph shown in Figure 18, which is a subgraph of G. By Lemma 13, if $P_x(y, z_1, \ldots, z_k)$ is not identifiable in a subgraph of G, then it is not identifiable in G, and therefore $P_x(v)$ is not identifiable in G.

Let $U = \{U_1, \ldots, U_{k+1}\}$. In Figure 18, we have

$$P(x, y, z_1, \ldots, z_k)$$
$$= \sum_{u} P(x|u_1)P(y|x, u_{k+1})P(z_1|u_1, u_2) \cdots P(z_k|u_k, u_{k+1})P(u_1) \cdots P(u_{k+1}),$$
(151)

50

and

$$P_x(y, z_1, \ldots, z_k)$$
$$= \sum_u P(y|x, u_{k+1})P(z_1|u_1, u_2) \cdots P(z_k|u_k, u_{k+1})P(u_1) \cdots P(u_{k+1}). \quad (152)$$

Let all variables $X, Y, Z_1, \ldots, Z_k, U_1, \ldots, U_{k+1}$ be binary variables. We will prove the nonidentifiability of $P_x(y, z_1, \ldots, z_k)$ by constructing two models such that in both models,

$$P(x, y, z_1, \ldots, z_k) = (1/2)^{k+2}, \quad \text{for all possible values of } x, y, z_1, \ldots, z_k, \quad (153)$$

while $P_x(y, z_1, \ldots, z_k)$ has different values in the two models. The construction involves the specification of all conditional probabilities in a parametric form, and shows two different parameterization both satisfying the set of 2^{k+2} equations in (153). We use the following parameterization, with five parameters, $a, b, c, d,$ and e.

$$P(u_i) = 1/2, \ u_i = 0, 1, \text{and } i = 1, \ldots, k+1 \quad (154)$$

x	u_1	$P(x|u_1)$
0	0	$1/2 + a$
0	1	$1/2 - a$

$\qquad(155)$

y	x	u_{k+1}	$P(y|x, u_{k+1})$
0	0	0	$1/2 + b$
0	0	1	$1/2 - b$
0	1	0	$1/2$
0	1	1	$1/2$

$\qquad(156)$

z_1	u_1	u_2	$P(z_1|u_1, u_2)$
0	0	0	$1/2 + c$
0	0	1	$1/2 - c$
0	1	0	$1/2 + d$
0	1	1	$1/2 - d$

$\qquad(157)$

51

z_i	u_i	u_{i+1}	$P(z_i\|u_i, u_{i+1})$
0	0	0	$1/2 + e$
0	0	1	$1/2 - e$
0	1	0	$1/2 - e$
0	1	1	$1/2 + e$

$$i = 2, \ldots, k. \qquad (158)$$

Substituting (154) into (151), Eq. (153) becomes

$$\frac{1}{2} = \sum_u P(x|u_1)P(y|x, u_{k+1})P(z_1|u_1, u_2) \cdots P(z_k|u_k, u_{k+1}). \qquad (159)$$

First we prove that if Eq. (159) is satisfied for $x = 0, y = 0, z_1 = 0, \ldots, z_k = 0$, then it is satisfied for all possible values of $x, y, z_1, \ldots, z_k$. We have that, for any a, b, c, d, e, the parameterization given in Eqs. (154)–(158) satisfies the following properties

$$\sum_{u_1} P(x|u_1) = 1. \qquad (160)$$

$$\sum_{u_{k+1}} P(y|x, u_{k+1}) = 1. \qquad (161)$$

$$\sum_{u_{i+1}} P(z_i|u_i, u_{i+1}) = 1, i = 1, \ldots, k. \qquad (162)$$

$$\sum_{u_i} P(z_i|u_i, u_{i+1}) = 1, i = 2, \ldots, k. \qquad (163)$$

(a) For $x = 1$ and any values of $y, z_1, \ldots, z_k$, Eq. (159) is satisfied:

$$\sum_u P(x = 1|u_1)P(y|x = 1, u_{k+1})P(z_1|u_1, u_2) \cdots P(z_k|u_k, u_{k+1})$$

$$= \frac{1}{2} \sum_u P(x = 1|u_1)P(z_1|u_1, u_2) \cdots P(z_k|u_k, u_{k+1}) \quad \text{(by } P(y|x = 1, u_{k+1}) = 1/2)$$

$$= \frac{1}{2} \quad \text{(by Eqs. (162) and (160))} \qquad (164)$$

(b) If, for a particular set of values $x, y, z_1, \ldots, z_k$, Eq. (159) is satisfied, then for the set of values $x, 1 - y, z_1, \ldots, z_k$, Eq. (159) is also satisfied:

$$\sum_u P(x|u_1)P(1 - y|x, u_{k+1})P(z_1|u_1, u_2)P(z_2|u_2, u_3) \cdots P(z_k|u_k, u_{k+1})$$

$$= \sum_u P(x|u_1)(1 - P(y|x, u_{k+1}))P(z_1|u_1, u_2)P(z_2|u_2, u_3) \cdots P(z_k|u_k, u_{k+1})$$

$$= \sum_u P(x|u_1)P(z_1|u_1, u_2)P(z_2|u_2, u_3) \cdots P(z_k|u_k, u_{k+1}) - \frac{1}{2} \quad \text{(by Eq. (159))}$$

$$= 1 - \frac{1}{2} \quad \text{(by Eqs. (162) and (160))}$$

$$= \frac{1}{2} \tag{165}$$

(c) If, for a particular set of values $x, y, z_1, \ldots, z_k$, Eq. (159) is satisfied, then for the set of values $x, y, z_1, \ldots, z_{i-1}, 1 - z_i, z_{i+1}, \ldots, z_k$, Eq. (159) is satisfied as well (for $i = 1, \ldots, k$):

$$\sum_u P(x|u_1)P(y|x, u_{k+1})P(z_1|u_1, u_2) \cdots P(1 - z_i|u_i, u_{i+1}) \cdots P(z_k|u_k, u_{k+1})$$

$$= \sum_u P(x|u_1)P(y|x, u_{k+1})P(z_1|u_1, u_2) \cdots P(z_{i-1}|u_{i-1}, u_i)P(z_{i+1}|u_{i+1}, u_{i+2})$$

$$\cdots P(z_k|u_k, u_{k+1}) - \frac{1}{2} \quad \text{(by Eq. (159))}$$

$$= 1 - \frac{1}{2} \quad \text{(by Eqs. (160)–(163))}$$

$$= \frac{1}{2} \tag{166}$$

From (a), (b), and (c), we obtain that if Eq. (159) is satisfied for $x = 0, y = 0, z_1 = 0, \ldots, z_k = 0$, then it is satisfied for all possible values of $x, y, z_1, \ldots, z_k$.

Next, we substitute the conditional probabilities given in Eqs. (154)–(158) into Eq. (159) for $x = 0, y = 0, z_1 = 0, \ldots, z_k = 0$. Define

$$f_{u_2, u_{k+1}} = \sum_{u_3, \ldots, u_k} P(z_2 = 0|u_2, u_3) \cdots P(z_k = 0|u_k, u_{k+1}) \tag{167}$$

53

We obtain

$$f_{00} = (1/2 + e)^{k-1} + \binom{k-1}{2}(1/2 + e)^{k-3}(1/2 - e)^2$$

$$+ \binom{k-1}{4}(1/2 + e)^{k-5}(1/2 - e)^4 + \cdots$$

$$= \sum_{i=0}^{i<k/2} \binom{k-1}{2i}(1/2 + e)^{k-1-2i}(1/2 - e)^{2i}. \tag{168}$$

From Eq. (163), we have

$$\sum_{u_2} f_{u_2, u_{k+1}} = 1. \tag{169}$$

From Eq. (162), we have

$$\sum_{u_{k+1}} f_{u_2, u_{k+1}} = 1. \tag{170}$$

Define

$$f = f_{00} - 1/2, \tag{171}$$

then $f_{u_2, u_{k+1}}$ is given as

u_2	u_{k+1}	$f_{u_2, u_{k+1}}$
0	0	$1/2 + f$
0	1	$1/2 - f$
1	0	$1/2 - f$
1	1	$1/2 + f$

Therefore, for $x = 0, y = 0, z_1 = 0, \ldots, z_k = 0$, Eq. (159) becomes

$$\frac{1}{2} = \sum_{u_1, u_{k+1}, u_2} P(x = 0|u_1)P(y = 0|x = 0, u_{k+1})P(z_1 = 0|u_1, u_2)f_{u_2, u_{k+1}}$$

$$= (1/2 + a)(1/2 + b)[(1/2 + c)(1/2 + f) + (1/2 - c)(1/2 - f)]$$
$$+ (1/2 + a)(1/2 - b)[(1/2 + c)(1/2 - f) + (1/2 - c)(1/2 + f)]$$
$$+ (1/2 - a)(1/2 + b)[(1/2 + d)(1/2 + f) + (1/2 - d)(1/2 - f)]$$
$$+ (1/2 - a)(1/2 - b)[(1/2 + d)(1/2 - f) + (1/2 - d)(1/2 + f)]$$
$$= 1/2 + 2bf(c + d + 2ac - 2ad) \tag{172}$$

which leads to

$$bf(c + d + 2ac - 2ad) = 0. \tag{173}$$

That is, with the parameterization given in (154)-(158), Eq. (153) holds if and only if Eq. (173) holds.

For $x = 0, y = 0, z_1 = 0, \ldots, z_k = 0$, $P_x(y, z_1, \ldots, z_k)$ is computed as

$$P_{x=0}(y = 0, z_1 = 0, \ldots, z_k = 0)$$
$$= \frac{1}{2^{k+1}} \sum_{u_1, u_{k+1}, u_2} P(y = 0 | x = 0, u_{k+1}) P(z_1 = 0 | u_1, u_2) f_{u_2, u_{k+1}}$$
$$= \frac{1}{2^{k+1}} [1 + 4bf(c + d)] \tag{174}$$

Let $-1/2 < e_0 < 1/2$ be a number such that $f \neq 0$ (see (171) and (168)). Consider the following two models:

Model 1 $a = 1/4, \ b = 0, \ c = d = 1/4, \ e = e_0$.

Model 2 $a = 1/4, \ b = 1/4, \ c = 1/12, \ d = -1/4, \ e = e_0$.

Eq. (173) holds in both models, hence the two models have the same distribution $P(x, y, z_1, \ldots, z_k) = (1/2)^{k+2}$. By Eq. (174), in Model 1, $P_{x=0}(y = 0, z_1 = 0, \ldots, z_k = 0) = (1/2)^{k+1}$, and in Model 2, $P_{x=0}(y = 0, z_1 = 0, \ldots, z_k = 0) = (1/2)^{k+1}(1 - f/6)$. Since $f \neq 0$, we have that $P_{x=0}(y = 0, z_1 = 0, \ldots, z_k = 0)$ takes different values in Model 1 and 2. Therefore $P_x(y, z_1, \ldots, z_k)$ is not identifiable. $\qquad \square$

B   Proof of Lemma 2

Lemma 2 $P_x(s)$ *is identifiable if and only if* $P_x(s)$ *is identifiable in the subgraph* $G_{An(S)}$.

Proof: (only if) By Lemma 13.

(if) Summing both sides of Eq. (6) over $V \setminus An(S)$, we have that the marginal distribution $P(an(S))$ decomposes exactly according to the graph $G_{An(S)}$. Hence if $P_x(s)$ is identifiable in $G_{An(S)}$, then it is computable from $P(an(S))$, and therefore is identifiable in G. $\qquad \square$

C Proof of Lemma 5

Lemma 5 *Let $A \subseteq B \subseteq V$. $Q[A]$ is computable from $Q[B]$ if and only if $Q[A]_{G_B}$ is computable from $Q[B]_{G_B}$.*

Proof: (only if) By Lemma 13.

(if) Proof by contradiction. Assume that $Q[A]$ is not computable from $Q[B]$, then there exist two models, M_1 and M_2, with the same causal graph G, satisfying

$$Q^{M_k}[B](b,c) = \sum_u \prod_{\{i|V_i \in B\}} P^{M_k}(v_i|pa'_i, c_i, u^i)P^{M_k}(u), \quad k = 1, 2, \tag{175}$$

where $PA'_i = PA_i \cap B$, $C_i = PA_i \setminus B$, and $C = \cup_i C_i$, such that

$$Q^{M_1}[B](b,c) = Q^{M_2}[B](b,c) > 0, \text{ for all values } b, c, \tag{176}$$

but

$$Q^{M_1}[A](b',c') \neq Q^{M_2}[A](b',c'), \text{ for some particular value } b', c'. \tag{177}$$

$Q[B]_{G_B}$ can be written as

$$Q[B]_{G_B}(b) = \sum_u \prod_{\{i|V_i \in B\}} P(v_i|pa'_i, u^i)P(u). \tag{178}$$

We construct two models, M'_1 and M'_2, with the same causal graph G_B as

$$P^{M'_k}(v_i|pa'_i, u^i) = P^{M_k}(v_i|pa'_i, C_i = c'_i, u^i), \quad k = 1, 2, \tag{179}$$

$$P^{M'_k}(u) = P^{M_k}(u), \quad k = 1, 2. \tag{180}$$

Then we have

$$Q[B]_{G_B}^{M'_k}(b) = Q[B]^{M_k}(b, c'), \text{ and } Q[A]_{G_B}^{M'_k}(b) = Q[A]^{M_k}(b, c'), \quad k = 1, 2. \tag{181}$$

From Eqs. (181), (176) and (177), we obtain

$$Q^{M'_1}[B]_{G_B}(b) = Q^{M'_2}[B]_{G_B}(b) > 0, \text{ for all values } b, \tag{182}$$

and

$$Q^{M'_1}[A]_{G_B}(b') \neq Q^{M'_2}[A]_{G_B}(b'), \text{ for the value } b', \tag{183}$$

which says that $Q[A]_{G_B}$ is not computable from $Q[B]_{G_B}$. $\qquad\square$