Ann Math Artif Intell (2008) 54:363–408
DOI 10.1007/s10472-008-9101-x

On the completeness of an identifiability algorithm for semi-Markovian models

Yimin Huang · Marco Valtorta

Published online: 13 May 2009

Abstract This paper addresses the problem of identifying causal effects from non-experimental data in a *causal Bayesian network*, i.e., a directed acyclic graph that represents causal relationships. The identifiability question asks whether it is possible to compute the probability of some set of (effect) variables given intervention on another set of (intervention) variables, in the presence of non-observable (i.e., hidden or latent) variables. It is well known that the answer to the question depends on the structure of the causal Bayesian network, the set of observable variables, the set of effect variables, and the set of intervention variables. Sound algorithms for identifiability have been proposed, but no complete algorithm is known. We show that the *identify* algorithm that Tian and Pearl defined for semi-Markovian models (Tian and Pearl 2002, 2002, 2003), an important special case of causal Bayesian networks, is both sound and complete. We believe that this result will prove useful to solve the identifiability question for general causal Bayesian networks.

Keywords Identifiability algorithm · Semi-Markovian models · Causal Bayesian network · Graphical models

Mathematics Subject Classifications (2000) 68T37 · 62B15 · 62F15 · 62-09

1 Introduction

This paper focuses on the feasibility of inferring the strength of cause-and-effect relationships from a causal graph [4], which is an acyclic directed graph expressing

Y. Huang · M. Valtorta (✉)
Department of Computer Science and Engineering,
University of South Carolina, Columbia, SC, USA
e-mail: mgv@cse.sc.edu

Y. Huang
e-mail: hym924@hotmail.com

nonexperimental data and causal relationships. Because of the existence of unmeasured variables, the following identifiability questions arise: "Can we assess the strength of causal effects from nonexperimental data and causal relationships? And if we can, what is the total causal effect in terms of estimable quantities?"

The questions just given can partially be answered using a graphical approach due to Pearl and his collaborators. More precisely, graphical conditions have been devised to show whether a causal effect, that is, the joint response of any set S of variables to interventions on a set T of action variables, denoted $P_T(S)$[1] is identifiable or not. Those results are summarized in [4]. For example, "back-door" and "front-door" criteria and *do*-calculus [5]; graphical criteria to identify $P_T(S)$ when T is a singleton [6]; graphical conditions under which it is possible to identify $P_T(S)$ where T and S are, possibly non-singleton, sets, subject to a special condition called Q-identifiability [7]. Some further study can be also found in [8] and [9].

Recently, J. Tian by himself and in collaboration with J. Pearl published a series of papers [1–3, 10] related to this topic. Their new methods combine the graphical characters of causal graph and the algebraic definition of causal effect. They used both algebraic and graphical methods to identify causal effects. The basic idea is first to transfer causal graphs to semi-Markovian graphs [2], then to use Algorithm 2 in [3] to calculate the causal effects we want to know.

Tian and Pearl's method is a great contribution to this study area, but there are still two open questions left. First, even though we believe, as Tian and Pearl do, that the semi Markovian models obtained from the transforming Projection algorithm in [2] are equal to the original causal graphs, and therefore the causal effects should be the same in both models, still, to the best of our knowledge, there is no formal proof for this equivalence. Second, the completeness question of the indentification algorithm in [3] (which we will simply call the *identify* algorithm from now on) is still open, so that it is unknown whether a causal effect is identifiable if the identify algorithm fails.

In this paper, we focus on the second question. Our conclusion shows that Tian and Pearl's identify algorithm on semi-Markovian models is sound and complete, which means that a causal effect on a semi-Markovian model is identifiable if and only if the given algorithm can run successfully and finally return an expression which is the target causal effect in terms of estimable quantities.

Using the result of this paper, it becomes possible to rewrite the identify algorithm on general Markovian models and prove that the new algorithm is still sound and complete. This work is not included in this paper, but we believe that we provide the foundations for the more general result.

In the next section we present the definitions and notation that we use in this paper. In section three, we present some important lemmas that will be used to support the analysis of the identify algorithm. In section four, we describe the algorithm that answers the identifiability question for a special causal effect case ($Q[S]$), and show that the algorithm is sound and complete. We present the identify algorithm for general causal effect $P_T(S)$ in section five and show that it is also sound and complete. Conclusions are included in section six.

[1] Pearl and Tian used notation $P(s|do(t))$ and $P(s|\hat{t})$ in [4] and $P_t(s)$ in [2, 3].

2 Definitions and notations

Markovian models are popular graphical models for encoding distributional and causal relationships. A Markovian model consists of a DAG G over a set of variables $V = \{V_1, \ldots, V_n\}$, called a *causal graph* and a probability distribution over V, which has some constraints on it that will be specified precisely below. We use $V(G)$ to indicate that V is the variable set of graph G. If it is clear in the context, we also use V directly. The interpretation of such kind of model consists of two parts. The probability distribution must satisfy two constraints. The first one is that each variable in the graph is independent of all its non-descendants given its direct parents. The second one is that the directed edges in G represent causal influences between the corresponding variables. A Markovian model for which only the first constraint holds is called a *Bayesian network*. This explains why Markovian models are also called causal Bayesian networks. As far as the second condition is concerned, some authors prefer to consider (3) (below) as definitional; others take (3) as following from more general considerations about causal links, and in particular the account of causality that requires that, when a variable is set, the parents of that variable be disconnected from it. See [11] and [4].

In this paper, capital letters, like V, are used for variable sets; lower-case letters, like v, stand for the instances of variable set V. Capital letters like X, Y and V_i are also used for single variables, and their values can be x, y and v_i. Normally, we use $F(V)$ to denote a function on variable set V. An instance of this function is denoted as $F(V)(V = v)$, or $F(V)(v)$, or just $F(v)$. Because each variable is in one-to-one correspondence to one node in the causal graph, we sometimes use node or node set instead of variable and variable set.

We use $Pa(V_i)$ to denote parent node set of variable V_i in graph G and $pa(V_i)$ as an instance of $Pa(V_i)$. $Ch(V_i)$ is V_i's children node set; $ch(V_i)$ is an instance of $Ch(V_i)$.

Based on the probabilistic interpretation, we get that the joint probability function $P(v) = P(v_1, \ldots, v_n)$ can be factorized as

$$P(v) = \prod_{V_i \in V} P(v_i|pa(V_i)) \tag{1}$$

The causal interpretation of Markovian model enables us to predict the intervention effects. Here, intervention means some kind of modification of factors in product (1). The simplest kind of intervention is fixing a subset $T \subseteq V$ of variables to some constants t, denoted by $do(T = t)$ or just $do(t)$, and then the post-intervention distribution

$$P_T(V)(T = t, V = v) = P_t(v) \tag{2}$$

is given by:

$$P_t(v) = P(v|do(t)) = \begin{cases} \prod_{V_i \in V \setminus T} P(v_i|pa(V_i)) & v \text{ consistent with } t \\ 0 & v \text{ inconsistent with } t \end{cases} \tag{3}$$

We note explicitly that the post-intervention distribution $P_T(V)(T = t, V = v) = P_t(v)$ is a probability distribution.

When all the variables in V are observable, since all $P(v_i|pa(V_i))$ can be estimated from nonexperimental data, all causal effects are computable. But when some variables in V are unobservable, the situation is much more complex.

Let $N(G)$ and $U(G)$ (or simply N and U when the graph is clear from the context) stand for the sets of observable and unobservable variables in graph G respectively, that is $V = N \cup U$. The observed probability distribution $P(n) = P(N = n)$, is a mixture of products:

$$P(n) = \sum_{U_k \in U} \prod_{V_i \in V} P(v_i|pa(V_i)) = \sum_{U_k \in U} \prod_{V_i \in N} P(v_i|pa(V_i)) \prod_{V_j \in U} P(v_j|pa(V_j)) \tag{4}$$

The post-intervention distribution $P_t(n) = P_{T=t}(N = n)^2$ is defined as:

$$P_t(n) = \begin{cases} \sum_{U_k \in U} \prod_{V_i \in N \setminus T} P(v_i|pa(V_i)) \prod_{V_j \in U} P(v_j|pa(V_j)) & n \text{ consistent with } t \\ 0 & n \text{ inconsistent with } t \end{cases} \tag{5}$$

Sometimes what we want to know is not the post-intervention distribution for the whole N, but the post-intervention distribution $P_t(s)$ of an observable variable subset $S \subseteq N$. For those two observable variable sets S and T, $P_t(s) = P_{T=t}(S = s)$ is given by:

$$P_t(s) = \begin{cases} \sum_{V_l \in (N \setminus S) \setminus T} \sum_{U_k \in U} \prod_{V_i \in N \setminus T} P(v_i|pa(V_i)) \prod_{V_j \in U} P(v_j|pa(V_j)) & s \text{ consistent with } t \\ 0 & s \text{ inconsistent with } t \end{cases} \tag{6}$$

The identifiability question is defined as whether the causal effect $P_T(S)$, that is all $P_t(s)$ given by (6), can be determined uniquely from the distribution $P(N = n)$ given by (4), and thus independently of the unknown quantities $P(v_i|pa(V_i))$s, where $V_i \in U$ or $V_j \in U$ for some $V_j \in Pa(V_i)$.

We give a formal definition of *identifiability* below, which follows [3].

A Markovian model consists of four elements

$$M = < N, U, G_{N \cup U}, P(v_i|pa(V_i)) >$$

where, (i) N is a set of observable variables; (ii) U is a set of unobservable variables; (iii) G is a directed acyclic graph with nodes corresponding to the elements of $V = N \cup U$; and (iv) $P(v_i|pa(V_i))$, is the conditional probability of variable $V_i \in V$ given its parents $Pa(V_i)$in G.

Definition 1 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 all the quantities $P_t(s)$ 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 $P^{M_1}(n) = P^{M_2}(n) > 0$ and $G(M_1) = G(M_2)$.

[2] In this paper, we only consider the situation in which $T \subseteq N$.

This definition means that, given the causal graph G, the quantity $P_t(s)$ can be determined from the observed distribution $P(n)$ alone; the probability tables that include unobservable variables are irrelevant.

Next, we define $Q[S]$ function and c-components in causal graphs. These definitions follow [2].

Normally, when we talk about S and T, we think they are both observable variable subsets of N and mutually disjoint. So, any configuration of S is consistent with any configuration of T, and (6) can be replaced by

$$P_t(s) = \sum_{V_l \in (N \backslash S) \backslash T} \sum_{U_k \in U} \prod_{V_i \in N \backslash T} P(v_i | pa(V_i)) \prod_{V_j \in U} P(v_j | pa(V_j)) \quad (7)$$

From now on, we will use this definition instead of (6).

We are sometimes interested in the causal effect on a set of observable variables S due to all other observable variables. In this case, keeping the convention that N stands for the set of all observable variables and T stands for the set of variables whose effect we want to compute, $T = N \backslash S$, and (7) simplifies to

$$P_{n \backslash s}(s) = \sum_{U_k \in U} \prod_{V_i \in S} P(v_i | pa(V_i)) \prod_{V_j \in U} P(v_j | pa(V_j)) \quad (8)$$

In formula (8), the subscript $n \backslash s$ indicates a configuration of the variable or variables in the set $N \backslash S$. For convenience and for uniformity with [2], we define

$$Q[S] = P_{N \backslash S}(S) \quad (9)$$

and interpret this equation as stating that the causal effect of $N \backslash S$ on S is $Q[S]$.

Note that $Q[S]$ is identifiable if $Q[S]^{M_1}(s) = Q[S]^{M_2}(s)$ for every pair of models M_1 and M_2 with $Q[N]^{M_1}(n) = Q[N]^{M_2}(n) > 0$ and $G(M_1) = G(M_2)$.

We define the *c-component relation* on the unobserved variable set U of graph G as follows. For any $U_1 \in U$ and $U_2 \in U$, they are related under the c-component relation if and only if one of conditions below is satisfied:

(i) there is an edge between U_1 and U_2,
(ii) U_1 and U_2 are both parents of the same observable node,
(iii) both U_1 and U_2 are in the c-component relation with respect to another node $U_3 \in U$.

Observe that the c-component relation in U is reflexive, symmetric and transitive, so it defines a partition of U. Based on this relation, we can therefore divide U into disjoint and mutually exclusive c-component related parts.

A *c-component* (short for "confounded component," [3]) of variable set V on graph G consists of all the unobservable variables belonging to the same c-component related part of U and all observable variables that have an unobservable parent which is a member of that c-component. According to the definition of c-component relation, it is clear that an observable node can only appear in one c-component. If an observable node has no unobservable parent, then itself is a c-component on V. Therefore, the c-components form a partition on all of the variables.

For any pair of variables V_1 and V_2 in causal graph G, if there is an unobservable node U_i which is a parent for both of them, then path $V_1 \leftarrow U_i \rightarrow V_2$ is called a *bidirected link*.[3] If for nodes $V_1, \ldots, V_n$, there are bidirected links between all $V_i, V_{i+1}, 1 \leqslant i < n$, then we say there is a *bidirected path* from V_1 to V_n.

We now introduce a way of reducing the size of causal graphs that preserves the answer to an identifiability question. It is more convenient to work with the reduced graphs than with the original, larger ones. Studying definition (4) and (5), we can see if there is an unobservable variable in graph G that has no child, then it can be summed out in both (4) and (5) and removed. Formally, if we have a model $M =< N, U, G_{N \cup U}, P(v_i|pa(V_i)) >$, $U' \in U$ and U' has no child in $G_{N \cup U}$, then the identification problem in M is equal to the identification problem in $M' =< N, U \backslash \{U'\}, G', P'(v_i|pa_i) >$, where G' is the subgraph of $G_{N \cup U}$ obtained by removing node U' and all links attached with it. $P'(v_i|pa(V_i))$ is obtained by removing all $P(u'|pa(U'))$ in the set of conditional probability tables $P(v_i|pa(V_i))$. The overall distribution (of all remaining variables) and the causal distribution (of only the observable variables) in these two models are still the same.

By repeating the transformation given above, any causal model can be transformed to a model in which each unobservable variable is an ancestor of one or more observable variables without changing the identifiability property. (This is analogous to barren node removal in Bayesian networks.) From now on in this paper, we assume that all models we study satisfy this property.

If in a Markovian model each unobserved variable is a root node with exactly two observed children, we call it a semi-Markovian model. Verma [12] defines a projection by which every Markovian model on graph G can be transferred to a semi-Markovian model on graph $PJ(G, V)$. Tian and Pearl [2] show that G and $PJ(G, V)$ have the same topological relations over V and the same partition of V into c-components. They conclude that if $P_T(S)$ is identified in $PJ(G, V)$, then it is identified in G with the same expression. This is a very important statement. From now on in this paper we will just deal with semi-Markovian models.

In semi-Markovian models, (7) can be rewritten as:

$$P_t(s) = \sum_{V_l \in (N \backslash S) \backslash T} \sum_{U_k \in U} \prod_{V_i \in N \backslash T} P(v_i|pa(V_i)) \prod_{V_j \in U} P(v_j) \qquad (10)$$

And (8) can be rewritten as:

$$P_t(s) = Q[s] = P_{n \backslash s}(s) = \sum_{U_k \in U} \prod_{V_i \in S} P(v_i|pa(V_i)) \prod_{V_j \in U} P(v_j) \qquad (11)$$

As in Tian and Pearl [3], for the sake of convenience, we represent a semi-Markovian model with a causal graph G without showing the elements of U explicitly, but represent the confounding effects of U variables using bidirected edges. We explicitly represent U nodes only when it is necessary.

So, from know on, unless otherwise noted, all the nodes we mention are observable nodes in graph G. We still use N to denote the set of observable nodes.

We conclude this section by giving several simple graphical definitions that will be needed later. For a given variable set $C \subseteq N$, let G_C denote the subgraph of G

[3] We use this term because the three-node structure can be replaced by the two observable nodes with a special *bidirected* edge between them.

composed only of variables in C and all the bidirected links between variable pairs in C. We define $An(C)$ be the union of C and the set of observable ancestors of the variables in C in graph G and $De(C)$ be the union of C and the set of observable descendents of the variables in C in graph G.

An observable variable set $S \subseteq N$ in graph G is called an *ancestral set* if it contains all its own observed ancestors (i.e., $S = An(S)$).

3 Theorems and lemmas

Because our definition of $Q[S]$ is equal to the definition of $Q[S]$ in [2], Lemma 1 in [2] is still correct, and therefore we have:

Theorem 1 *Let $W \subseteq C \subseteq N$. If W is an ancestral set in G_C, then*

$$\sum_{V_i \in C \setminus W} Q[C] = Q[W] \tag{12}$$

We recall that subgraph G_C includes all variables in C and the subset of the unobservable variables in G for which their children are all in C. The lemma says that in such a subgraph, if W is a set of observable variables whose ancestor set includes no other observable variables in the subgraph, then $Q[W]$ can be calculated directly from $Q[C]$ by marginalizing variables in $C \setminus W$. In particular, note that if $Q[C]$ is identifiable, then $Q[W]$ is also identifiable. We will exploit this observation later on.

Another very important theorem is given in [2]. We only use the first two parts of it, which are:

Theorem 2 *Let $H \subseteq N$, and let $H'_1, \ldots, H'_l$ be c-components in the subgraph G_H. Let $H_i = H'_i \cap N$, $1 \leq i \leq l$. Then we have*

(i) *$Q[H]$ can be decomposed as*

$$Q[H] = \prod_{i=1}^{l} Q[H_i] \tag{13}$$

(ii) *Each $Q[H_i]$ is computable from $Q[H]$. Let k be the number of variables in H, and let a topological order of variables in H be $V_1 < \ldots < V_k$ in G_H. Let $H^{(i)} = \{V_1, \ldots, V_i\}$ be the set of variables in H ordered before V_i (including V_i), $i = 1, \ldots, k$, and $H^{(0)} = \phi$. Then each $Q[H_j], j = 1, \ldots, l$, is given by*

$$Q[H_j] = \prod_{\{i | V_i \in H_j\}} \frac{Q[H^{(i)}]}{Q[H^{(i-1)}]} \tag{14}$$

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

$$Q[H^{(i)}] = \sum_{H \setminus H^{(i)}} Q[H] \tag{15}$$

Theorem 2 means that if $Q[H]$ is identifiable, then each $Q[H_i]$, $1 \leq i \leq l$, is also identifiable. In the special case for which $H = N$, $Q(H) = Q(N) = P(N)$, which is obviously identifiable, and therefore Theorem 2 implies that $Q[N'_i \cap N]$ is always identifiable for each c-component N'_i of a given causal graph G.

Lemma 1 *Let $S, T \subset N$ be two disjoint sets of observable variables. If $P_T(S)$ is not identifiable in G, then $P_T(S)$ is not identifiable in the graph resulting 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 resulting from removing a directed or bidirected edge from G.*

Intuitively, this lemma says that unidentifiability does not change by adding links. This property is mentioned in [4]. A formal proof of this lemma for semi-Markovian model can be found in [3].

Lemma 2 *Let $S, T \subset N$ be two disjoint sets of observable variables. If S_1 and T_1 are subsets of S, T, and $P_{T_1}(S_1)$ is not identifiable in a subgraph of G, which does not include nodes $S\backslash S_1 \cup T\backslash T_1$, then $P_T(S)$ is not identifiable in the graph G.*

Proof Assume that $P_{T_1}(S_1)$is not identifiable in a subgraph of G, which we will name G', and which does not include nodes $S\backslash S_1 \cup T\backslash T_1$. We can add all nodes in G but not in G' into G' as isolated nodes. Then we have (trivially) that $P_T(S)$ is not identifiable in this new graph. According to Lemma 1, $P_T(S)$ is not identifiable in graph G either. □

Lemma 3 *Let $A \subset B \subset N$. $Q[A]$ is computable from $Q[B]$ if and only if $Q[A]_{G_B}$ is computable from $Q[B]_{G_B}$*

Recall that $Q[A] = P_{V\backslash A}(A)$. The only if part of this lemma follows from Lemma 2. A formal proof of the if part can be found in [3].

4 Identify algorithm for $Q[S]$

Let S be a subset of observable variables (i.e., $S \subset N$). Recall that $Q[S] = P_{N\backslash S}(S)$. Based the theorems in the previous section, Tian and Pearl [3] gave an algorithm to solve the identifibility problem of $Q[S]$ and showed that this algorithm is sound. We present their algorithm here and show that it is also complete. We begin with a lemma.

Lemma 4 *Assume that N is partitioned into c-components $N_1, \ldots, N_k$ in G, and S is partitioned into c-components $S_1, \ldots, S_l$ in graph G_S. Because each S_j, $j = 1, \ldots, l$, is a c-component in G_S, which is a subgraph of G, it must be included in exactly one N_j, $N_j \in \{N_1, \ldots, N_k\}$. $Q[S]$ is identifiable if and only if each $Q[S_j]$ is identifiable in graph G_{N_j}.*

Proof First note that, because of Theorem 2 (part i), in any model on graph G, we have

$$Q[S] = \prod_{j=1}^{l} Q[S_j] \tag{16}$$

Only if part:

From Lemma 3, it follows that, if each $Q[S_j]$ is identifiable in graph G_{N_j}, then each $Q[S_j]$ is identifiable from $Q[N_j]$ on graph G. When we have $Q[N]$, according to Theorem 2 (part ii), we can compute all the $Q[N_j]$s. So, each $Q[S_j]$ is identifiable from $Q[N]$. Based on (16), $Q[S]$ is identifiable.

If part:

If one $Q[S_j]$ is unidentifiable in $Q[N_j]$ in graph G_{N_i}, then from Lemma 2, we have $Q[S]$ is unidentifiable. □

Let us now consider how to compute $Q[S_j]$ from $Q[N_j]$. This discussion will lead to an algorithm, expressed below as function identify.

Let $F = An(S_j)_{G_{N_j}}$.

If $F = S_j$, that is, if S_j is an ancestral set in G_{N_j}, then by Theorem 1, $Q[S_j]$ is computable as: $Q[S_j] = \sum_{N_j \setminus S_j} Q[N_j]$.

If $F = N_j$, we will prove (Theorem 3, below) that $Q[S_j]$ is not identifiable in G_{N_j}.

If $S_j \subset F \subset N_j$, by Theorem 1, we know $Q[F] = \sum_{N_j \setminus F} Q[N_j]$.

Assume that in the graph G_F, S_j is contained in a c-component H. Note that S_j must belong to one c-component. By Theorem 1, $Q[H]$ is computable from $Q[F]$ and is given by $Q[H] = \sum_{H \setminus S_j} Q[F]$. We obtain that the problem of whether $Q[S_j]$ is computable from $Q[N_i]$ is reduced to whether $Q[S_j]$ is computable from $Q[H]$.

Based on Lemma 3, we know that $Q[S_j]$ is computable from $Q[N_j]$ if and only if $Q[S_j]$ is computable from $Q[N_j]$ in G_{N_j}.

Using Lemma 3 again, we know that $Q[S_j]$ is computable from $Q[N_j]$ in G_{N_j} if and only if $Q[S_j]$ is identifiable form $Q[H]$ in graph G_H.

We now restate Tian and Pearl's algorithm [3] to obtain $Q[C]$ from $Q[T]$.

Function Identify (C,T,Q)

INPUT: $C \subseteq T \subseteq N$, $Q = Q[T]$, G_T and G_C are both composed of one single c-component.

OUTPUT: Expression for $Q[C]$ in terms of Q or FAIL.

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

i) *If $A = C$, output $Q[C] = \sum_{T \setminus C} Q[T]$.*
ii) *If $A = T$, output FAIL.*
iii) *If $C \subset A \subset T$*

1. Assume that in G_A, C is contained in a c-component T_1.
2. Compute $Q[T_1]$ from $Q[A] = \sum_{T \setminus A} Q[T]$ with Theorem 2
3. Output Identify(C,T_1,$Q[T_1]$).

From the discussions above, we know that cases i) and iii) are correct. Case ii) is handled by the theorem below.

Theorem 3 *In a semi-Markovian graph G, if*

1. *G itself is a c-component, and*
2. $S \subset N$ *in G, and* G_S *has only one c-component, and*
3. *All variables in* $N \backslash S$ *are ancestors of S,*

then $Q[S]$ *is unidentifiable in G.*

The proof of this theorem is in Appendix A.

Based on the analysis above we have

Theorem 4 *The identify algorithm for computing* $Q[S]$ *in causal graph G is sound and complete.*

From Theorem 4 above, the corollaries below follow.

Corollary 1 *Let* $S \subset N$ *in graph G, e be an outgoing link from one S node, and graph* G' *be the same as graph G except that it does not have link e. Then* $Q[S]$ *is identifiable in graph G if and only if* $Q[S]$ *is identifiable in graph* G'.

Proof Since e is a link exiting an S node, graph G and G' have the same c-component partition. Any c-component in G is also a c-component in G', and vice versa. Graph G_S and G'_S also have the same c-component partition. Any c-component in G_S is also a c-component in G'_S, and vice versa. From Algorithm Identify(C,T,Q), Algorithm Computing $Q[S]$, and Theorem 4, we know that $Q[S]$ is identifiable in graph G if and only if $Q[S]$ is identifiable in graph G'. □

From Corollary 1, we have the following, which will be used in the next section:

Corollary 2 *Let* $S \subset N$ *in graph G and graph* G' *be obtained by removing all outgoing links from S nodes in graph G. Then* $Q[S]$ *is identifiable in graph G if and only if* $Q[S]$ *is identifiable in graph* G'.

5 Identify algorithm for $P_T(S)$

Lemma 5 *Assume* $S \subset N$ *and* $T \subset N$ *are disjunct node sets in graph G,* $< X_1, X_2 >$ *is a directed link in G,* $X_1 \in S$, *and* $X_2 \in S$. *Assume that graph* G' *is obtained by removing link* $< X_1, X_2 >$ *from graph G. If* $P_T(S)$ *is unidentifiable in graph* G', *then* $P_T(S \backslash \{X_1\})$ *is unidentifiable in G.*

The proof of this lemma is in Appendix B.

A *direct ancestor set* of S in G is a variable set D such that $S \subseteq D \subseteq N$, and if node $X \in D$, then $X \in S$ or there is a directed path from X to a node in S, and all the nodes on that path are in D.

Lemma 6 *Assume D is a direct ancestor set of node set S on graph G.* $\sum_{D \backslash S} Q[D]$ *is identifiable if and only if* $Q[D]$ *is identifiable.*

Proof If part:

By definition, if $Q[D]$ is identifiable, $\sum_{D\backslash S} Q[D]$ is identifiable.

If $Q[D]$ is unidentifiable, then we know from Corollary 2 that $Q[D]$ is unidentifiable in graph G', where G' is obtained by removing from G all outgoing links from nodes in D.

Since D is a directed ancestor set of S, we can find an order of nodes in $D\backslash S$, say $X_1, \ldots, X_k$, for which X_i, $1 \leq i \leq k$, is a parent of at least one node in $S \cup \{X_1, \ldots, X_{i-1}\}$ in graph G. Assume that for X_i, $1 \leq i \leq k$, the link outgoing from X_i that is removed from G to get G' is e_i,that graph G_i is obtained by adding link e_i to graph G_{i-1}, and that $G_0 = G'$.

Note that $Q[D] = P_{N\backslash D}(D)$ is unidentifiable in G'. From Lemma 5, $P_{N\backslash D}(D\backslash\{X_1\})$ is unidentifiable in graph G_1. Using this lemma again, we have $P_{N\backslash D}(D\backslash\{X_1, X_2\})$ is unidentifiable in graph G_2, and repeating, we have $P_{N\backslash D}(S)$ is unidentifiable in graph G_k. Since G_k is a subgraph of G, according to Lemma 1, $P_{N\backslash D}(S)$ is unidentifiable in G too. and $P_{N\backslash D}(S) = \sum_{D\backslash S} P_{N\backslash D}(D) = \sum_{D\backslash S} Q[D]$. □

Based on the lemmas above, we can get a general algorithm to solve the identifiability problem on semi-Markovian models.

Let variable set N in causal graph G be partitioned into c-components $N_1, \ldots, N_k$, and S and T be disjoint observable variable sets in G. According to Theorem 2, we have

$$P(N) = Q[N] = \prod_{i=1}^{k} Q[N_i] \tag{17}$$

where each $Q[N_i]$, $1 \leq i \leq k$ is computable from $Q[N]$.

What we want to compute is:

$$P_t(s) = \sum_{N\backslash(T\cup S)} P_t(n\backslash t) = \sum_{N\backslash(T\cup S)} Q[N\backslash T] \tag{18}$$

Algorithm Identify

INPUT: two disjoint observable variable sets $S, T \subset N$. OUTPUT: the expression for $P_T(S)$ or FAIL.

1. *Find all c-components of G:$N_1, \ldots, N_k$.*
2. *Compute all $Q[N_i]$,$1 \leq i \leq k$, by Theorem 2.*
3. *Let $D = An(S)_{G_{N\backslash T}}$*
4. *Let c-components in graph G_D be $D_1, \ldots, D_l$.*
5. *For each D_j,$1 \leq j \leq l$, where $D_j \subseteq N_i$, $1 \leq i \leq k$, we compute $Q[D_k]$ by calling the function identify(D_j,N_i,$Q[N_i]$). If the function returns FAIL, then stop and output FAIL.*
6. *Output $P_T(S) = \sum_{D\backslash S} \prod_{j=1}^{l} Q[D_j]$*

Let $D = An(S)_{G_{N\backslash T}}$. Since D is an ancestral set in graph $G_{N\backslash T}$, Theorem 1 allows us to conclude that $\sum_{N\backslash(T\cup D)} Q[N\backslash T] = Q[D]$. Therefore, we can rewrite $P_t(s)$ from (18) as:

$$P_t(s) = \sum_{N\backslash(T\cup S)} Q[N\backslash T] = \sum_{D\backslash S} \sum_{N\backslash(T\cup D)} Q[N\backslash T] = \sum_{D\backslash S} Q[D] \tag{19}$$

Since D is a directed ancestor set of S, according to Lemma 6, $\sum_{D\backslash S} Q[D]$ is identifiable if and only if $Q[D]$ is identifiable. Now the identifiability problem of $P_T(S)$ is transferred to the identifiability problem of $Q[D]$, which can be solved by the algorithm in the last section.

Summarizing the discussion following Lemma 6, we present the identify algorithm [3].

Our discussion above shows:

Theorem 5 *The identify algorithm for computing $P_T(S)$ is sound and complete.*

6 Conclusion

We prove that the identification algorithm given by J.Tian and J.Pearl, which can be used on semi-Markovian graphs, a special case of causal Bayesian networks, is complete. This complements the proof of soundness in [3] and is a stepping stone towards the solution of the longstanding problem of finding a sound and complete algorithm for the general identifiability question in general Bayesian networks. We conjecture that a straightforward extension of the same algorithm is sound and complete for general causal Bayesian networks.

Appendix A: Proof of Theorem 3

Recall that G_S is the subgraph of G that includes all nodes in the observable node set S and all bidirected links between two S nodes.

Theorem 3 *In a semi-Markovian graph G, if*

1. *G itself is a c-component.*
2. *$S \subset N$ in G, and G_S has only one c-component.*
3. *All variables in $N\backslash S$ are ancestors of S.*

then $Q[S]$ is unidentifiable in G.

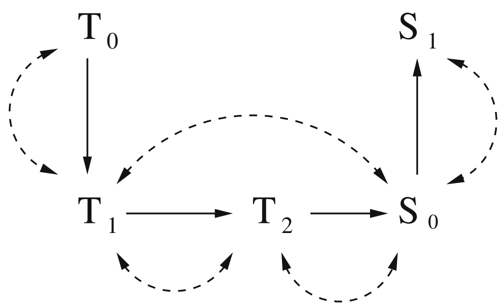


Fig. 1 A graph that satisfies the properties of Theorem 3

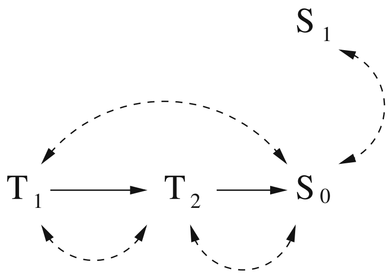


Fig. 2 An unidentifiable subgraph of Fig. 1

See Fig. 1 for an example of a graph that has the three properties in the premise of Theorem 3. Tian and Pearl [3] have proved that this theorem is true when T just includes one node. Here we show that this theorem is true in the general case.

A.1 General unidentifiable subgraph

For a given G that satisfies the properties given in Theorem 3, assume G' is a subgraph of G that satisfies the three properties below

1. G' is a c-component.
2. Let the observable node set in G' be N', and let $S' = N' \cap S$. Then, S' is not empty and $G_{S'}$ is a c-component.
3. $N' \setminus S'$ is not empty and all nodes in $N' \setminus S'$ are ancestors of S' in G'.

Then we say that G' is an *unidentifiable subgraph* of G. From Lemma 1 and Lemma 2, if $Q[S']$ is unidentifiable in G', $Q[S]$ is unidentifiable in G. See Fig. 2 for an example.

Assume G^m is an unidentifiable subgraph of G and no subgraph of G^m obtained by removing edges from G^m is an unidentifiable subgraph of G. We say G^m is a *general unidentifiable subgraph*. See Fig. 3 for an example. For any semi-Markovian graph G we study here, we can find at least one general unidentifiable subgraph, and we may therefore focus on general unidentifiable subgraphs.

From now on, in this appendix, we assume the graph G we studying is a general unidentifiable subgraph.

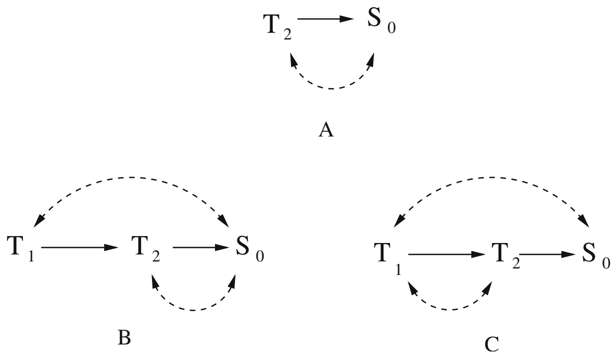


Fig. 3 Three general unidentifiable subgraphs of Fig. 1

Any general unidentifiable subgraph has the four properties below:

Graph Property 1 *If we take each bidirected link as an edge, then G_N by itself is a free tree.*

Recall that a free tree (sometimes called an unrooted tree) is a connected undirected graph with no cycles. Note that G_N can be obtained by removing all links between observable nodes from G. This property says that graph G is connected by bidirected links. If $|N| = m$, then we have just $m - 1$ bidirected links in G.

Graph Property 2 *If we take each bidirected link as an edge in G_S, then G_S by itself is a free tree.*

This property says that subgraph G_S is also connected by bidirected links. If $|S| = n$, then we have just $n - 1$ bidirected links in G_S.

Graph Property 3 *For each $T_i \in T = N(G)\backslash S$, there is a unique directed path from it to an S node.*

This property is true because if there are two paths, we can break one of them and T_i is still an ancestor of S, so G is not a general unidentifiable subgraph.

This property also tells us there are just $|T|$ directed links in G, and each T_i has just one directed link out from it.

Graph Property 4 *There are no directed links out of S nodes.*

A.2 Extension of S node

From graph property 4, we know that no node in S has outgoing links. But there are three kind of links that can enter an S node S_j. The first type includes directed links from T nodes to S_j, the second type includes bidirected links between T nodes and S_j, and the third type includes bidirected links between S_j and other S nodes.

Lemma 7 *Assume that e is a first type or second type link into node $S_j \in S$. Add an extra S node S'_j to graph G, make e point to S'_j instead of S_j and add a bidirected link*

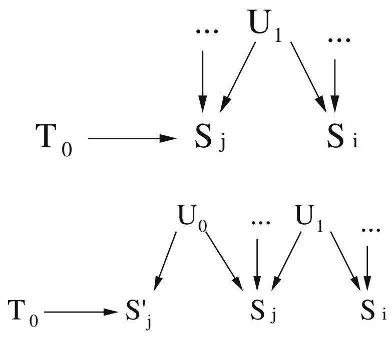


Fig. 4 S node extension: case A

between S_j and S'_j. Call the new graph G'. If $Q[S \cup \{S'_j\}]$ is unidentifiable in G' then $Q[S]$ is unidentifiable in G.

Proof Note that in G', S'_j has only two links into it. One is the e we are dealing with and the other is the bidirected link between S_j and S'_j.

A) If e is a first type link, then we conclude that for S_j in G', in addition the bidirected link between S_j and S'_j, S_j has at least one other bidirected link get into it. (See Fig. 4).

In G', we call the observable parent of S'_j T_0, the unobservable node on the bidirected link between S'_j and S_j U_0, and the another unobservable node, which is a parent of S_j, U_1. U_1 has two observable children: one is S_j, and the other we call S_i. (See Fig. 4).
If $Q[S \cup \{S'_j\}]$ is unidentifiable in G' then we have two models M_1 and M_2 on G' that

$$P^{M_1}(N(G) \cup \{S'_j\}) = P^{M_2}(N(G) \cup \{S'_j\}) \tag{20}$$

but for some (t, s, s'_j),

$$P_t^{M_1}(s, s'_j) \neq P_t^{M_2}(s, s'_j), \tag{21}$$

where t is an instance of variable set $T(G) = N(G) \backslash S(G)$, s is an instance of variable set S, s'_j is a value for variable S'_j. We assume in s, $S_i = s_i$ and $S_j = s_j$.

Now, we create two models M'_1 and M'_2 on graph G based on models M_1 and M_2.

For any node X in G, which is not in $\{S_j, U_1, S_i\}$, $k = 1, 2$, we define

$$P^{M'_k}(x|pa(X)) = P^{M_k}(x|pa(X)) \tag{22}$$

The state space of S_j in M'_k is given by $S(S'_j) \times S(S_j)$, where $S(S'_j)$ and $S(S_j)$ are the state spaces of S'_j and S_j in M_k.

Note that the parent set of S_j in G is the parent set of S_j in G' minus U_0 plus T_0.

The state space of U_1 in M'_k is defined as $S(U_0) \times S(U_1)$, where $S(U_0)$ and $S(U_1)$ are the state spaces of U_0 and U_1 in M_k.

The state space of node S_i in M'_k is the same as the state space of S_i in M_k.
Now we define:

$$P^{M'_k}(u'_1) = P^{M'_k}((u_0, u_1)) = P^{M_k}(u_0) \times P^{M_k}(u_1) \tag{23}$$

Here u'_1 is an instance of U_1 in M'_k, u_0 and u_1 are instances for U_0 and U_1 in M_k. We define

$$\begin{aligned} &P^{M'_k}((s_j, s'_j)|t_0, (u_0, u_1), pa'(S_j)) \\ &\quad = P^{M_k}(s'_j|t_0, u_0) \times P^{M_k}(s_j|u_0, u_1, pa'(S_j)), \end{aligned} \tag{24}$$

where $pa'(S_j)$ is an instance of the parent set of S_j in G except for U_1 and T_0. Note that $pa'(S_j)$ is also an instance of parent set of S_j in G' except U_0 and U_1.

We also define

$$P^{M'_k}(s_i|pa'(S_i), (u_0, u_1)) = P^{M'_k}(s_i|pa'(S_i), u_1), \tag{25}$$

where $pa'(S_i)$ is an instance of the parent set of S_i on G except U_1.

From these definitions, it follows that

$$P^{M'_1}(n', (s_j, s'_j)) = P^{M_1}(n', s_j, s'_j) = P^{M_2}(n', s_j, s'_j) = P^{M'_1}(n', (s_j, s'_j)), \quad (26)$$

where, n' is an instance of $N\backslash\{S_j\}$ in G.

But for any $(t, s', (s_j, s'_j))$,

$$P_t^{M'_1}(s', (s_j, s'_j)) = P_t^{M_1}(s', s_j, s'_j) \neq P_t^{M_2}(s', s_j, s'_j) P_t^{M'_2}(s', (s_j, s'_j)), \quad (27)$$

where, $s' = s\backslash\{s_j\}$. Therefore, $Q[S]$ is unidentifiable in G.

B) If e is a link of the second type, note that in G', S_j may just has only one unobservable parent, the one on the bidirected link between S_j and S'_j. This happens when S just has one node.
Call U_1 the unobservable node on the bidirected link between S'_j and S_j, and call U_0 the unobservable node that is parent of S'_j and of the T node T_0.
Just as in case A, we can construct new models of G based on models for G'. We define models for G by letting the state space of U_1 be the product of U_0 and U_1 in models for G', and by letting the state space of S_j be the product of the state spaces of S'_j and S_j in models for G' (Fig. 5).
From this point on, the proof of the lemma for case B is analogous to that for case A. □

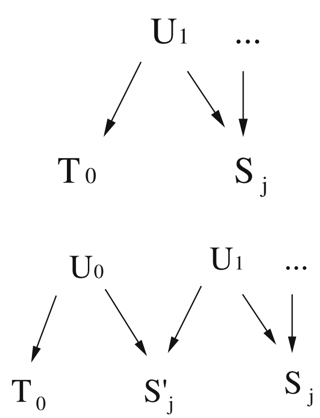
From the lemma above, and noting that this kind of extension will not affect the four graph properties of general unidentifiable subgraphs, the graph G we are studying satisfies also the property below:

Graph Property 5 *Any S node connected with a T node through a directed link or a bidirected link has just two incoming links. One link is connected to a T node, and the other link is a bidirected link connected to another S node.*

In Fig. 6, we present the process of S node extension on graph B of Fig. 3.

From now on, we assume the graph G we study satisfies all these five graph properties.

Fig. 5 S node extension: case B



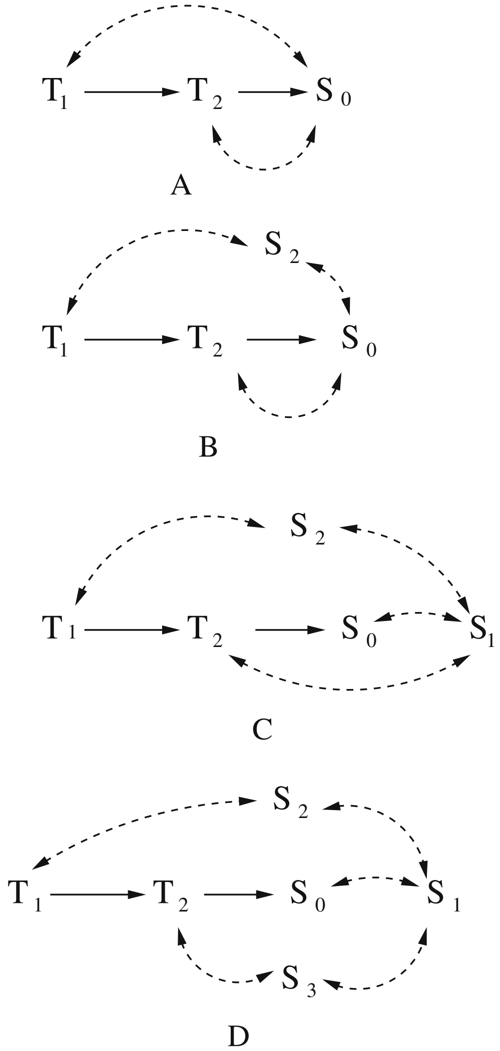


Fig. 6 S node extension example

A.2.1 Mathematical properties

Next, we need some math knowledge.

Math Property 1 *Assume we have a number a,* $0.5 < a < 1$*, then for any c,* $1-a < c < a$*, we can always find a number b,* $0 < b < 1$*, to make that* $ab + (1-a)(1-b) = c$*.*

Proof From $ab + (1-a)(1-b) = c$, we can get $b = (c + a - 1)/(2a - 1)$. Since $c + a - 1 > 0$ and $c + a < 2a$, we have $0 < b < 1$. □

Math Property 2 *For given* $0.5 < m < 1$, $n > 0$*, if we have* $0.5 < m + n < 1$*, then we can find* a, b, c*, such that* $0.5 < a < 1$*,* $0 < b < 1$*,* $0 < c < 1$*,* $c \neq 1 - b$*,* $ab + (1 - a)(1 - b) = m$*, and* $ac + (1 - a)(1 - c) = n$*.*

Proof Assign a value in $(1 - n/2, 1)$ to a. Note that $0.5 < m$ and $m + n < 1$, and therefore $0.5 < a < 1$ and $a > m$. From math property 1, we can find b such that $ab + (1 - a)(1 - b) = m$. Since $1 - a < n/2 < n < 1 - n < a$, using math property 1 again, we can find c such that $ac + (1 - a)(1 - c) = n$.

If we have $c = 1 - b$, then

$$\begin{aligned} m + n &= ab + (1 - a)(1 - b) + ac + (1 - a)(1 - c) \\ &= ab + (1 - a)(1 - b) + a(1 - b) + (1 - a)b \\ &= a + 1 - a = 1 \end{aligned} \tag{28}$$

But this is impossible because $m + n < 1$. □

Math Property 3 *If we have* a, b*, and* c *such that* $0.5 < b < a < 1$*, and* $ca + (1 - c)(1 - a) = b$*, then* $0.5 < c < 1$*. If we have* a, b*, and* c *such that* $0 < a < b < 0.5$*, and* $ca + (1 - c)(1 - a) = b$*, then* $0.5 < c < 1$*.*

Proof For the first part, from $ca + (1 - c)(1 - a) = b$, we have $c = (b + a - 1)/(2a - 1)$. From this, $b + a - 1 > 0$, and $2a - 1 > 0$, we obtain $c > 0$. Also, $b + a - 1 < 2a - 1$, so $c < 1$, and since $(2a - 1)/2 = a - 1/2 < a - 1/2 + b - 1/2 = b + a - 1$, we obtain $c > 0.5$.

For the second part, from $ca + (1 - c)(1 - a) = b$, we have $c = (1 - b - a)/(1 - 2a)$. From this, $1 - b - a > 0$, and $1 - 2a > 0$, we obtain $c > 0$. Also, $1 - b - a < 1 - 2a$, so $c < 1$, and since $(1 - 2a)/2 = 1/2 - a < 1/2 - a + 1/2 - b = 1 - b - a$, we obtain $c > 0.5$. □

Math Property 4 *If we have two numbers* a, b*, with* $0 < a < 0.5$ *and* $0.5 < b < 1$*, then* $ab + (1 - a)(1 - b) < 0.5$*.*

Proof We have

$$\begin{aligned} &0.5 - (ab + (1 - a)(1 - b)) \\ &\quad = 0.5 - (ab + 1 - a - b + ab) \\ &\quad = b - 2ab - 0.5 + a \\ &\quad = b(1 - 2a) - 0.5(1 - 2a) \\ &\quad = (1 - 2a)(b - 0.5) > 0 \end{aligned} \tag{29}$$

□

Math Property 5 *If we have a number* a *such that* $0.5 < a < 1$*, and two numbers* $b, c \in (0, 1)$ *then* $ab + (1 - a)(1 - b) = ac + (1 - a)(1 - c)$ *if and only if* $b = c$

Proof We have

$$\begin{aligned} & ab + (1-a)(1-b) \\ & \quad = ac + (1-a)(1-c) \Longleftrightarrow ab - ac + (1-a)(1-b) - (1-a)(1-c) \\ & \quad = 0 \Longleftrightarrow a(b-c) + (1-a)(b-c) \\ & \quad = 0 \Longleftrightarrow b - c = 0 \Longleftrightarrow b = c \end{aligned} \tag{30}$$

□

Math Property 6 *Assume that we have positive numbers* c, d, $0.5 < c < 1$ *and* $c + d < 1$. *Then, for any number* $n \in [0.5, c)$ *we can always find a number* a, $0 < a < 1$, *such that:* $a \times c + (1-a) \times d = n$

Proof From $a \times c + (1-a) \times d = n$, we get $a = (n-d)/(c-d)$. From $c + d < 1$ and $c > 0.5$, we obtain $d < 0.5$, and therefore $c > d$, and $c - d > 0$. We also have $n - d > 0$ and $n - d < c - d$ when $n \in [0.5, c)$. Therefore $0 < a = (n-d)/(c-d) < 1$. □

A.3 EG Graph and EG^S graph

To prove that all graphs that satisfy the 5 graph properties are unidentifiable, we first show that a class of special graphs, which we call EG graphs, are unidentifiable. Then we extend the result to show that all graphs that satisfy the five graph properties are also unidentifiable.

Let G be a graph that satisfies the five graph properties. We define how to construct EG from G first.

Based on graph property 5, the node set S of G can be divided into three disjunct sets: $S = S^d \cup S^m \cup S^i$. Here S^d contains exactly the S nodes that have a T node as parent. S^i contains exactly the S nodes that have bidirected links with T nodes. $S^m = S \backslash \{S^d \cup S^i\}$ contains exactly the S nodes that have no directed link or bidirected link from any T node.

Note that $|S^i| > 0$, because G is a c-component.

Assume that in graph G, $S^i = \{S^i_1, S^i_2, \ldots, S^i_{n_1}\}$, and these nodes are connected with T nodes $T_1, T_2, \ldots, T_{n_1}$ with bidirected links. Graph $EG(G)$ is obtained by adding $n_1 - 1$ bidirected links between (T_1, T_j), $j = 2, \ldots, n_1$ on G. See Fig. 7 for an example.

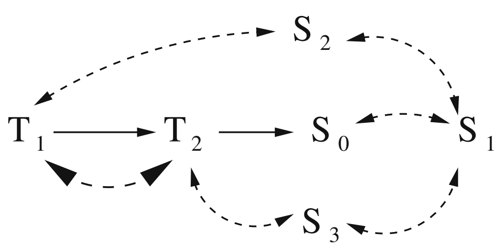


Fig. 7 EG graph for extension result of Fig. 6

So, for any graph G that satisfies the five graph properties given above, we can generate an EG graph $EG(G)$. Any graph that can be constructed in the way just described from a graph G that satisfies the 5 properties is called a EG graph. If in graph G, $|S^i| = 1$, then $EG(G) = G$, and we call any graph that satisfies this property an EG^S graph. We have $EG^S \subset EG$.

Note that for any EG graph G, when we take bidirected links as edges, G_T is a free tree and a c-component. For observable nodes $T_1, T_2 \in T$ in graph G, there is a unique bidirected path from T_1 to T_2 that includes only nodes in T.

Also note that for any EG graph with $|S_i| = n_1$, if we remove $n_1 - 1$ S_i nodes and the bidirected links attached with them, we get an EG^S graph. We will exploit this property in our model construction later (Fig. 8).

A.4 Unidentifiability of EG graphs

A.4.1 Model construction

Assume that, in graph G, $|S^i| = n_1$, $|S^m| = n_2$, $|S^d| = n_3$, and $|T| = n_4$. Based on the graph property 1 and the construction of EG graphs, the number of unobservable nodes (equivalently, bidirected links) in graph $EG(G)$ is $m = n_1 + n_2 + n_3 + n_4 - 1 + n_1 - 1$.

To show that any EG graph is unidentifiable, we create two models M_1 and M_2 and show that they have different causal effects on $P_t(s)$ but the same probabilities on the observable variables.

We define a function $cf(v)$, where v is an instance of vector $v = (v_1, \ldots, v_k)$, as

$$cf(v) = \sum_{i=1}^{k} v_i \tag{31}$$

Our construction for M_1 and M_2 is as below:

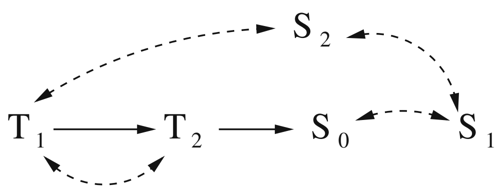
For the models we create, we assume all the variables are binary, with state space (0, 1), and for each unobservable node U_j, $P^{M_i}(u_j = 0) = 1/2$, $i \in \{1, 2\}$, and $j \in \{1, \ldots, m\}$.

We assign a value $0 < \nu_x < 1$ to each observable node $X \in T \cup S^m$, and

$$\begin{cases} P^{M_i}(X = x|pa(X)) = \nu_x & if\ cf((pa(X), x))\ mod\ 2 = 0 \\ P^{M_i}(X = x|pa(X)) = 1 - \nu_x & if\ cf((pa(X), x))\ mod\ 2 = 1, \end{cases} \tag{32}$$

where $(pa(X), x)$ is a vector obtained by adding x at the end of vector $pa(X)$.

Fig. 8 EG^S graph obtained by removing node S_3 in Fig. 7



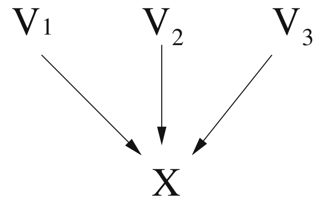


Fig. 9 A node with three parents

Example 1 In Fig. 9, node X has three parents, and the CPT of X is as follows:

X	V_1	V_2	V_3	$P^{M_k}(x\|v_1, v_2, v_3)$
0	0	0	0	ν_x
0	0	0	1	$1 - \nu_x$
0	0	1	0	$1 - \nu_x$
0	0	1	1	ν_x
0	1	0	0	$1 - \nu_x$
0	1	0	1	ν_x
0	1	1	0	ν_x
0	1	1	1	$1 - \nu_x$
1	0	0	0	$1 - \nu_x$
1	0	0	1	ν_x
1	0	1	0	ν_x
1	0	1	1	$1 - \nu_x$
1	1	0	0	ν_x
1	1	0	1	$1 - \nu_x$
1	1	1	0	$1 - \nu_x$
1	1	1	1	ν_x

(33)

We also assign a value $0 < \nu_x < 1$ to each node $X \in S^d$. Note that X has just two parents. Assume $T_x \in T$ is a parent of X, and U_x is the other parent. We define:

X	T_x	U_x	$P^{M_k}(x\|t_x, u_x)$
0	0	0	ν_x
0	0	1	$1 - \nu_x$
0	1	0	$1/2$
0	1	1	$1/2$

(34)

Note that $P^{M_k}(X = 1|t_x, u_x) = 1 - P^{M_k}(X = 0|t_x, u_x)$.

We assign two values $0 < \nu_x^1 < 1$ and $0 < \nu_x^2 < 1$, $\nu_x^1 \neq 1 - \nu_x^2$ with each node $X \in S^i$. Note that X has two unobservable parents. Assume U_1 is the parent on the bidirected link between X and a T node, and U_2 is the other parent, which is on the bidirected link between X and an S node. We define:

X	U_1	U_2	$P^{M_k}(x\|u_1, u_2)$
0	0	0	ν_x^1
0	0	1	$1 - \nu_x^1$
0	1	0	ν_x^2
0	1	1	$1 - \nu_x^2$

(35)

Note that $P^{M_k}(X=1|u_1,u_2) = 1 - P^{M_k}(X=0|u_1,u_2)$.

A.4.2 Construction properties

Here are some properties of this construction.

First, for each node $X \in T \cup S^m$ and for any unobservable node $U' \in Pa(X)$,

$$\sum_{U'} P^{M_i}(X=x|pa(X)) = 1 \tag{36}$$

Second, for each node $X \in S^d$ and for $U_x \in Pa(X)$,

$$\sum_{U_x} P^{M_i}(X=x|t_x,u_x) = 1 \tag{37}$$

Third, for each node $X \in S^i$ and for $U_2 \in Pa(X)$,

$$\sum_{U_2} P^{M_i}(X=x|u_1,u_2) = 1 \tag{38}$$

This means that if we marginalize over the U_2 node, which is the U node on the S side, we obtain 1.

Recall that $n_1+n_2+n_3+n_4$ is the number of observable variables, and therefore $|\, U \,| = 2n_1+n_2+n_3+n_4-2$ in any EG graph.

Lemma 8 *Under the construction above, if we can find parameter values for which*

$$P^{M_k}(T=0,S=0) = \sum_{U} \prod_{v\in T\cup S\cup U} P(v|pa(v)) = (1/2)^{n_1+n_2+n_3+n_4}, \tag{39}$$

then for any (s,t)*, we have* $P^{M_k}(T=t,S=s) = (1/2)^{n_1+n_2+n_3+n_4}$*, and* $P^{M_1}(N) = P^{M_2}(N)$ *is always satisfied.*

Proof Since $P(u) = 1/2$ for all unobservable variables, we just need to show that when $t=0, s=0$,

$$\sum_{U} \prod_{V\in T\cup S} P(v|pa(v)) = 1/2 \times 2^{n_1-1} \tag{40}$$

holds for any (t,s) pair if it holds for $t=0, s=0$.

(a) For a particular set of values $(s,t) = (s_1,\ldots,s_{n_1+n_2+n_3},t_1,\ldots,t_i,\ldots,t_{n_4})$, if T_i is a parent of a S node, and $t_i=1$, then (40) is satisfied.

Assume the S node which is child of T_i is S_i,notes when $t_i=1$, $P^{M_k}(t_i|pa(S_i)) = 1/2$, which is a constant and can be put out. In the remain part, we can always have a U_i, which only appears as one observable node X_j's parent, we can repeatedly remove $P(X_j|Pa(X_j))$ and finally get 1,and n_1-1 extra U nodes we added when we construct EG graph, so (40) is satisfied.

(b) If for a particular set of values $(s,t) = (s_1,\ldots,s_{n_1+n_2+n_3},t_1,\ldots,t_{n_4})$, (40) is satisfied, then for the set of values

$$(s_1,\ldots,s_{i-1},1-s_i,s_{i+1},\ldots,s_{n_1+n_2+n_3},t_1,\ldots,t_{n_4}) \tag{41}$$

Equation (39) is also satisfied, because

$$\sum_U \prod_{V \in T \cup S} P(v|pa(v))((S, T) = (s_1, \ldots, s_{n_1+n_2+n_3}, t_1, \ldots, t_{n_4}))$$

$$+ \sum_U \prod_{V \in T \cup S} P(v|pa(v))((S, T)$$

$$= (s_1, \ldots, s_{i-1}, 1 - s_i, s_{i+1}, \ldots, s_{n_1+n_2+n_3}, t_1, \ldots, t_{n_4}))$$

$$= 2^{n_1 - 1} \tag{42}$$

First for the node S_i, we know $P(S_i = 0|pa(S_i) + P(S_i = 1|pa(S_i) = 1$ for any given $pa(S_i)$. so this $P(s_i|pa(S_i)$ can be removed. Then we can always select U_i, which only appears as one observable node X_j's parent, repeatedly remove $P(X_j|Pa(X_j))$ from (42), and finally obtain 2^{n_1-1}.

(c) If for a particular set of values $(s, t) = (s_1, \ldots, s_{n_1+n_2+n_3}, t_1, \ldots, t_{n_4})$, (39) is satisfied, then for the set of values $(s_1, \ldots, s_{n_1+n_2+n_3}, t_1, \ldots, t_{i-1}, 1 - t_i, t_{i+1}, \ldots, t_{n_4})$ (39) is also satisfied, when T_i is not a parent of any S node.

We prove this by showing

$$\sum_U \prod_{V \in T \cup S} P(v|pa(v))((S, T) = (s_1, \ldots, s_{n_1+n_2+n_3}, t_1, \ldots, t_{n_4}))$$

$$= \sum_U \prod_{V \in T \cup S} P(v|pa(v))$$

$$((S, T) = (s_1, \ldots, s_{n_1+n_2+n_3}, t_1, \ldots, t_{i-1}, 1 - t_i, t_{i+1}, \ldots, t_{n_4})) \tag{43}$$

Since T_i is an ancestor of S, there must be a directed path from T_i to an S node, and T_i must have a child in T. Assume T_j is the observable child of T_i. From the construction of EG, we know that we can find an unique bidirected path from T_i to T_j and that all the observable nodes on that path are T nodes. We name the unobservable variable set on that path $U_{i,j}$.

For the instantiation of $P^{M_k}(s_1, \ldots, s_{n_1+n_2+n_3}, t_1, \ldots, t_{n_4}, u_{i,j}, u/u_{i,j})$, $u_{i,j}$ is an instance of variable set $U_{i,j}$, $u \backslash u_{i,j}$ is an instance of $U \backslash U_{i,j}$, and based on our construction we know that it equals $P^{M_k}(s_1, \ldots, s_{n_1+n_2+n_3}, t_1, \ldots, t_{i-1}, 1 - t_i, t_{i+1}, \ldots, t_{n_4}, u'_{i,j}, u/u_{i,j})$, where $u'_{i,j}$ is given by reversing all the values in $u_{i,j}$.

This is because: for node T_i,

$$P^{M_k}(T_i = t_i|pa'(T_i), u_i) = P^{M_k}(T_i = 1 - t_i|pa'(T_i), 1 - u_i) \tag{44}$$

where u_i is an instance of unobservable node $U_i \in U_{i,j}$, and $pa'(T_i)$ is an instance of $Pa'(T_i) = Pa(t_i) \backslash \{U_i\}$, and for any node X which has two unobservable parents $U_0 \in U_{i,j}$, $U_1 \in U_{i,j}$,

$$P^{M_k}(X = x|pa'(X), u_0, u_1) = P^{M_k}(X = x|pa'(X), 1 - u_0, 1 - u_1), \tag{45}$$

where u_0,u_1 are instances of U_0,U_1, and $pa'(X)$ is an instance of $Pa'(X) = Pa(X) \backslash \{U_0, U_1\}$.

For node T_j,

$$P^{M_k}(T_j = t_j|pa'(T_j), t_i, u_j) = P^{M_k}(T_j = t_j|pa'(T_j), 1 - t_i, 1 - u_j), \tag{46}$$

where u_j is an instance of unobservable node $U_j \in U_{i,j}$, $pa'(T_j)$ is an instance of $Pa'(T_j) = Pa(t_i) \backslash \{T_i, U_j\}$

This equation gives us a one-one map between $P^{M_k}(s, t, u)$ and

$$P^{M_k}(s, t_1, \ldots, t_{i-1}, 1 - t_i, t_{i+1}, \ldots, t_{n_4}, u), \tag{47}$$

so (43) is satisfied. □

Before we determine the values attached with the observable nodes in M_k, $k = 1, 2$, we give a lemma/

Lemma 9 *Let M_k be one of the models we create on an EG graph G. Let between (T_1, T_2) and (T_2, T_3),$T_1, T_2, T_3 \in T$ be bidirected links. Let M'_k, be defined on a graph equal to G but with bidirected link (T_1, T_3) instead of (T_1, T_2). If in both M_k and M'_k, the variables attached all nodes are the same, and in model M_k, (40) is satisfied, then (40) is also satisfied in M'_k, and we have $P^{M_k}(N) = P^{M'_k}(N)$ and $Q^{M_k}[S] = Q^{M'_k}[S]$.*

Proof First, note that we just need to consider the situation in which $P(N = 0)$. From cases a, b, and c in the proof of Lemma 8, we know we just need to consider $\sum_u P(s = 0, t = 0)$ in these two different models. We assume that in the first graph the unobservable node in bidirected link (T_1, T_2) is U_{12}, the unobservable node in bidirected link (T_2, T_3) is U_{23}. In the second graph the unobservable node in bidirected link (T_1, T_3) is U'_{13}, the unobservable node in bidirected link (T_2, T_3) is U'_{23}.

For any instantiation u' of $U \backslash \{U_{12}, U_{23}\}$, we have

$$\begin{aligned}
P^M(S=0, T=0, u', U_{12}=0, U_{23}=0) &= P^{M'}(S=0, T=0, u', U_{13}=0, U_{23}=0)\\
P^M(S=0, T=0, u', U_{12}=0, U_{23}=1) &= P^{M'}(S=0, T=0, u', U_{13}=0, U_{23}=1)\\
P^M(S=0, T=0, u', U_{12}=1, U_{23}=0) &= P^{M'}(S=0, T=0, u', U_{13}=1, U_{23}=1)\\
P^M(S=0, T=0, u', U_{12}=1, U_{23}=1) &= P^{M'}(S=0, T=0, u', U_{13}=1, U_{23}=0)
\end{aligned} \tag{48}$$

So, $\sum_u P^M(S = 0, T = 0) = \sum_u P^{M'}(S = 0, T = 0)$. □

A.4.3 Unidentifiability of EG^S graph

Note that any EG^S graph is also a EG graph and we follow the same model construction we defined above.

Graph $G_{S_d \cup S_m}$ is a subgraph of a EG^S graph. (It is the same when we take it as a subgraph of the graph G, which generates the EG^S graph.). It just includes observable nodes in $S_d \cup S_m$ plus all bidirected links between them. Note that when we treat bidirected links as edges, $G_{S_d \cup S_m}$ is a free tree.

Fig. 10 shows the $G_{S_d \cup S_m}$ graph of EG^S in Fig. 8

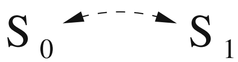


Fig. 10 $G_{S_d \cup S_m}$ for EG^S graph Fig. 7

Lemma 10 *In graph* $G_{S_d \cup S_m}$, $\sum_U \prod_{X \in S_d \cup S_m} P^{M_k}(X = 0|pa(X))$ *can take any value in* (0.5, 1).

Proof From the graph properties, we know that for any EG^S graph G, $G_{S_d \cup S_m}$ is a free tree when we take the bidirected links as edges. We prove this lemma by induction.

First, when there is just one node in $S_d \cup S_m$, $G_{S_d \cup S_m}$ just has that one node, and there are no unobservable nodes. And as we defined before, that observable node is binary. $\sum_U \prod_{X \in S_d \cup S_m} P^{M_k}(X = 0|pa(X)) = P(X = 0) = \nu_x$, which can be any value in (0.5, 1).

The inductive assumption is that when there are k nodes in $G_{S_d \cup S_m}$,

$$a = \sum_U \prod_{X \in S_d \cup S_m} P^{M'_k}(X = 0|pa(X)) \tag{49}$$

can be any value in (0, 1).

Any particular $G_{S_d \cup S_m}$ has $k+1$ nodes can be obtained by adding an observable node S_i in another $G_{S_d \cup S_m}$ with k nodes. We can assume that the added S_i is a leaf in the free tree and assume the unobservable parent of S_i in $G_{S_d \cup S_m}$is U_i.

Based on our construction property, in the new graph with S_i and U_i, and in the old graph plus U_i, $\sum_{U_i=\{0,1\}} \sum_U \prod_{X \in S_d \cup S_m} P(X = 0|pa(X)) = 1$

So, in the new graph, for $S_d = 0$, $S_m = 0$ and $S_i = 0$,

$$\begin{aligned}
&\sum_{U_i} \sum_{U} \prod_{X \in S_d \cup S_m \cup \{S_i\}} P(X = 0|pa(X)) \\
&= \sum_{U_i = 0} \sum_{U} \prod_{X \in S_d \cup S_m \cup \{S_i\}} P(X = 0|pa(X)) \\
&\quad + \sum_{U_i = 1} \sum_{U} \prod_{X \in S_d \cup S_m \cup \{S_i\}} P(X = 0|pa(X)) \\
&= \nu_{S_i} \times a + \sum_{U_i = 1} \sum_{U} \prod_{X \in S_d \cup S_m} P(X = 0|pa(X))(1 - a) \\
&= \nu_{S_i} \times a + \left(\sum_{U_i = 0,1} \sum_{U} \prod_{X \in S_d \cup S_m} P(X = 0|pa(X)) \right. \\
&\quad \left. - \sum_{U_i = 0} \sum_{U} \prod_{X \in S_d \cup S_m} P(X = 0|pa(X)) \right) (1 - a) \\
&= \nu_{S_i} \times a + (1 - \nu_{S_i}) \times (1 - a)
\end{aligned} \tag{50}$$

For any value $b \in (0.5, 1)$, we can set $a = (1 + b)/2$, and therefore $a \in (0.5, 1)$. Based on math property 1, we can now choose $\nu_{S_i} \in (0, 1)$ in such a way that (50) is b. □

Example 2 Consider the graph $G_{S_d \cup S_m}$ shown in Fig. 10. To make

$$\sum_{U} \prod_{X \in S_d \cup S_m} P^{M_k}(X = 0 | pa(X)) \tag{51}$$

equal to 0.8, we can select, for example, $\nu_{S_0} = 0.9$ and $\nu_{S_1} = 7/8 = 0.875$. To make it equal to 0.9, we can select, for example, $\nu_{S_0} = 0.95$ and $\nu_{S_1} = 0.9444444$. To make it equal to 0.95, we can select, for example, $\nu_{S_0} = 0.975$ and $\nu_{S_1} = 0.97368421$.

Next, we study graph $G_{S^d \cup S^m \cup \{S_1^i\}}$. This is the subgraph of an EG^S graph obtained by adding node S_1^i and the bidirected link between it and a S node to graph $G_{S^d \cup S^m}$.

We know that the S_1^i node has two U parents in the EG^S graph, U_2 on the bidirected link to a S node and U_1 on the bidirected link to a T node.

We name G^{ini} the graph obtained by adding node U_1 and the directed link from U_1 to S_1^i to $G_{S^d \cup S^m \cup \{S_1^i\}}$.

Then we have the lemma below:

Lemma 11 *For any value* $0.5 < a < 1$ *and* $0.5 < b < a$*, in any* G^{ini} *graph* G*, we can force*

$$\sum_{U} \prod_{X \in S^d \cup S^m \cup \{S_1^i\}} P^{M_k}(X = 0 | pa(X)) = a \tag{52}$$

and

$$\sum_{U_1 = 0} \sum_{U \setminus \{U_1\}} \prod_{X \in S^d \cup S^m \cup \{S_1^i\}} P^{M_k}(X = 0 | pa(X)) = b \tag{53}$$

Proof Assume in the graph $G_{S^d \cup S^m}$, which is a subgraph for the given G^{int} graph,

$$\sum_{U} \prod_{X \in S_d \cup S_m} P(X = 0 | Pa(X)) = c \tag{54}$$

then in the G^{ini} graph, just like in quantity (50), we have

$$\sum_{U_1 = 0} \sum_{U \setminus \{U_1\}} \prod_{X \in S^d \cup S^m \cup \{S_1^i\}} P(X = 0 | pa(X)) = \nu^1_{S_1^i} c + \left(1 - \nu^1_{S_1^i}\right)(1 - c) \tag{55}$$

and

$$\sum_{U_1 = 1} \sum_{U \setminus \{U_1\}} \prod_{X \in S^d \cup S^m \cup \{S_1^i\}} P(X = 0 | pa(X)) = \nu^2_{S_1^i} c + \left(1 - \nu^2_{S_1^i}\right)(1 - c) \tag{56}$$

We want to find $\nu^1_{S_1^i}$ and $\nu^2_{S_1^i}$, so that quantity (55) is b, quantity (56) is $a - b$, and $\nu^1_{S_1^i} \neq \nu^2_{S_1^i}$. From Lemma 10 we know that c can be any value in (0.5, 1), and based on math property 2 we know that the desired result can always be achieved. □

Example 3 Consider the G^{ini} graph in Fig. 11, and $b = 0.7$, $a = 0.8$. to satisfy (52) and (53), we can set $\nu^1_{S_2} = 0.722222$, $\nu^2_{S_2} = 0.055555556$, $\nu_{S_0} = 0.975$ and $\nu_{S_1} = 0.97368421$.

For $b = 0.6$, $a = 0.9$, we can set $\nu^1_{S_2} = 0.642857$, $\nu^2_{S_2} = 0.2142857$, $\nu_{S_0} = 0.925$ and $\nu_{S_1} = 0.9117647$.

For a G^{ini} graph, we denote

$$\sum_U \prod_{X \in S^d \cup S^m \cup \{S^i_1\}} P^{M_k}(X = 0|pa(X)) \tag{57}$$

as N_0 and

$$\sum_{U_1 = 0} \sum_{U \backslash \{U_1\}} \prod_{X \in S^d \cup S^m \cup \{S^i_1\}} P^{M_k}(X = 0|pa(X)) \tag{58}$$

as M_0. Our discussion above shows that we can set values on S nodes to make N_0, M_0 be any values for which $0.5 < M_0 < N_0 < 1$.

Next, we focus on the EG^S graphs, which form a subset of the EG graphs.

Note that in EG^S graph G, G_T and G_S are both c-components, and these two c-components are bidirectly connected by one and only one bidirected link, which goes through S^i_1. The 5 graphical properties are still satisfied in G.

We define on any EG graph G

$$M(G) = \sum_U \prod_{X \in N(G)} P^{M_k}(x|pa(X))(s = 0, t = 0) \tag{59}$$

and

$$N(G) = \sum_U \prod_{X \in S(G)} P^{M_k}(x|pa(X))(s = 0, t = 0) \tag{60}$$

Lemma 12 *For any EG^S graph G, and any $0.5 < n < 1$, there is a model with $N(G) = n$ and in which $M(G)$ is any value in $[0.5, n)$.*

Proof We prove this lemma by induction. First consider there is just one T node T_1 in G (see Fig. 12). Assume the unobservable parent of T_1 is U_1. For given $0.5 < n < 1$

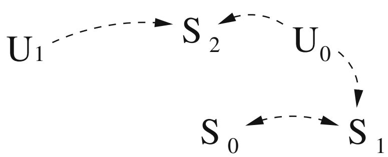


Fig. 11 G^{ini} graph gotten from Fig. 8

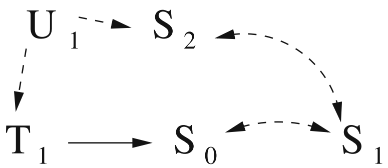


Fig. 12 An EG^S graph with just one T node

and any $0.5 \le m < n$, let $m' = (m+n)/2$. Then from Lemma 11, we can force in the G^{ini} graph obtained from G that $N_0 = n$ and $M_0 = m'$. Note that $N(G) = N_0 = n$ and

$$\begin{aligned}
M(G) &= \sum_{U} \prod_{X \in S \cup T} P^{M_k}(X=0|pa(X)) \\
&= \sum_{U_1=0} \sum_{U \backslash \{U_1\}} \prod_{X \in S \cup T} P^{M_k}(X=0|pa(X)) \\
&\quad + \sum_{U_1=1} \sum_{U \backslash \{U_1\}} \prod_{X \in S \cup T} P^{M_k}(X=0|pa(X)) \\
&= \nu_{T_1} \times \sum_{U_1=0} \sum_{U \backslash \{U_1\}} \prod_{X \in S} P^{M_k}(X=0|pa(X)) \\
&\quad +(1-\nu_{T_1}) \times \sum_{U_1=1} \sum_{U \backslash \{U_1\}} \prod_{X \in S} P^{M_k}(X=0|pa(X)) \\
&= \nu_{T_1} \times M_0 + (1-\nu_{T_1}) \times \left(\sum_{U} \prod_{X \in S} P^{M_k}(X=0|pa(X)) \right. \\
&\quad \left. - \sum_{U_1=0} \sum_{U \backslash \{U_1\}} \prod_{X \in S} P^{M_k}(X=0|pa(X)) \right) \\
&= \nu_{T_1} \times M_0 + (1-\nu_{T_1}) \times (N_0 - M_0)
\end{aligned} \tag{61}$$

Based on math property 2, we know there must be a ν_{T_1} for which $M(G) = 1/2$. And for any positive number in $[0.5, n)$, we can always find a value for ν_{T_1} to make $M(G)$ equal that number. The proof continues with the inductive step after an example.

Example 4 For Fig. 12, if we want to set $N(G) = 0.8$, and $M(G) = 0.6$, we can set: $\nu^1_{S_2} = 0.722222$, $\nu^2_{S_2} = 0.05555556$, $\nu_{S_0} = 0.975$, $\nu_{S_1} = 0.9736841$, and $\nu_{T_1} = 0.83333$.

If we want $N(G) = 0.8$, and $M(G) = 0.5$, we can set: $\nu^1_{S_2} = 0.625$, $\nu^2_{S_2} = 0.125$, $\nu_{S_0} = 0.95$, $\nu_{S_1} = 0.9444444$, and $\nu_{T_1} = 0.75$.

If we want $N(G) = 0.7$, and $M(G) = 0.5$, we can set: $\nu^1_{S_2} = 0.6111111$, $\nu^2_{S_2} = 0.055555556$, $\nu_{S_0} = 0.975$, $\nu_{S_1} = 0.97368421$, and $\nu_{T_1} = 0.8$.

Assume that this lemma is true for each graph EG^S with $|T| = k$. Now consider an EG^S graph G with $|T| = k+1$.

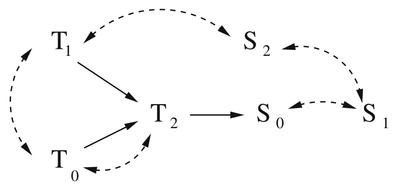


Fig. 13 Bidirected link free tree leaf X has more than one parent

From graph property 1, when we take G_n as a free tree, in EG^S graph G, we can find a T node X, which is a leaf of the free tree.

A) This T node X has no observable parent. Since it is a leaf of the free tree, we know there is only one bidirected link into it. Clearly, that bidirected link connects it with another T node.

From Lemma 9, we can change the bidirected link until it is between X and its child. When X is T_0, Fig. 14 gives an example of this situation. Note when we remove X and the bidirected link attached with it, we will get a EG^S graph with $|T| = k$.

B) This T node X has only one observable parent, as node T_2 in Fig. 8. Note that X has one observable parent and one unobservable parent. Because we just consider the case that all observables are 0, so, for X's only child V, $P(v|Pa(v) \cap N = 0, pa(v) \cap U)$, where $Pa(v) \cap N$ is the observable parents set of V and $Pa(v) \cap U$ is the unobservable parents set of V, is unchanged before and after we add X and the bidirected link attached with it into the original which by itself is an EG^S graph with $|T| = k$.

C) This X node has more than one observable parent, as node T_2 in Fig. 13.

Consider the tree of directed links between observable nodes, and reverse the direction of these links. On this tree, we can find at least two leaves, which are X's observable ancestors. In our example, they are nodes T_0 and T_1. Based on the definition of EG^S node, we know at least one of them has no bidirected link to any S nodes. We take that node as the new X we select, and from Lemma 9 we know that if there are more than one bidirected links into this new X, we can always find an equivalent EG^S graph with just one bidirected link into this new X. Fig. 14 shows the equivalent graph of Fig. 13. Note that we are back in the situation of case A).

In both case A) and case B), consider the graph G' obtained by removing X and the bidirected link attached with it from G. This subgraph G' is still an EG^S graph with $|T| = k$. Based on the inductive assumption, for any given $0.5 < n < 1$, any $0.5 \leq m < n$, and $m' = (m + n)/2$, we can always have $N(G') = n$ and $M(G') = m'$.

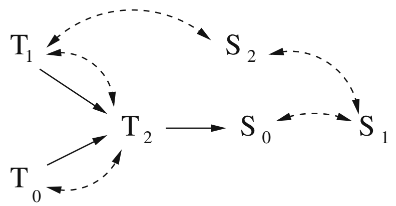


Fig. 14 Graph equivalent to that of Fig. 13

Note that we have $N(G) = N(G')$. Assuming that the observable parent of X is U_1, we have

$$
\begin{aligned}
M(G) &= \sum_{U} \prod_{X \in S \cup T} P^{M_k}(X = 0 | pa(X)) \\
&= \sum_{U_1 = 0} \sum_{U \backslash \{U_1\}} \prod_{X \in S \cup T} P^{M_k}(X = 0 | pa(X)) \\
&\quad + \sum_{U_1 = 1} \sum_{U \backslash \{U_1\}} \prod_{X \in S \cup T} P^{M_k}(X = 0 | pa(X)) \\
&= \nu_x \times M(G') + (1 - \nu_x) \left(\sum_{U} \prod_{X \in S \cup T \backslash \{X\}} P^{M_k}(X = 0 | pa(X)) \right. \\
&\quad \left. - \sum_{U_1 = 0} \sum_{U \backslash \{U_1\}} \prod_{X \in S \cup T \backslash \{X\}} P^{M_k}(X = 0 | pa(X)) \right) \\
&= \nu_x \times M(G') + (1 - \nu_x)(N(G') - M(G')) \\
&= \nu_x \times m' + (1 - \nu_x)(n - m'). \qquad (62)
\end{aligned}
$$

Note that in the above equation we have

$$\sum_{U} \prod_{X \in S \cup T \backslash \{X\}} P^{M_k}(X = 0 | pa(X)) = N(G'), \qquad (63)$$

because if we just insert node U_1 and the link from it to a T node T_1 in G', from (36), we have that U_1 and $P(T_1 = 0 | pa(T_1))$ can be removed from the above equation. Since G_T is a c-component, this kind of removing can continue until all T nodes and U nodes on bidirected links between the T nodes are removed, and we finally get $N(G')$.

Based on math property 6, we can always find a solution ν_x to make $M(G) \in [1/2, m')$. So $M(G)$ can be m, which is in $[0.5, m')$. □

With Lemma 8 and Lemma 12, we have already proved that any EG^S graph G is unidentifiable. We can generate two models M_1 and M_2 following our construction process, and select different N values for them, but force in both models the M value to be $1/2$, which means (39) holds.

Example 5 For Fig. 8, if we want to set $N(G) = 0.8$, and $M(G) = 0.6$, we can set: $\nu_{S_2}^1 = 0.72222222$, $\nu_{S_2}^2 = 0.055555556$, $\nu_{S_0} = 0.975$, $\nu_{S_1} = 0.97368421$, $\nu_{T_1} = 0.9$ and $\nu_{T_2} = 0.91666667$.

If we want to set $N(G) = 0.8$, and $M(G) = 0.5$, we can set: $\nu_{S_2}^1 = 0.72222222$, $\nu_{S_2}^2 = 0.055555556$, $\nu_{S_0} = 0.975$, $\nu_{S_1} = 0.97368421$, $\nu_{T_1} = 0.75$ and $\nu_{T_2} = 0.8333333$.

If we want to set $N(G) = 0.9$, and $M(G) = 0.5$, we can set: $\nu_{S_2}^1 = 0.75$, $\nu_{S_2}^2 = 0.125$, $\nu_{S_0} = 0.95$, $\nu_{S_1} = 0.94444444$, $\nu_{T_1} = 0.6666667$ and $\nu_{T_2} = 0.8$.

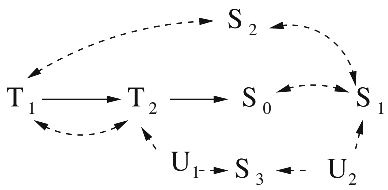


Fig. 15 *EG* graph with two named *U* nodes

Example 6 For Fig. 13, if we want to set $N(G) = 0.8$, and $M(G) = 0.6$, we can set: $\nu^1_{S_2} = 0.7222222$, $\nu^2_{S_2} = 0.05555556$, $\nu_{S_0} = 0.975$, $\nu_{S_1} = 0.97368421$,$\nu_{T_0} = 0.94444444$, $\nu_{T_1} = 0.9285714$ and $\nu_{T_2} = 0.9375$.

If we want to set $N(G) = 0.8$, and $M(G) = 0.5$, we can set: $\nu^1_{S_2} = 0.625$, $\nu^2_{S_2} = 0.125$, $\nu_{S_0} = 0.95$, $\nu_{S_1} = 0.94444444$, $\nu_{T_0} = 0.91666667$, $\nu_{T_1} = 0.875$ and $\nu_{T_2} = 0.9$.

If we want to set $N(G) = 0.9$, and $M(G) = 0.5$, we can set: $\nu^1_{S_2} = 0.75$, $\nu^2_{S_2} = 0.125$, $\nu_{S_0} = 0.95$, $\nu_{S_1} = 0.9444444$, $\nu_{T_0} = 0.8666667$, $\nu_{T_1} = 0.7142857$ and $\nu_{T_2} = 0.818181818$.

Example 7 All setting for Fig. 13 can also be used on Fig. 14.

A.4.3 Unidentifiability of EG graph

In a general EG graph G, assume $|S^i| > 1$, For each node $X \in \{S^i_2, \ldots, S^i_{n_1-1}\}$, there is a bidirected link between X and a T node and a bidirected link between X and a S node. Note that when we remove X and the two bidirected links attached with it, the result is still a EG graph. As we mentioned before, by repeatingly removing all nodes in $\{S^i_2, \ldots, S^i_{n_1-1}\}$, we finally obtain an EG^S graph.

In the example of Fig. 15, if we remove node S_3, U_1 and U_2, we obtain an EG^S graph.

Lemma 13 *In any EG graph G with* $|S^i| = n_1$, *we can find* a, b *such that* $0.5 < b < a < 1$, *and make* $M(G) = b \times 2^{n_1-1}$, $N(G) = a \times 2^{n_1-1}$.

Proof We will prove this lemma by induction.

When $|S^i| = 1$, G is an EG^S graph, and the result follows from Lemma 12.

Assume for all EG graphs with $|S^i| = k$ this lemma is still true, and consider an EG graph with $|S^i| = k + 1$. Note that if we remove an S^i node X and the bidirected links attached to it from G, we obtain an EG graph G' with $|S^i| = k$.

From the inductive assumption, we know that we can have $0.5 < b < a < 1$, $M(G') = b \times 2^{k-1}$, and $N(G') = a \times 2^{k-1}$. Assume the U node on the bidirected link connecting X with a T node is U_1 and the U node on the bidirected link connected X with a S node is U_2, and remember that the CPT we create for X is

X	U_1	U_2	$P^{M_k}(x\|u_1, u_2)$
0	0	0	ν^1_x
0	0	1	$1 - \nu^1_x$
0	1	0	ν^2_x
0	1	1	$1 - \nu^2_x$

(64)

In the graph G,

$$
\begin{aligned}
M(G) = \sum_{U} \prod_{X \in S \cup T} & P^{M_k}(X=0|pa(X)) \\
= \sum_{U_1=0, U_2=0} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T} & P^{M_k}(X=0|pa(X)) \\
+ \sum_{U_1=1, U_2=0} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T} & P^{M_k}(X=0|pa(X)) \\
+ \sum_{U_1=0, U_2=1} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T} & P^{M_k}(X=0|pa(X)) \\
+ \sum_{U_1=1, U_2=1} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T} & P^{M_k}(X=0|pa(X)) \\
= \nu_x^1 \times \sum_{U_1=0, U_2=0} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T \backslash \{X\}} & P^{M_k}(X=0|pa(X)) \\
+ \left(1-\nu_x^1\right) \times \sum_{U_1=1, U_2=0} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T \backslash \{X\}} & P^{M_k}(X=0|pa(X)) \\
+ \nu_x^2 \times \sum_{U_1=0, U_2=1} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T \backslash \{X\}} & P^{M_k}(X=0|pa(X)) \\
+ \left(1-\nu_x^2\right) \times \sum_{U_1=1, U_2=1} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T \backslash \{X\}} & P^{M_k}(X=0|pa(X))
\end{aligned}
\tag{65}
$$

We have

$$
\sum_{U_1=0, U_2=0} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T \backslash \{X\}} P^{M_k}(X=0|pa(X)) = M(G') \tag{66}
$$

and

$$
\begin{aligned}
& \sum_{U_1=1, U_2=0} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T \backslash \{X\}} P^{M_k}(X=0|pa(X)) \\
& = \sum_{U_2=0} \sum_{U \backslash \{U_2\}} \prod_{X \in S \cup T \backslash \{X\}} P^{M_k}(X=0|pa(X)) \\
& \quad - \sum_{U_1=0, U_2=0} \sum_{U \backslash \{U_1, U_2\}} \prod_{X \in S \cup T \backslash \{X\}} P^{M_k}(X=0|pa(X)) \\
& = N(G') - M(G'),
\end{aligned}
\tag{67}
$$

where,

$$\sum_{U_2=0}\sum_{U\backslash\{U_2\}}\prod_{X\in S\cup T\backslash\{X\}} P^{M_k}(X=0|pa(X)) = N(G') \tag{68}$$

This is true because when we marginalize away U_1, based on (36), the CPT of a T node which is a child of U_1 can be removed from the left side of the above equation. Because G_T is a c-component, we can repeat this kind of removing and finally get $N(G')$.

We also have

$$\begin{aligned}
&\sum_{U_1=0,U_2=1}\sum_{U\backslash\{U_1,U_2\}}\prod_{X\in S\cup T\backslash\{X\}} P^{M_k}(X=0|pa(X))\\
&=\sum_{U_1=0}\sum_{U\backslash\{U_1\}}\prod_{X\in S\cup T\backslash\{X\}} P^{M_k}(X=0|pa(X))\\
&\quad-\sum_{U_1=0,U_2=0}\sum_{U\backslash\{U_1,U_2\}}\prod_{X\in S\cup T\backslash\{X\}} P^{M_k}(X=0|pa(X))\\
&=1-M(G')
\end{aligned} \tag{69}$$

We have

$$\sum_{U_1=0}\sum_{U\backslash\{U_1\}}\prod_{X\in S\cup T\backslash\{X\}} P^{M_k}(X=0|pa(X)) = 2^{k-1} \tag{70}$$

This is true because when we marginalize away U_2, based on (36), (37) and (38), the CPT of the S node which is a child of U_2 can be removed from the left side of the above equation first. Since G is a c-component, we can repeat this kind of removal and finally remove all $N(G)$ nodes. Note that in G', the number of unobservable nodes minus the number of observable nodes equals $k-1$ and all the unobservable nodes are binary. So, we finally obtain (70).

We also have

$$\begin{aligned}
&\sum_{U_1=1,U_2=1}\sum_{U\backslash\{U_1,U_2\}}\prod_{X\in S\cup T\backslash\{X\}} P^{M_k}(X=0|pa(X))\\
&=\sum_{U_1=1}\sum_{U\backslash\{U_1\}}\prod_{X\in S\cup T\backslash\{X\}} P^{M_k}(X=0|pa(X))\\
&\quad-\sum_{U_1=1,U_2=0}\sum_{U\backslash\{U_1,U_2\}}\prod_{X\in S\cup T\backslash\{X\}} P^{M_k}(X=0|pa(X))\\
&=2^{k-1}-(N(G')-M(G'))
\end{aligned} \tag{71}$$

This is because, by the argument just given,

$$\sum_{U_1=1}\sum_{U\backslash\{U_1\}}\prod_{X\in S\cup T\backslash\{X\}} P^{M_k}(X=0|pa(X)) = 2^{k-1}. \tag{72}$$

We finally obtain

$$
\begin{aligned}
M(G) &= \nu_x^1 M(G') + \nu_x^2 (N(G') - M(G')) \\
&\quad +(1-\nu_x^1)(2^{k-1} - M(G')) + (1-\nu_x^2)(2^{k-1} - N(G') + M(G')) \\
&= \nu_x^1 \times 2^{k-1} b + \nu_x^2 \times 2^{k-1}(a-b) \\
&\quad +(1-\nu_x^1) \times 2^{k-1}(1-b) + (1-\nu_x^2) \times 2^{k-1}(1-a+b) \\
&= 2^{k-1}(\nu_x^1 b + \nu_x^2(a-b) + (1-\nu_x^1)(1-b) + (1-\nu_x^2)(1-a+b))
\end{aligned} \tag{73}
$$

Note that here, for any given $0 < \alpha < min(0.5 - a + b, (b - 0.5)/2)$. Since $b - \alpha > 0.5$, $1 - b < 0.5 < b - \alpha < b$, based on math property 1, we can find a $0 < \nu_x^1 < 1$ to make $\nu_x^1 b + (1-\nu_x^1)(1-b) = b - \alpha$. Because we also have $a - b < 0.5 - \alpha < 0.5 < 1 - (a - b)$, still based on math property 1, we can also find a ν_x^2 to make $\nu_x^2(a - b) + (1-\nu_x^2)(1-a+b) = 0.5 - \alpha$. From math property 3 we have that $\nu_x^1 > 0.5$ and $\nu_x^2 > 0.5$ here. So, $\nu_x^1 \neq 1 - \nu_x^2$. When we use these ν_x^1 and ν_x^2 in the equation above, we have $M(G) = 2^{k-1}(b + 1/2 - 2\alpha)$. For $0.5 < b - 2\alpha < 1$, let $b' = (b + 1/2 - 2\alpha)/2$. We have $0.5 < b' < 1$ and $M(G) = 2^k b'$.

$$
\begin{aligned}
N(G) &= \sum_{U} \prod_{X \in S} P^{M_k}(X=0|pa(X)) \\
&= \sum_{U_1=0, U_2=0} \sum_{U \setminus \{U_1, U_2\}} \prod_{X \in S} P^{M_k}(X=0|pa(X)) \\
&\quad + \sum_{U_1=0, U_2=1} \sum_{U \setminus \{U_1, U_2\}} \prod_{X \in S} P^{M_k}(X=0|pa(X)) \\
&\quad + \sum_{U_1=1, U_2=0} \sum_{U \setminus \{U_1, U_2\}} \prod_{X \in S} P^{M_k}(X=0|pa(X))
\end{aligned}
$$

$$
\begin{aligned}
&+ \sum_{U_1=1, U_2=1} \sum_{U \setminus \{U_1, U_2\}} \prod_{X \in S} P^{M_k}(X=0|pa(X)) \\
&= \nu_x^1 \times \sum_{U_1=0, U_2=0} \sum_{U \setminus \{U_1, U_2\}} \prod_{X \in S \setminus \{X\}} P^{M_k}(X=0|pa(X)) \\
&+ \left(1-\nu_x^1\right) \times \sum_{U_1=1, U_2=0} \sum_{U \setminus \{U_1, U_2\}} \prod_{X \in S \setminus \{X\}} P^{M_k}(X=0|pa(X)) \\
&+ \nu_x^2 \times \sum_{U_1=0, U_2=1} \sum_{U \setminus \{U_1, U_2\}} \prod_{X \in S \setminus \{X\}} P^{M_k}(X=0|pa(X)) \\
&+ \left(1-\nu_x^2\right) \sum_{U_1=1, U_2=1} \sum_{U \setminus \{U_1, U_2\}} \prod_{X \in S \setminus \{X\}} P^{M_k}(X=0|pa(X)) \\
&= \nu_x^1 N(G') + (1-\nu_x^1)(2^k - N(G')) + \nu_x^2 N(G') + (1-\nu_x^2)(2^k - N(G')) \\
&= 2^{k-1}\left(\left(\nu_x^1 + \nu_x^2\right) a + \left(2 - \nu_x^1 - \nu_x^2\right)(2-a)\right)
\end{aligned} \tag{74}
$$

Here we have

$$\sum_{U_1=0,U_2=1}\sum_{U\setminus\{U_1,U_2\}}\prod_{X\in S\setminus\{X\}}P^{M_k}(X=0|pa(X))$$
$$=\sum_{U_1=0}\sum_{U\setminus\{U_1\}}\prod_{X\in S\setminus\{X\}}P^{M_k}(X=0|pa(X))$$
$$-\sum_{U_1=0,U_2=0}\sum_{U\setminus\{U_1,U_2\}}\prod_{X\in S\setminus\{X\}}P^{M_k}(X=0|pa(X))=$$
$$=2^k-N(G) \tag{75}$$

and

$$\sum_{U_1=1,U_2=1}\sum_{U\setminus\{U_1,U_2\}}\prod_{X\in S\setminus\{X\}}P^{M_k}(X=0|pa(X))$$
$$=\sum_{U_1=1}\sum_{U\setminus\{U_1\}}\prod_{X\in S\setminus\{X\}}P^{M_k}(X=0|pa(X))$$
$$-\sum_{U_1=1,U_2=0}\sum_{U\setminus\{U_1,U_2\}}\prod_{X\in S\setminus\{X\}}P^{M_k}(X=0|pa(X))=$$
$$=2^k-N(G) \tag{76}$$

If we let $(\nu_x^1+\nu_x^2)/2=x$, we have $0.5<x<1$, and

$$N(G)=2^k(xa+(1-x)(2-a)) \tag{77}$$

Note that $0.5<a<1, 0.25<a/2<0.5$. If we let $a'=xa+(1-x)(2-a)$, we have

$$a'=xa+(1-x)(2-a)>xa+(1-x)a=a>0.5 \tag{78}$$

and

$$a'=xa+(1-x)(2-a)=2(x\times a/2+(1-x)(1-a/2)) \tag{79}$$

From math property 4 we have

$$x\times a/2+(1-x)(1-a/2)<0.5 \tag{80}$$

So, finally we have

$$0.5<a'=2(x\times a/2+(1-x)(1-a/2))<2\times 0.5=1 \tag{81}$$

Note that

$$\begin{aligned}
&a' - b' \\
&\quad = (\nu_x^1 + \nu_x^2)/2 \times a + (2 - \nu_x^1 - \nu_x^2)/2 \times (2 - a) \\
&\qquad -(\nu_x^1 b + (1 - \nu_x^1)(1 - b) + \nu_x^2(a - b) + (1 - \nu_x^2)(1 - a + b))/2 = \\
&\quad = 1/2(\nu_x^1 a + \nu_x^2 a + (1 - \nu_x^1)(2 - a) + (1 - \nu_x^2)(2 - a) \\
&\qquad -(\nu_x^1 b + \nu_x^2(a - b) + (1 - \nu_x^1)(1 - b) + (1 - \nu_x^2)(1 - a + b)) = \\
&\quad = 1/2(\nu_x^1(a - b) + \nu_x^2 b + (1 - \nu_x^1)(1 - a + b) + (1 - \nu_x^2)(1 - b)) \\
&\quad > 0
\end{aligned} \tag{82}$$

So we have $0.5 < b' < a' < 1$. □

Example 8 For Fig. 15, if we set $\nu_{S_2}^1 = 0.75$, $\nu_{S_2}^2 = 0.125$, $\nu_{S_0} = 0.95$, $\nu_{S_1} = 0.9444444$, $\nu_{T_1} = 0.9$, $\nu_{T_2} = 0.875.\nu_{S_3}^1 = 0.9$ and $\nu_{S_3}^2 = 0.55$, we have $N(G) = 0.955$, and $M(G) = 0.53$.

We now provide a lemma that, when combined with Lemma 8, shows that any *EG* graph is unidentifiable.

Lemma 14 *For any EG graph G, if $S^i = n_1$ we can create two models such that in both of them $M(G) = 1/2 \times 2^{n_1 - 1}$, but the $N(G)$ are not equal.*

Proof When in G, $|S^i| = 1$, from Lemma 12, this lemma have be proved.

When in G $|S^i| = k + 1$, assume node $X \in S^i$, and graph G' is obtained by removing X and the bidirected links attached to it from G. G' is still an *EG* graph, and from Lemma 13 we know we can have a model satisfying $0.5 < b < a < 1$, $M(G') = b \times 2^{k-1}$, $N(G') = a \times 2^{k-1}$. Here we show we can get two pairs (ν_x^1, ν_x^2) such that make in both of them $M(G) = 1/2 \times 2^k$ but $N(G)$ is not equal. From the proof of Lemma 13, we know

$$M(G) = 2^{k-1}(\nu_x^1 b + \nu_x^2(a - b) + (1 - \nu_x^1)(1 - b) + (1 - \nu_x^2)(1 - a + b)) \tag{83}$$

and

$$N(G) = 2^{k-1}((\nu_x^1 + \nu_x^2)a + (2 - \nu_x^1 - \nu_x^2)(2 - a)) \tag{84}$$

For any given $0 < \alpha < min(0.5 - a + b, b - 0.5)$, $b - \alpha > 0.5, 1 - b < 0.5 < 0.5 + \alpha < b$, based on math property 1, we can find a ν_x^1 to make $\nu_x^1 b + (1 - \nu_x^1)(1 - b) = 0.5 + \alpha$. Because we also have $a - b < 0.5 - \alpha < 0.5 < 1 - (a - b)$, still based on math property 1, we can find a ν_x^2 to make $\nu_x^2(a - b) + (1 - \nu_x^2)(1 - a + b) = 0.5 - \alpha$, then we have $M(G) = 1/2 \times 2^k$. From math property 3 we have property $\nu_x^1 > 0.5$ and $\nu_x^2 > 0.5$ here. So, $\nu_x^1 > \neq 1 - \nu_x^2$.

For different values of α satisfying $0 < \alpha < min(0.5 - a + b, b - 0.5)$, we can select more than one pair of (ν_x^1, ν_x^2) for which $M(G) = 1/2 \times 2^k$. Assume that (c, d) and (c', d') are two of those pairs. We have $c \neq c'$ and $c + d \neq c' + d'$.

First we know for (c, d) and (c', d'), we have

$$cb + d(a - b) + (1 - c)(1 - b) + (1 - d)(1 - a + b) = 1 \tag{85}$$

and

$$c'b + d'(a-b) + (1-c')(1-b) + (1-d')(1-a+b) = 1 \tag{86}$$

If $c + d = c' + d'$, let $c' = c + \gamma, \gamma \neq 0$, then $d' = d - \gamma$, and place them into (86):

$$\begin{aligned}
&(c+\gamma)b + (d-\gamma)(a-b) + (1-(c+\gamma))(1-b) + (1-(d-\gamma))(1-a+b) = 1 \iff \\
&cb + d(a-b) + (1-c)(1-b) + (1-d)(1-a+b) \\
&+\gamma b - \gamma(a-b) - \gamma(1-b) + \gamma(1-a+b) = 1 \iff \\
&\gamma(4b - 2a) = 0 \iff \\
&2b = a \ (Wrong)
\end{aligned} \tag{87}$$

So, we have $c + d \neq c' + d'$,
Note:

$$\begin{aligned}
N(G) &= 2^{k-1}((\nu_x^1 + \nu_x^2)a + (2 - \nu_x^1 - \nu_x^2)(2-a)) \\
&= 2^{k+1}((\nu_x^1 + \nu_x^2)/2 \times a/2 + (1 - (\nu_x^1 + \nu_x^2)/2)(1 - a/2))
\end{aligned} \tag{88}$$

From math property 5, we know with (c, d) and (c', d'), we will get different $N(G)$ values. □

Example 9 For this example, the EG^S graph is obtained by removing node S_2 from the EG graph of Fig. 15. We can set $\nu_{S_3}^1 = 0.75$, $\nu_{S_3}^2 = 0.125$, $\nu_{S_0} = 0.95$, $\nu_{S_1} = 0.9444444$, $\nu_{T_2} = 0.9$ and $\nu_{T_1} = 0.875$.

With these values, if we select $\nu_{S_2}^1 = 0.9$, $\nu_{S_2}^2 = 0.7$, we can obtain model 1 with $M^1(G) = 1/2$, and $N^1(G) = 0.94$.

If we select $\nu_{S_2}^1 = 0.8$, $\nu_{S_2}^2 = 0.65$, we can obtain model 2 with $M^2(G) = 1/2$, and $N^2(G) = 0.955$.

A.5 Unidentifiability of G

So far we have proved with our construction that we can create two models M_1 and M_2 to show any EG graph G' is unidentifiable. We need to show that any graph G that satisfies the five graph properties is unidentifiable. We start by showing the following lemma:

Lemma 15 *Assume EG graph G_0 is obtained by adding bidirected links $\{e_1, \ldots, e_k\}$ to graph G, which satisfies the 5 graph properties. Graphs $\{G_1, \ldots, G_k\}$ are defined as: G_i, $1 \leq i \leq k$, is obtained by removing e_i from G_{i-1}, and $G_k = G$. Then, each G_i, $1 \leq i \leq k$, is unidentifiable.*

Proof We want to show that for any G_i, $1 \leq i \leq k$, we can find two models M^1 and M^2 on G_i, which satisfy:

$$\sum_{U(G_i)} \prod_{N \cup U(G_i)} P^{M^1}(x|pa(X))(s,t) = \sum_{U(G_i)} \prod_{N \cup U(G_i)} P^{M^2}(x|pa(X))(s,t) \tag{89}$$

but for one (s', t')

$$\sum_{U(G_i)} \prod_{T \cup U(G_i)} P^{M^1}(x|pa(X))(s', t') \neq \sum_{U(G_i)} \prod_{T \cup U(G_i)} P^{M^2}(x|pa(X))(s', t') \tag{90}$$

Note that all the G_i models have the same S^i,S^d,S^m and T sets, and the same G_s graph.

For G_0, which is an EG graph, we know that we can have two model M^1 and M^2, such that:

All nodes are binary and, especially, all observable nodes are binary.

For any unobservable node $U_j \in U(G_0)$ we have

$$P^{M^1}(U_j = u_j) = P^{M^2}(U_j = u_j) = \alpha \tag{91}$$

here, α means a constant, which, for G_0, is $1/2$.

For any node $X \in T \cup S^m$, for any unobservable node $U' \in Pa(X)$, we have

$$\sum_{U'} P^{M_1}(X = x|pa(X)) = \sum_{U'} P^{M_2}(X = x|pa(X)) = \alpha \tag{92}$$

For any node $X \in S^d$, for $U_x \in Pa(X)$, we have

$$\sum_{U_x} P^{M_1}(X = x|t_x, u_x) = \sum_{U_x} P^{M_2}(X = x|t_x, u_x = 1) = \alpha \tag{93}$$

For any node $X \in S^i$, we have for $U_2 \in Pa(X)$, one of U_2's child is an S node, and

$$\sum_{U_2} P^{M_1}(X = x|u_1, u_2) = \sum_{U_2} P^{M_2}(X = x|u_1, u_2) = \alpha \tag{94}$$

In all the equations above, α is a constant, although it may be different for different X and marginalized U nodes. In G_0, all the αs are equal to 1, and we have that (89) and (90) are satisfied.

Assume that on graph G_i, all the equations from (89) to (94) above are satisfied. We will now remove each of the edges added to the G graph to obtain the EG graph. Assume that the bidirected $< T_1, T_2 >$ is the extra link we remove from G_i to get graph G_{i+1}. U_0 is the unobservable node on that link in G_i. T_1 is connected with S node S_1 through a bidirected link. T_2 is connected with S node S_k through another bidirected link.

We know G_S is a c-component and a bidirected link free tree. So, there is a unique bidirected path in G_S from S_1 to S_k. By adding to this path the bidirected link from T_1 to S_1 and T_2 to S_k, we have a unique bidirected path from T_1 to T_2, and all observable nodes in this path are S nodes. We name them $S_1, \ldots, S_k$ in order, and the unobservable nodes on this path are named $U_1, \ldots, U_{k+1}$, where U_1 is on the bidirected link between T_1 and S_1, U_2 is on the bidirected link between S_1 and $S_2, \ldots, U_{k+1}$ is on the bidirected link between S_k to T_2.

We construct two models M_1' and M_2' on graph G_{i+1}. The construction is based on models M_1 and M_2 on graph G_i.

In model M'_k, $k = 1, 2$, for all the unobservable variables that are not in $\{U_1, \ldots, U_{k+1}\}$, we define their state space to be the same as the state space in M_k. For each X that belongs to this class,

$$P^{M'_k}(X = x) = P^{M_k}(X = x), k = 1, 2. \tag{95}$$

We rename all the unobservable variables that are in $\{U_1, \ldots, U_{k+1}\}$ as $\{U'_1, \ldots, U'_{k+1}\}$ in M'_k. We define their state space in M'_k as the product of their state space in M_k and the state space of U_0 in M_k, which is (0, 1). For each X that belongs to this class,

$$P^{M'_k}(X = (x, u_0)) = P^{M_k}(X = x) \times P^{M_k}(U_0 = u_0) = P^{M_k}(X = x)/2. \tag{96}$$

The state space of all the observable variables is unchanged, i.e., it is the same as in M_k. Therefore, all the observable variables are still binary variables.

For each observable variable X not in $\{T_1, T_2, S_1, \ldots, S_k\}$, we map an instance $pa(X)$ in model M'_k to an instance $pa'(X)$ in model M_k like this: if Y is an observable node in $Pa(x)$, or Y is an unobservable node but is not in $\{U'_1, \ldots, U'_{k+1}\}$, and $Y = y$ is in $pa(X)$, then $Y = y$ is also in $pa'(X)$. If Y is an unobservable node in $Pa(X)$ and Y is in $\{U'_1, \ldots, U'_{k+1}\}$, we denote the value of Y in $pa(X)$ as $Y = (u^Y, u_0^Y)$, and we have $Y = u^Y$ in $pa'(X)$. We define

$$P^{M'_k}(x|pa(X)) = P^{M_k}(x|pa'(X)). \tag{97}$$

For $X = T_1$, we define

$$P^{M'_k}(x|pa'(X), U'_1 = (u_1, u_0)) = P^{M_k}(x|pa'(X), U_1 = u_1, U_0 = u_0). \tag{98}$$

Here, $pa'(X)$ is an instance of $Pa(T_1)$, except for U_0 and U_1.

For $X = T_2$, we define

$$P^{M'_k}(x|pa'(X), U'_{k+1} = (u_{k+1}, u_0)) = P^{M_k}(x|pa'(X), U_{k+1} = u_{k+1}, U_0 = u_0). \tag{99}$$

Here, $pa'(X)$ is an instance of $Pa(T_2)$, except for U_0 and U_{k+1}.

For observable variable S_i in $\{S_1, \ldots, S_k\}$, we define

$$\begin{aligned} P^{M'_k}(s_i|pa(S_i)) &= P^{M'_k}(s_i|pa'(S_i), U'_i = (u_i, u_0^i), U'_{i+1} = (u_{i+1}, u_0^{i+1})) \\ &= \begin{cases} P^{M_k}(s_i|pa'(S_i), U_i = u_i, U_{i+1} = u_{i+1}) & u_0^i = u_0^{i+1} \\ 1/2 & u_0^i \neq u_0^{i+1} \end{cases} \end{aligned} \tag{100}$$

Here, $pa'(S_i)$ is an instance of $Pa(S_i)$ except for U_i and U_{i+1}.

Note that with this construction (91), (92), (93), and (94) still hold, and for any node S_i in $S_1, \ldots, S_k$, we have, for fixed u_0,

$$\sum_{U'_i = (U_i, u_0)} P^{M_i}(s_i|pa(S_i)) = \alpha \tag{101}$$

$$\sum_{U'_{i+1} = (U_{i+1}, u_0)} P^{M_i}(s_i|pa(S_i)) = \alpha \tag{102}$$

From the two equations above and (91), (92), (93) and (94), we have in both $M'_k, k = 1, 2$

$$
\begin{aligned}
P(s,t) &= \sum_{U} \prod_{U \cup N} P(x|pa(X))(s,t) \\
&= \sum_{U'} \sum_{\{U'_1,\ldots,U'_{k+1}\}} \prod_{U' \cup N} P(x|pa(X)) \prod_{\{U'_1,\ldots,U'_{k+1}\}} P(u)(s,t) \\
&= \sum_{U'} \sum_{U_1} \sum_{U_0^1} \sum_{\{U'_2,\ldots,U'_{k+1}\}} \prod_{U' \cup N} P(x|pa(X)) \prod_{\{U_1,U_0^1,U'_2,\ldots,U'_{k+1}\}} P(u)(s,t) \\
&= \sum_{U'} \sum_{U_1,U_2} \sum_{U_0^1,U_0^2} \sum_{\{U'_3,\ldots,U'_{k+1}\}} \\
&\quad \times \prod_{U' \cup N} P(x|pa(X)) \prod_{\{U_1,U_2,U_0^1,U_0^2,U'_3,\ldots,U'_{k+1}\}} P(u)(s,t) \\
&= \sum_{U'} \sum_{U_1,U_2} \sum_{U_0^1,U_0^2} \sum_{\{U'_3,\ldots,U'_{k+1}\}} \\
&\quad \times \prod_{U' \cup N} P(x|pa(X)) \prod_{\{U_1,U_2,U_0^1,U_0^2,U'_3,\ldots,U'_{k+1}\}} P(u)(s,t,u_0^1 \neq u_0^2) \\
&\quad + \sum_{U'} \sum_{U_1,U_2} \sum_{U_0^1,U_0^2} \sum_{\{U'_3,\ldots,U'_{k+1}\}} \\
&\quad \times \prod_{U' \cup N} P(x|pa(X)) \prod_{\{U_1,U_2,U_0^1,U_0^2,U'_3,\ldots,U'_{k+1}\}} P(u)(s,t,u_0^1 = u_0^2) \\
&= \sum_{U'} \sum_{U_1,U_2} \sum_{U_0^1,U_0^2} \sum_{\{U'_3,\ldots,U'_{k+1}\}} \\
&\quad \times \prod_{U' \cup N} P(x|pa(X)) \prod_{\{U_1,U_2,U_0^1,U_0^2,U'_3,\ldots,U'_{k+1}\}} P(u)(s,t,u_0^1 \neq u_0^2) \\
&\quad + \sum_{U'} \sum_{U_1,U_2,U_3} \sum_{U_0^1,U_0^2,U_0^3} \sum_{\{U'_4,\ldots,U'_{k+1}\}} \prod_{U' \cup N} P(x|pa(X)) \\
&\quad \times \prod_{\{U_1,U_2,U_3,U_0^1,U_0^2,U_0^3,U'_4,\ldots,U'_{k+1}\}} P(u)(s,t,u_0^1 = u_0^2 \neq u_0^3) \\
&\quad + \sum_{U'} \sum_{U_1,U_2,U_3} \sum_{U_0^1,U_0^2,U_0^3} \sum_{\{U'_4,\ldots,U'_{k+1}\}} \prod_{U' \cup N} P(x|pa(X)) \\
&\quad \times \prod_{\{U_1,U_2,U_3,U_0^1,U_0^2,U_0^3,U'_4,\ldots,U'_{k+1}\}} P(u)(s,t,u_0^1 = u_0^2 = u_0^3)
\end{aligned}
$$

$$= \sum_{U'} \sum_{U_1, U_2} \sum_{U_0^1, U_0^2} \sum_{\{U_3', \ldots, U_{k+1}'\}}$$

$$\times \prod_{U' \cup N} P(x|pa(X)) \prod_{\{U_1, U_2, U_0^1, U_0^2, U_3', \ldots, U_{k+1}'\}} P(u) \left(s, t, u_0^1 \neq u_0^2\right)$$

$$+ \cdots + \sum_{U'} \sum_{U_1, \ldots, U_{k+1}} \sum_{U_0^1, \ldots, U_0^{k+1}} \prod_{U' \cup N} P(x|pa(X))$$

$$\times \prod_{\{U_1, \ldots, U_{k+1}\}} P(u)(s, t, u_0^1 = \ldots = u_0^{k+1}), \tag{103}$$

where, $U' = U \backslash \{U_1, \ldots, U_{k+1}\}$. Note that, in the last expression of the above equation, all the terms except the last one are equal to a constant, so they are equal in M_1' and M_2'. Based on (89), we know the last term is also equal in both models. So, we have $P^{M_1'}(s, t) = P^{M_2'}(s, t)$ for any (s, t). That is (89) still holds in models M_1' and M_2', and in both M_k', $k = 1, 2$, for $(S = 0, T = 0)$,

$$P_t(s) = \sum_{U} \prod_{U \cup S} P(x|pa(X))(s, t)$$

$$= \sum_{U'} \sum_{\{U_1', \ldots, U_{k+1}'\}} \prod_{U' \cup S} P(x|pa(X)) \prod_{\{U_1', \ldots, U_{k+1}'\}} P(u)(s, t)$$

$$= \sum_{U'} \sum_{U_1} \sum_{U_0^1} \sum_{\{U_2', \ldots, U_{k+1}'\}} \prod_{U' \cup S} P(x|pa(X)) \prod_{\{U_1, U_0^1, U_2', \ldots, U_{k+1}'\}} P(u)(s, t)$$

$$= \sum_{U'} \sum_{U_1, U_2} \sum_{U_0^1, U_0^2} \sum_{\{U_3', \ldots, U_{k+1}'\}}$$

$$\times \prod_{U' \cup S} P(x|pa(X)) \prod_{\{U_1, U_2, U_0^1, U_0^2, U_3', \ldots, U_{k+1}'\}} P(u)(s, t)$$

$$= \sum_{U'} \sum_{U_1, U_2} \sum_{U_0^1, U_0^2} \sum_{\{U_3', \ldots, U_{k+1}'\}}$$

$$\times \prod_{U' \cup S} P(x|pa(X)) \prod_{\{U_1, U_2, U_0^1, U_0^2, U_3', \ldots, U_{k+1}'\}} P(u) \left(s, t, u_0^1 \neq u_0^2\right)$$

$$+ \sum_{U'} \sum_{U_1, U_2} \sum_{U_0^1, U_0^2} \sum_{\{U_3', \ldots, U_{k+1}'\}}$$

$$\times \prod_{U' \cup S} P(x|pa(X)) \prod_{\{U_1, U_2, U_0^1, U_0^2, U_3', \ldots, U_{k+1}'\}} P(u) \left(s, t, u_0^1 = u_0^2\right)$$

$$= \sum_{U'} \sum_{U_1, U_2} \sum_{U_0^1, U_0^2} \sum_{\{U_3', \ldots, U_{k+1}'\}}$$

$$\times \prod_{U' \cup S} P(x|pa(X)) \prod_{\{U_1, U_2, U_0^1, U_0^2, U_3', \ldots, U_{k+1}'\}} P(u) \left(s, t, u_0^1 \neq u_0^2\right)$$

$$+ \sum_{U'} \sum_{U_1, U_2, U_3} \sum_{U_0^1, U_0^2, U_0^3} \sum_{\{U_4', \ldots, U_{k+1}'\}} \prod_{U' \cup S} P(x|pa(X))$$

$$\times \prod_{\{U_1, U_2, U_3, U_0^1, U_0^2, U_0^3, U_4', \ldots, U_{k+1}'\}} P(u)(s, t, u_0^1 = u_0^2 \neq u_0^3)$$

$$+ \sum_{U'} \sum_{U_1, U_2, U_3} \sum_{U_0^1, U_0^2, U_0^3} \sum_{\{U_4', \ldots, U_{k+1}'\}} \prod_{U' \cup S} P(x|pa(X))$$

$$\times \prod_{\{U_1, U_2, U_3, U_0^1, U_0^2, U_0^3, U_4', \ldots, U_{k+1}'\}} P(u)(s, t, u_0^1 = u_0^2 = u_0^3)$$

$$= \sum_{U'} \sum_{U_1, U_2} \sum_{U_0^1, U_0^2} \sum_{\{U_3', \ldots, U_{k+1}'\}}$$

$$\times \prod_{U' \cup S} P(x|pa(X)) \prod_{\{U_1, U_2, U_0^1, U_0^2, U_3', \ldots, U_{k+1}'\}} P(u)(s, t, u_0^1 \neq u_0^2)$$

$$+ \cdots + \sum_{U'} \sum_{U_1, \ldots, U_{k+1}} \sum_{U_0^1, \ldots, U_0^{k+1}} \prod_{U' \cup S} P(x|pa(X))$$

$$\times \prod_{\{U_1, \ldots, U_{k+1}\}} P(u)(s, t, u_0^1 = \ldots = u_0^{k+1}) \tag{104}$$

Note that in the last expression of the above equation, all the terms except the last one are equal to a constant, so they are equal in M_1' and M_2'. Based on (90), we know that the last term is not equal in M_1' and M_2'. So, we have $P_t^{M_1'}(s)(S = 0, T = 0) \neq P_t^{M_2'}(s)(S = 0, T = 0)$, which means (90) still holds for M_k', $k = 1, 2$.

By repeating this construction, we conclude that all graph models $\{G_0, G_1, \ldots, G_k\}$ are unidentifiable. □

Example 10 The Hugin files for the models constructed in this appendix for the G graph of Fig. 6 (D) can be downloaded from http://www.cse.sc.edu/~mgv/reports/exampleD1TS.net and http://www.cse.sc.edu/~mgv/reports/exampleD2TS.net.

Appendix B: Proof of Lemma 5

Lemma 5 *Assume $S \subset N$ and $T \subset N$ are disjunct node sets in graph G, $< X_1, X_2 >$ is a directed link in G, $X_1 \in S$, and $X_2 \in S$. Assume that graph G' is obtained by* *removing link* $< X_1, X_2 >$ *from graph G. If* $P_T(S)$ *is unidentifiable in graph* G'*, then* $P_T(S\backslash\{X_1\})$ *is unidentifiable in G.*

Proof By definition, in graph G', $P_t(s)$ is given by:

$$P_t(s) = \sum_{V_l \in (N\backslash S)\backslash T} \sum_{U_k \in U} \prod_{V_i \in N\backslash T} P(v_i|pa(V_i)) \prod_{V_j \in U} P(v_j) \tag{105}$$

In graph G,

$$P_t(s\backslash\{x_1\}) = \sum_{V_l \in (N\backslash S)\backslash T \cup \{X_1\}} \sum_{U_k \in U} \prod_{V_i \in N\backslash T} P(v_i|pa(V_i)) \prod_{V_j \in U} P(v_j) \tag{106}$$

When $P_T(S)$ is unidentifiable in graph G', we know there are two models M_1 and M_2 on G' such that: $P^{M_1}(n) = P^{M_2}(n)$, which means:

$$\sum_{U_k \in U} \prod_{V_i \in N} P^{M_1}(v_i|pa(V_i)) \prod_{V_j \in U} P^{M_1}(v_j) = \sum_{U_k \in U} \prod_{V_i \in N} P^{M_2}(v_i|pa(V_i)) \prod_{V_j \in U} P^{M_2}(v_j) \tag{107}$$

but for at least one (s, t) ,$P_t^{M_1}(s) \neq P_t^{M_2}(s)$.

Now based on M_1 and M_2, we create models M'_1 and M'_2 on graph G. First, we define a probability function F from $S(X_1)$ to $(0, 1)$, where $S(X_1)$ is the state space of X_1 in model M_i, $i = 1, 2$. For any $a \in S(X_1)$, $P(F(a) = 0) > 0$, $P(F(a) = 1) > 0$ and $P(F(a) = 0) + P(F(a) = 1) = 1$.

For any node X, which is an unobservable node or a node in $N\backslash(\{X_2\} \cup CH(X_2))$, we define, for $i = 1, 2$

$$P^{M'_i}(x|pa(X)) = P^{M_i}(x|pa(X)) \tag{108}$$

The state space of X_2 in M'_i is defined as $S(X_2) \times \{0, 1\}$, where $S(X_2)$ is the state space of X_2 in M_i,$i = 1, 2$.

For $x_2 \in S(X_2)$, $i = 1, 2$, $k = 0, 1$, we define

$$P^{M'_i}((x_2, k)|pa(X_2), x_1) = P^{M_i}(x_2|pa(X_2)) \times P(F(x_1) = k), \tag{109}$$

where $Pa(X_2)$ is the parent set of X_2 in graph G'. So, $Pa(X_2) \cup \{X_1\}$ is the parent set of X_2 in graph G.

Note that, for a given $(pa(X_2), x_1)$, we have

$$\sum_{x_2, k} P^{M'_i}((x_2, k)|pa(X_2), x_1) = \sum_{x_2} P^{M_i}(x_2|pa(X_2)) \times \sum_k P(F(x_1) = k) = 1. \tag{110}$$

For any node $X \in Ch(X_2)$, we define

$$P^{M_i'}(x|pa'(X),(x_2,k)) = P^{M_i}(x|pa'(X),x_2), \tag{111}$$

where $Pa'(X) = Pa(X)\backslash\{X_2\}$ is the parent set of X in graph G, except for node X_2. Then for any instance n of N in model M_1' and M_2', if $X_1 = x_1$ and $X_2 = (x_2, k)$ in n, we have

$$\begin{aligned}
P^{M_1'}(n) &= \sum_{U_k \in U} \prod_{V_i \in N} P^{M_1'}(v_i|pa(V_i)) \prod_{V_j \in U} P^{M_1'}(v_j) \\
&= \sum_{U_k \in U} \prod_{V_i \in N} P^{M_1}(v_i|pa(V_i)) \prod_{V_j \in U} P^{M_1}(v_j)(n) \times P(F(x_1)=k) \\
&= \sum_{U_k \in U} \prod_{V_i \in N} P^{M_2}(v_i|pa(V_i)) \prod_{V_j \in U} P^{M_2}(v_j)(n) \times P(F(x_1)=k) \\
&= \sum_{U_k \in U} \prod_{V_i \in N} P^{M_2'}(v_i|pa(V_i)) \prod_{V_j \in U} P^{M_2'}(v_j) \\
&= P^{M_2'}(n).
\end{aligned} \tag{112}$$

We know that for a given (s,t),$P_t^{M_1}(s) \neq P_t^{M_2}(s)$, and we assume that $X_1 = x_1$ and $X_2 = x_2$ in s.

Note that $\sum_{X_1} P_t^{M_i}(s\backslash\{x_1\}) \leq 1$, because after setting the values of the T nodes, the result model is still a Bayesian network.

Assume that $P_t^{M_1}(s) = a > P_t^{M_2}(s) = b > 0$. If we define $P(F(x_1) = 0) = 0.5$, but $P(F(x) = 0) = (a-b)/4$ for all $x \in S(X_1)$,$x \neq x_1$, we have that for $(s\backslash\{x_2\},(x_2,0),t)$

$$\begin{aligned}
P_t^{M_1'}(s\backslash\{x_1\})(S\backslash\{X_1\} &= (s\backslash\{x_1,x_2\},(x_2,0)),\, T=t) \\
&= \sum_{V_l \in (N\backslash S)\backslash T \cup \{X_1\}} \sum_{U_k \in U} \prod_{V_i \in V\backslash T} P^{M_1'}(v_i|pa(V_i)) \\
(S\backslash\{X_1\} &= (s\backslash\{x_1,x_2\},(x_2,0)),\, T=t) \\
&> \sum_{X_1 = x_1} \sum_{V_l \in (N\backslash S)\backslash T} \sum_{U_k \in U} \prod_{V_i \in V\backslash T} P^{M_1'}(v_i|pa(V_i)) \\
&\quad (S\backslash\{X_1\} = (s\backslash\{x_1,x_2\},(x_2,0)),\, T=t) \\
&= \sum_{V_l \in (N\backslash S)\backslash T} \sum_{U_k \in U} \prod_{V_i \in V\backslash T} P^{M_1}(v_i|pa(V_i))(S=s,\, T=t) \\
&\quad \times P(F(x_1)=0) = 0.5a,
\end{aligned} \tag{113}$$

but

$$
\begin{aligned}
P_t^{M_2'}(s\backslash\{x_1\})(S\backslash\{X_1\} &= (s\backslash\{x_1,x_2\},(x_2,0)),T=t) \\
&= \sum_{V_l \in (N\backslash S)\backslash T\cup\{X_1\}} \sum_{U_k \in U} \prod_{V_i \in V\backslash T} P^{M_2'}(v_i|pa(V_i)) \\
&\quad (S\backslash\{X_1\} = (s\backslash\{x_1,x_2\},(x_2,0)),T=t) \\
&= \sum_{X_1=x_1} \sum_{V_l \in (N\backslash S)\backslash T} \sum_{U_k \in U} \prod_{V_i \in V\backslash T} P^{M_2'}(v_i|pa(V_i)) \\
&\quad (S\backslash\{X_1\} = (s\backslash\{x_1,x_2\},(x_2,0)),T=t) \\
&\quad + \sum_{X_1\neq x_1} \sum_{V_l \in (N\backslash S)\backslash T} \sum_{U_k \in U} \prod_{V_i \in V\backslash T} P^{M_2'}(v_i|pa(V_i)) \\
&\quad (S\backslash\{X_1\} = (s\backslash\{x_1,x_2\},(x_2,0)),T=t) \\
&< \sum_{V_l \in (N\backslash S)\backslash T} \sum_{U_k \in U} \prod_{V_i \in V\backslash T} P^{M_2}(v_i|pa(V_i))(S=s,T=t) \\
&\quad \times P(F(x_1)=0) \\
&\quad + \sum_{V_l \in (N\backslash S)\backslash T\cup\{X_1\}} \sum_{U_k \in U} \prod_{V_i \in V\backslash T} P^{M_2}(v_i|pa(V_i)) \\
&\quad (S\backslash\{X_1\} = s\backslash\{x_1\},T=t)\times P(F(X_i\neq x_1)=0) \\
&\leq 0.5b + \sum_{X_1} P_t^{M_2}(s\backslash\{x_1\})\times(a-b)/4 \\
&\leq 0.5b+(a-b)/4<0.5a
\end{aligned}
\tag{114}
$$

From the models M_1' and M_2' thus constructed, we know $P_T(S\backslash\{X_1\}$ is unidentifiable in G. □

References

1. Tian, J., Pearl, J.: A general identification condition for causal effects. In: Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI-02), pp. 567–573, Edmonton, 28 July–1 August 2002
2. Tian, J., Pearl, J.: On the testable implications of causal models with hidden variables. In: Proceedings of the Eighteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-02), pp. 519–527, Edmonton, July 2002
3. Tian, J., Pearl, J.: On the identification of causal effects, technical report 290-L. Tech. rep., Cognitive Systems Laboratory, University of California at Los Angeles (2003) (Extended version available at http://www.cs.iastate.edu/jtian/r290-L.pdf)
4. Pearl, J.: Causality: Models, Reasoning, and Inference. Cambridge University Press, New York (2000)
5. Pearl, J.: Causal diagrams for empirical research. Biometrika **82**, 669–710, 12 (1995)

6. Galles, D., Pearl, J.: Testing identifiability of causal effects. In: Proceedings of the Eleventh Annual Conference on Uncertainty in Artificial Intelligence (UAI-95), pp. 185–195, Montreal, 18–20 August 1995
7. Pearl, J., Robins, J.M.: Probabilistic evaluation of sequential plans from causal models with hidden variables. In: Proceedings of the Eleveth Annual Conference on Uncertainty in Artificial Intelligence (UAI-95), pp. 444–453, Montreal, 18–20 August 1995
8. Kuroki, M., Miyakawa, M.: Identifiability criteria for causal effects of joint interventions. J. Jpn. Stat. Soc. **29**(2), 105–117 (1999)
9. Robins, J.: Causal inference from complex longitudinal data. In: Berkane, M. (ed.) Latent Variable Modeling with Applications to Causality, Lecture Notes in Statistics, vol. 120, pp. 69–117. Springer, New York (1997)
10. Tian, J.: Identifying conditional causal effects. In: Proceedings of the Twentieth Annual Conference on Uncertainty in Artificial Intelligence (UAI-04), pp. 561–568, Banff, July 2004
11. Lauritzen, S.: Causal inference from graphical models. In: Barndorff-Nielsen, O., Klueppelberg, C. (eds.) Complex Stochastic Systems, pp. 63–107. Chapman and Hall/CRC, London (2001)
12. Verma, T.: Graphical aspects of causal models, technical report, R-191. Tech. Rep., Cognitive Systems Laboratory, University of California at Los Angeles (1993)