Appendix for Causal Discovery from Subsampled Time Series with Proxy Variables

Table of contents

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

A Preliminary

A.1 Explanation of Figure 2(c): MAG

Table 1: Detailed explanation of edges in the MAG

Edge in the MAG	Inducing path in the full time DAG
$A(t_1) \to A(t_3)$	$A(t_1) \to A(t_2) \to A(t_3)$
$A(t_1) \to M(t_3)$	$A(t_1) \to A(t_2) \to M(t_3)$ $A(t_1) \to M(t_2) \to M(t_3)$
$A(t_1) \to B(t_3)$	$A(t_1) \to M(t_2) \to B(t_3)$
$M(t_1) \to M(t_3)$	$M(t_1) \to M(t_2) \to M(t_3)$
$M(t_1) \to B(t_3)$	$M(t_1) \to M(t_2) \to B(t_3)$ $M(t_1) \to B(t_2) \to B(t_3)$
$B(t_1) \to B(t_3)$	$B(t_1) \to B(t_2) \to B(t_3)$
$A(t_3) \leftrightarrow M(t_3)$	$A(t_3) \leftarrow A(t_2) \to M(t_3)$
$M(t_3) \leftrightarrow B(t_3)$	$M(t_3) \leftarrow M(t_2) \to B(t_3)$

A.2 Proof of Remark 2.5: Edge orientations in the MAG

Remark 2.5. *In the MAG, the directed edge $A \to B$ means A is the ancestor of B. The bidirected edge $A \leftrightarrow B$ means there is an unobserved confounder U between A and B.*

Proof. Since the meaning of the directed edge is directly induced from the definition, we focus on explaining the bidirected edge.

Specifically, $A \leftrightarrow B$ means there is an inducing path p relative to $\mathbf{L}$ between A and B. Since $A \notin \mathbf{An}(B)$ and $B \notin \mathbf{An}(A)$, the inducing path p must contains non-mediators, *i.e.*, colliders or confounders. Suppose that there are r confounders on p, since between every two confounders there is a collider, the number of colliders on p is $r-1$. Denote the confounders on p as $U_1, ..., U_r$, the colliders as $C_1, ..., C_{r-1}$, we have $p = A \leftarrow \cdots \leftarrow U_1 \to \cdots \to C_1 \leftarrow \cdots \leftarrow \cdots \to \cdots \to C_{h-1} \leftarrow \cdots \leftarrow U_h \to \cdots \to B$. According to the definition of inducing path, each collider C_i on p is the ancestor of either A or B. Based on this, the following algorithm is assured to find a latent confounder between A and B.

Algorithm 1: Find the latent confounder

for $i = 1, ..., r-1$ **do**
    if $C_i \in \mathbf{An}(B)$ **then**
        return U_i ; /* $A \leftarrow \cdots \leftarrow C_{i-1} \leftarrow \cdots \leftarrow U_i \to \cdots \to C_i \to \cdots \to B$ */
    end
end
return U_r ; /* $A \leftarrow \cdots \leftarrow C_{r-1} \leftarrow \cdots \leftarrow U_r \to \cdots \to B$ */

$\square$

B Structural identifiability

B.1 Proof of Proposition 3.1: Identifiability of the MAG

Proposition 3.1. *Assuming model (1) and Assumption 2.2, then the MAG over the observed variable set* $\mathbf{O}$ *is identifiable, i.e., its skeleton and edge orientations can be uniquely derived from the joint distribution* $\mathbb{P}(\mathbf{O})$.

To prove Proposition 3.1, we first introduce the following lemma, which studies the property of the path between a vertex at t and a vertex at $t+\tau$ ($\tau \geq 0$).

Lemma B.1. *Suppose that there is a path between* $A(t), B(t+\tau)$ ($\tau \geq 0$) *denoted as* $p = A(t), V_1, ..., V_r, B(t+\tau)$, *if* $\exists\, V_i$ *on* p *such that* $time(V_i) \geq t+\tau$, *then there is a collider on* p *at* $time \geq t+\tau$.

Proof. We discuss two possible cases and show that p always contains a collider at $time \geq t+\tau$.

1. $\forall\, V_i, time(V_i) \leq t+\tau$. In this case, there is a vertex V_j on p such that $time(V_j) = t+\tau$. Consider its adjacent vertices V_{j-1}, V_{j+1} on the path. According to our assumption, we have $time(V_{j-1}) = time(V_{j+1}) = t+\tau-1$. Hence, according to the causal precedence assumption, the edges among V_{j-1}, V_j, V_{j+1} are $V_{j-1}(t+\tau-1) \to V_j(t+\tau) \leftarrow V_{j+1}(t+\tau+1)$, which means $V_j(t+\tau)$ is a collider on p at $time = t+\tau$.
2. $\exists\, V_i, time(V_i) > t+\tau$. Let $j := \arg\max_i time(V_i)$ and $t^* := time(V_j)$. We have $t^* > t+\tau$. Consider the vertices V_{j-1}, V_j, V_{j+1}, the edges among them are $V_{j-1}(t^*-1) \to V_j(t^*) \leftarrow V_{j+1}(t^*-1)$. Hence, $V_j(t^*)$ is a collider on p at $time > t+\tau$.

□

Equipped with Lemma B.1, we now introduce the proof of Proposition 3.1below:

Proof of Proposition 3.1. According to [1], the skeleton of the MAG is identifiable under Assumption 2.2.

The orientation of edges in the MAG is either directed (→) or bi-directed (↔). Since we assume that the cause must precedes the effect, instantaneous edges in the MAG must be bi-directed. In the following, we show that lagged edges in the MAG must be directed and cross k time steps, *e.g.*, $A(t) \to B(t+k)$.

Suppose that $A(t_1)$ and $B(t_2)$ ($t_2 > t_1$) are adjacent in the MAG. We first prove that $t_2 = t_1 + k$. Prove by contradiction. Since $A(t_1)$ and $B(t_2)$ are adjacent in the MAG, there is an inducing path $p = A(t_1), V_1, ..., V_r, B(t_2)$ between them in the full time DAG. Suppose that $t_2 \neq t_1 + k$, in other words, $t_2 \geq t_1 + 2k$. Since the order of the SVAR process is 1, there exist V_i on p such that $time(V_i) = t_1 + k, time(V_{i+1}) = t_1 + k + 1$. Hence, $V_i(t_1+k)$ is neither latent nor a collider[1]. As a result, the path p is not an inducing path, which is a contradiction.

Next, we prove that the edge between $A(t), B(t+k)$ is directed. We prove this by showing that $A(t) \leftrightarrow B(t+k)$ can not be true. Prove by contradiction. Suppose that we have $A(t) \leftrightarrow B(t+k)$ in the MAG, then there is an inducing path $p = A(t), V_1, ..., V_r, B(t+k)$ between $A(t), B(t+k)$, and $A(t) \notin \mathbf{An}(B(t+k))$[2] in the full time DAG. There are three kinds of possible inducing paths:

1. $\exists\, V_i$ such that $time(V_i) \geq t+k$;
2. $\forall\, V_i$, $time(V_i) < t+k$ and $time(V_i) > t$;
3. $\forall\, V_i$, $time(V_i) < t+k$ and $\exists\, V_j$ such that $time(V_j) \leq t$.

In the following, we will show neither of these paths exist, which is a contradiction and shows that the edge between $A(t), B(t+k)$ is directed.

[1] $V_i(t_1+k) \to V_{i+1}(t_1+k+1)$ means $V_i(t_1+k)$ can not be a collider.
[2] $B(t+k)$ can not be the ancestor of $A(t)$ according to the causal precedence assumption.

1. $\exists\, V_i$ such that $time(V_i) \geq t+k$. In this case, according to Lemma B.1, there is a collider on p at $time \geq t+k$. Since the assumed causal precedence, the collider can not be the ancestor of $A(t)$ or $B(t+k+1)$. Hence, the path p is not an inducing path.

2. $\forall\, V_i$, $time(V_i) < t+k$ and $time(V_i) > t$. In other words, we have $time(V_i) = t+1, .., t+k-1$. We show that in this case, we have $A(t) \in \mathbf{An}(B(t+k))$ in the full time DAG, which contradicts with $A(t) \leftrightarrow B(t+k)$ in the MAG.

   Specifically, when $k=2$, the path is $p = A(t) \rightarrow V_1(t+1) \rightarrow B(t+2)$. Hence, $A(t) \in \mathbf{An}(B(t+k))$. When $k \geq 3$, consider the first $\leftarrow$ edge along the path[3], we have $A(t) \rightarrow \cdots \rightarrow V_{i-1}(t_i+1) \leftarrow V_i(t_i)$ and $i \geq 3$. Therefore, $V_{i-1}(t_i+1)$ is the descendent of $A(t)$ and a collider on p. According to the definition of the inducing path, $V_{i-1}(t_i+1)$ is the ancestor of $B(t+k)$, which means $A(t) \in \mathbf{An}(B(t+k))$.

3. $\forall\, V_i$, $time(V_i) < t+k$ and $\exists\, V_j$ such that $time(V_j) \leq t$. In this case, we further consider two possibilities,

   a. $\forall\, V_i, t \leq time(V_i) < t+k$. In this case, there is a vertex V_j on p such that $time(V_j) = t$ and we have $V_{j-1}(t+1) \leftarrow V_j(t) \rightarrow V_{j+1}(t+1)$. Therefore, $V_j(t)$ is neither a latent variable nor a collider, which means p is not an inducing path.

   b. $\forall\, V_i, time(V_i) < t+k$ and $\exists\, V_i$ such that $time(V_i) < t$. In this case, we have $p = A(t), ..., V_i(t-\tau), ..., B(t+k)$ $(\tau > 0)$. As a result,there is a vertex V_j on p between $V_i(t-\tau)$ and $B(t+k)$ such that $time(V_j) = t$ and $time(V_{j+1} = t+1$. Therefore, $V_j(t)$ is neither a latent variable nor a collider, which means p is not an inducing path.

To conclude, we have shown that if there is an inducing path between $A(t), B(t+k)$, then $A(t) \in \mathbf{An}(B(t+k))$, which means all lagged edges in the MAG are directed ones. □

B.2 Proof of Proposition 3.3: MAG to summary DAG

Proposition 3.3.*If there are $A(t) \rightarrow B(t+k)$ and $A(t+k) \leftrightarrow B(t+k)$ in the MAG, then, in the summary DAG, there is either*

1. *$A \rightarrow B$,*
2. *a directed path from A to B with length $l \leq k-2$, or*
3. *a directed path from A to B with length $l = k-1$ and a confounding structure between them with lengths $(r \leq k-2, q \leq k-2)$.*

Proof. 1. We first show that If there is $A \rightarrow B$ in the summary DAG, then, there are $A(t) \rightarrow B(t+k)$ and $A(t+k) \leftrightarrow B(t+k)$ in the MAG:

   Suppose that there is $A \rightarrow B$ in the summary DAG. Then, there is a directed inducing path $p_1 = A(t) \rightarrow A(t+1) \rightarrow \cdots \rightarrow A(t+k-1) \rightarrow B(t+k)$ from $A(t)$ to $B(t+k)$ in the full time DAG. Thus, we have $A(t) \rightarrow B(t+k)$ in the MAG.

   Besides, there is an inducing path $p_2 = A(t+k) \leftarrow A(t+k-1) \rightarrow B(t+k)$ in the full time DAG, which means we have $A(t+k) \leftrightarrow B(t+k)$ in the MAG.

2. We then show that if there is $A(t) \rightarrow B(t+k)$ in the MAG, then, there is a directed path from A to B with length $0 \leq l \leq k-1$ (a directed path with $l=0$ means $A \rightarrow B$) in the summary DAG:

   Suppose that there is $A(t) \rightarrow B(t+k)$ in the MAG. Then, there is a directed path $A(t) \rightarrow V_1(t+1) \rightarrow \cdots \rightarrow V_{k-1}(t+k-1) \rightarrow B(t+k)$ in the full time DAG. Denote the number of V_i such that $V_i = A$ as r, then, we have $0 \leq r \leq k-1$. Hence, in the summary DAG, there is a directed path from A to B with length $l = k-r-1$ and we have $0 \leq l \leq k-1$.

3. We finally show that if there is $A(t+k) \leftrightarrow B(t+k)$ in the MAG, then, in the summary DAG, at least one of the following structures exists:

[3] If all edges among the path are $\rightarrow$, then $A(t) \in \mathbf{An}(B(t+k))$.

a. A directed path[4] p between A and B of length $0 \leq l \leq k-2$ (a directed path with $l=0$ means $A \to B$);
b. A confounding structure c between A and B of length $(r \leq k-2, q \leq k-2)$.

Suppose that there is $A(t+k) \leftrightarrow B(t+k)$ in the MAG, we will show that there is a latent confounder U between $A(t+k)$ and $B(t+k)$ such that $t+1 \leq time(U) \leq t+k-1$.

In this regard, if $U = A$ or $U = B$, there is a directed path between A and B with length $0 \leq l \leq k-2$; Otherwise when $U \neq A$ and $U \neq B$, there is a confounding structure between A and B with length $(r \leq k-2, q \leq k-2)$.

In the following, we prove that there is a latent confounder U between $A(t+k)$ and $B(t+k)$ such that $t+1 \leq time(U) \leq t+k-1$. Prove by contradiction. Suppose that $time(U) \leq t$ or $time(U) \geq t+k$. If $time(U) \leq t$, then on the directed inducing path from U to A, there is a vertex at time t, which means the path can not be an inducing path. If $time(U) \geq t+k$, then there is a collider on the inducing path between U and A according to Lemma B.1, which means the path is not directed and U is not a latent confounder. Hence, the time of the latent confounder U is between $t+1$ and $t+k-1$.

To conclude, combining the results in 2. and 3., we prove the proposition. □

B.3 Proof of Theorem 3.5: Identifiability of the summary DAG

For two vertex A, B, let the vertex set $\mathbf{M}$ contain A and any $M_i \neq B$ such that $A(t) \to M_i(t+k)$ in the MAG and M_i is not B's descendant[5]. Let the vertex set $\mathbf{S}$ contains any $S_i \neq A$ such that $S_i(t) \to B(t+k)$ or $S_i(t) \to M_j(t+k)$ for some $M_j \in \mathbf{M}$ in the MAG.

Theorem 3.5. *Assuming model (1), Assumption 2.2, and Assumptions 2.6,2.7,2.11, then the summary DAG is identifiable. Specifically,*

1. *There is $A \to B$ in the summary DAG iff there are $A(t) \to B(t+k), A(t+k) \leftrightarrow B(t+k)$ in the MAG, and the set $\mathbf{M}(t+1) \cup \mathbf{S}(t)$ is not sufficient to d-separate $A(t), B(t+k)$ in the full time DAG.*
2. *The condition "the set $\mathbf{M}(t+1) \cup \mathbf{S}(t)$ is not sufficient to d-separate $A(t), B(t+k)$ in the full time DAG" can be tested by the proxy variable $\mathbf{M}(t+k)$ of the unobserved set $\mathbf{M}(t+1)$.*

Proof. 1. $\Rightarrow$. Suppose that there is $A \to B$ in the summary DAG. According to Proposition 3.3, we have $A(t) \to B(t+k)$ and $A(t+k) \leftrightarrow B(t+k)$ in the MAG. In addition, in the full time DAG, there is a directed path $p = A(t) \to B(t+1) \to \cdots \to B(t+k)$ from $A(t)$ to $B(t+k)$, which is not d-separated by $\mathbf{M}(t+1) \cup \mathbf{S}(t)$ (because $B \notin \mathbf{M}$).

$\Leftarrow$. Suppose that there are $A(t) \to B(t+k)$ and $A(t+k) \leftrightarrow B(t+k)$ in the MAG. According to Proposition 3.3, at least one of the following structures exists:

a. $A \to B$;
b. A directed path p_{AB} from A to B of length $0 < l \leq k-2$;
c. A directed path p_{AB} from A to B of length $l = k-1$ and a confounding structure c_{AB} between A and B of length $(r \leq k-2, q \leq k-2)$.

In the following, we show that if the set $\mathbf{M}(t{+}1) \cup \mathbf{S}(t)$ is not sufficient to d-separate $A(t), B(t{+}k)$ in the full time DAG, $A \to B$ must exist in the summary DAG. We prove this by its contrapositive statement, *i.e.*, if $A \to B$ does not exist, the set $\mathbf{M}(t+1) \cup \mathbf{S}(t)$ is sufficient to d-separate $A(t)$ and $B(t+k)$.

Consider the path between $A(t)$ and $B(t+k)$ in the full time DAG $p = A(t), V_1, ..., V_l, B(t+k)$. There are three possible cases:

a. $\exists\, V_i, time(V_i) \geq t+k$;
b. $\forall\, V_i, time(V_i) < t+k$ and $\exists\, V_j$ such that $time(V_j) \leq t$;

[4]The path can be from A to B or from B to A.
[5]This can be justified from the MAG. Specifically, if M_i is the descendant of B, then there is $B(t) \to V_1(t+k), V_1(t) \to V_2(t+k), ..., V_l(t) \to M_i(t+k)$ in the MAG.

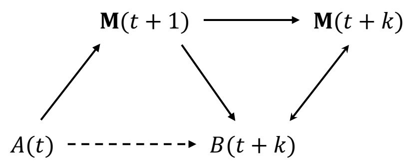

Figure 1: Causal graph over $A(t), B(t+k), \mathbf{M}(t+1), \mathbf{M}(t+k)$.

c. $\forall\, V_i, t < time(V_i) < t + k$.

Next, we will show that if $A \to B$ does not exist, the set $\mathbf{M}(t+1) \cup \mathbf{S}(t)$ can *d*-separate all these three kinds of paths.

a. For the path that $\exists\, V_i, time(V_i) \geq t + k$. According to Lemma B.1, there is a collider at $time \geq t + k$. Because the collider and its descendants are not in $\mathbf{M}(t+1) \cup \mathbf{S}(t)$, the path is *d*-separated.
b. For the path that $\forall\, V_i, time(V_i) < t + k$ and $\exists\, V_j$ such that $time(V_j) \leq t$. There is a vertex V_j on the path such that $V_j(t) \to V_{j+1}(t+1) \to V_{j+2}(t+2) \to \cdots$[6] If the path from $V_j(t)$ to $B(t+k)$ is a directed path, we have $V_j \in \mathbf{S}$ and the path is *d*-separated. Otherwise, there is a collider on the path at $time \geq t + 2$. Neither the collider nor its descendants are in $\mathbf{M}(t+1) \cup \mathbf{S}(t)$, the path is also *d*-separated.
c. For the path that $\forall\, V_i, t < time(V_i) < t + k$. We have $A(t) \to V_1(t+1) \to V_2(t+2) \to \cdots$. If the path from $V_1(t+1)$ to $B(t+k)$ is a directed one, $V_1 \in \mathbf{M}$[7]. Otherwise, there is a collider on the path at $time \geq t + 2$. Neither the collider nor its descendants are in $\mathbf{M}(t+1) \cup \mathbf{S}(t)$, the path is also *d*-separated.

2. In the following, we show that the condition "the set $\mathbf{M}(t+1) \cup \mathbf{S}(t)$ is not sufficient to *d*-separate $A(t), B(t+k)$ in the full time DAG" can be justified by testing $A(t) \perp\!\!\!\perp B(t+k) | \mathbf{M}(t+1) \cup \mathbf{S}(t)$ with proxy variables.

a. We first explain that $A(t), B(t+k), \mathbf{M}(t+1), \mathbf{M}(t+k)$ have the causal graph shown in Fig. 1.
   - $A(t) \to \mathbf{M}(t+1)$. By definition, A is the ancestor of vertices in $\mathbf{M}$.
   - $\mathbf{M}(t+1) \to \mathbf{M}(t+k)$. By Assumption 2.10the self causation edge always exists.
   - $\mathbf{M}(t+1) \to B(t+k)$. By definition, B is not the ancestor of any vertex in $\mathbf{M}$.
   - $\mathbf{M}(t+1) \leftrightarrow B(t+1)$. According to Proposition 3.1, the instantaneous edge must be bi-directed.

   Indeed, the directed edge $\mathbf{M}(t+1) \to B(t+k)$ and the bi-directed edge $\mathbf{M}(t+k) \leftrightarrow B(t+k)$ may not exist. However, according to the theory of single proxy causal discovery, as long as we have $A(t) \perp\!\!\!\perp \mathbf{M}(t+k) | \mathbf{M}(t+1), \mathbf{S}(t)$ (which will be proved in the following), we can test whether $A(t) \to B(t+k)$ with the proxy variable $\mathbf{M}(t+k)$.
b. We next show that $A(t) \perp\!\!\!\perp \mathbf{M}(t+k) | \mathbf{M}(t+1), \mathbf{S}(t)$, *i.e.*, $\mathbf{M}(t+k)$ can act as a legal proxy variable of $\mathbf{M}(t+1)$. We prove this by showing that all path between $A(t)$ to any $Z_i(t+k) \in \mathbf{M}(t+k)$ can be *d*-separated by $\mathbf{M}(t+1) \cup \mathbf{S}(t+1)$. Specifically, we consider two different cases:
   - When $A \not\to B$ in the summary DAG. To prove this conclusion, again, consider three kinds of paths $A(t), V_1, ..., V_l, Z_i(t+k)$ between $A(t)$ and $Z_i(t+k)$,
     i. For the path that $\exists\, V_i, time(V_i) \geq t + k$. It must contain a collider at $time \geq t + k$. Hence, neither the collider nor its descendant is in $\mathbf{M}(t+1) \cup \mathbf{S}(t)$ and the path is *d*-separated.
     ii. For the path that $\forall\, V_i, time(V_i) < t + k$ and $\exists\, V_j$ such that $time(V_j) \leq t$. Then, there is a vertex V_j on the path such that $V_j(t) \to V_{j+1}(t+1) \to V_{j+2}(t+2) \to \cdots$. If the path between $V_j(t)$ and $Z_i(t+k)$ is a directed one, then we have $V_j(t) \in \mathbf{S}(t)$ and the path is *d*-separated. Otherwise, there is a collider on the path between $V_{j+2}(t+2)$ and

[6]Strictly, when $k = 2$, V_{j+2} is $B(t+2)$.
[7]$A(t) \to V_1(t+1)$ means we have $A(t) \to V_1(t+k)$ in the MAG. V_1 is the ancestor of B, so $V_1 \notin \mathbf{De}(B)$. Hence, $V_i \in \mathbf{M}$.

$Z_i(t+k)$. Thus, neither the collider nor its descendant is in $\mathbf{M}(t+1) \cup \mathbf{S}(t)$ and the path is also d-separated.

iii. For the path that $\forall\, V_i, t < time(V_i) < t+k$. In this case, the path must be $A(t) \rightarrow V_1(t+1) \rightarrow V_2(t+2) \rightarrow \cdots$. If the path between $V_1(t1)$ and $Z_i(t+k)$ is a directed one, we have $V_1(t+1) \in \mathbf{M}(t+1)$, because of the following facts: $V_1 \neq B$, $A \rightarrow V_1$, $V_1 \notin \mathbf{De}(B)$[8]. As a result, the path is d-separated. Otherwise, when the path between $V_1(t1)$ and $Z_i(t+k)$ is not a directed one, there is a collider on the path at $time \geq t+2$. Therefore, neither the collider nor its descendant is in $\mathbf{M}(t+1) \cup \mathbf{S}(t)$ and the path is d-separated.

- When $A \rightarrow B$ in the summary DAG. In this case, except for the above analysis, we need to extra show that every path $p = A(t) \rightarrow B(t+1), V_1, ..., V_l, Z_i(t+k)$ can be d-separated by $\mathbf{M}(t+1) \cup \mathbf{S}(t)$. Specifically, since we define $Z_i \notin \mathbf{De}(B)$, p can not be a directed path. Therefore, starting from $B(t+1)$ along the path, at least one of the edges is $\leftarrow$. As a result, there is a collider on p which is either $B(t+1)$ itself or at $time \geq t+2$. For both cases, neither the collider nor its descendant is in $\mathbf{M}(t+1) \cup \mathbf{S}(t)$ and the path is d-separated.

□

[8]If $V_1 \in \mathbf{De}(B)$, since V_1 is the ancestor of Z_i, we have $Z_i \in \mathbf{De}(B)$, which contradicts with the definition of the vertex set $\mathbf{M}$.

C Experiment

C.1 Extra results of Section 5.1: Synthetic study

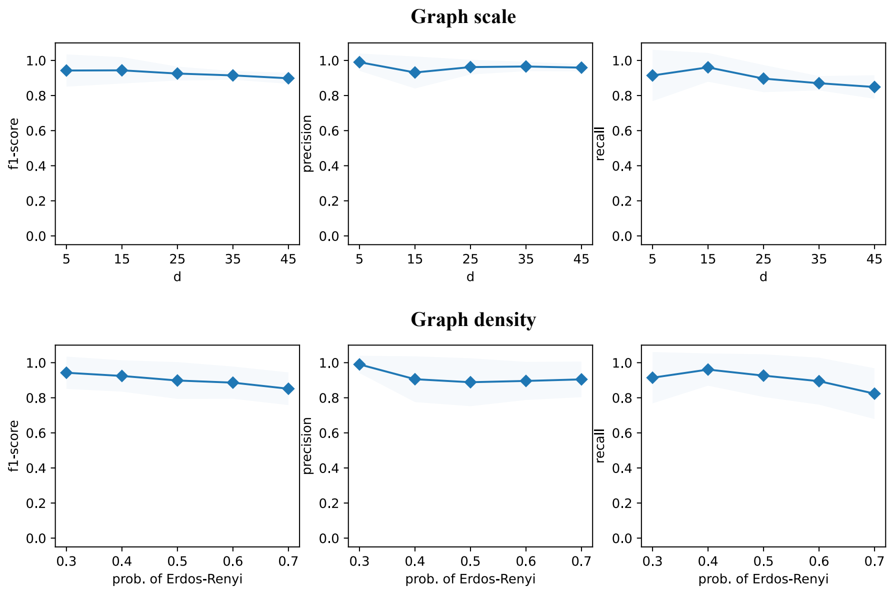

Figure 2: Performance of our method under different graph scales (d denotes the variable number) and densities ($\mathrm{prob.}$ denotes the edge probability in the Erdos-Renyi model).

C.2 Extra results of Section 5.2: Discovering causal pathways in Alzheimer's disease

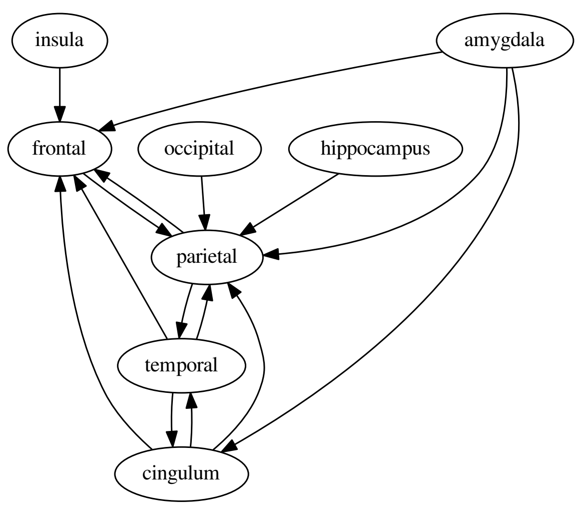

Figure 3: Summary DAG over eight meta-regions in AD recovered by the NG-EM [2] baseline.

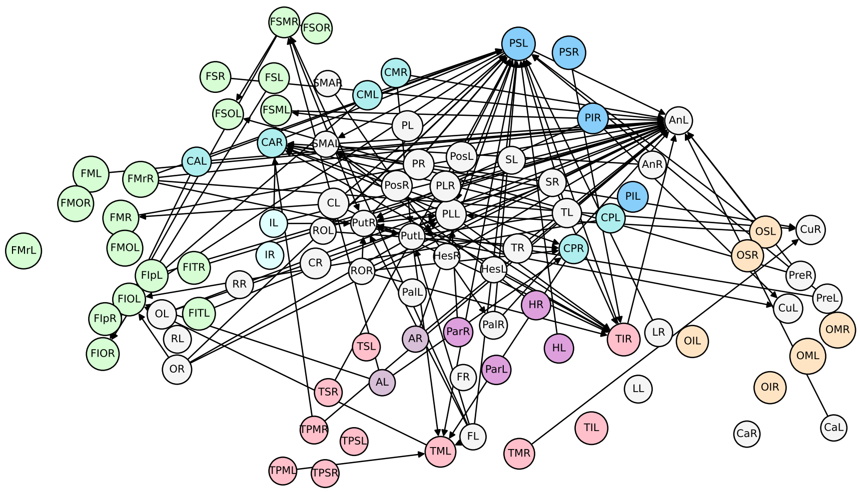

Figure 4: Summary DAG in AD recovered by the NG-EM [2] baseline.

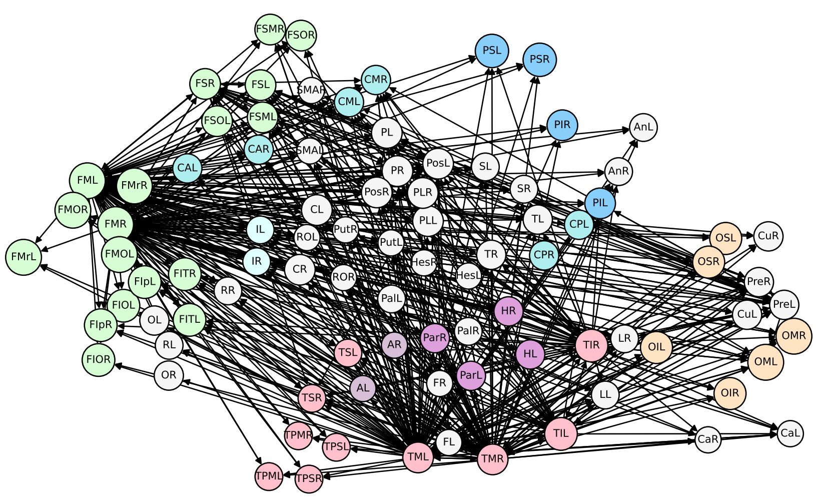

Figure 5: Summary DAG in AD recovered by the Dynotears [3] baseline.

Table 2: Indices for brain regions partition in Fig. 7.

AAL Index	Abbreviation	Full name	AAL Index	Abbreviation	Full name
1	PL	Precentral L	2	PR	Precentral R
3	FSL	Frontal Sup L	4	FSR	Frontal Sup R
5	FSOL	Frontal Sup Orb L	6	FSOR	Frontal Sup Orb R
7	FML	Frontal Mid L	8	FMR	Frontal Mid R
9	FMOL	Frontal Mid Orb L1	10	FMOR	Frontal Mid Orb R1
11	FIOL	Frontal Inf Oper L	12	FIOR	Frontal Inf Oper R
13	FITL	Frontal Inf Tri L	14	FITR	Frontal Inf Tri R
15	FIpL	Frontal Inf Orb L	16	FIpR	Frontal Inf Orb R
17	ROL	Rolandic Oper L	18	ROR	Rolandic Oper R
19	SMAL	Supp Motor Area L	20	SMAR	Supp Motor Area R
21	OL	Olfactory L	22	OR	Olfactory R
23	FSML	Frontal Sup Medial L	24	FSMR	Frontal Sup Medial R
25	FMrL	Frontal Mid Orb L2	26	FMrR	Frontal Mid Orb R2
27	RL	Rectus L	28	RR	Rectus R
29	IL	Insula L	30	IR	Insula R
31	CAL	Cingulum Ant L	32	CAR	Cingulum Ant R
33	CML	Cingulum Mid L	34	CMR	Cingulum Mid R
35	CPL	Cingulum Post L	36	CPR	Cingulum Post R
37	HL	Hippocampus L	38	HR	Hippocampus R
39	ParL	ParaHippocampal L	40	ParR	ParaHippocampal R
41	AL	Amygdala L	42	AR	Amygdala R
43	CaL	Calcarine L	44	CaR	Calcarine R
45	CuL	Cuneus L	46	CuR	Cuneus R
47	LL	Lingual L	48	LR	Lingual R
49	OSL	Occipital Sup L	50	OSR	Occipital Sup R
51	OML	Occipital Mid L	52	OMR	Occipital Mid R
53	OIL	Occipital Inf L	54	OIR	Occipital Inf R
55	FL	Fusiform L	56	FR	Fusiform R
57	PosL	Postcentral L	58	PosR	Postcentral R
59	PSL	Parietal Sup L	60	PSR	Parietal Sup R
61	PIL	Parietal Inf L	62	PIR	Parietal Inf R
63	SL	SupraMarginal L	64	SR	SupraMarginal R
65	AnL	Angular L	66	AnR	Angular R
67	PreL	Precuneus L	68	PreR	Precuneus R
69	PLL	Paracentral Lobule L	70	PLR	Paracentral Lobule R
71	CL	Caudate L	72	CR	Caudate R
73	PutL	Putamen L	74	PutR	Putamen R
75	PalL	Pallidum L	76	PalR	Pallidum R
77	TL	Thalamus L	78	TR	Thalamus R
79	HesL	Heschl L	80	HesR	Heschl R
81	TSL	Temporal Sup L	82	TSR	Temporal Sup R
83	TPSL	Temporal Pole Sup L	84	TPSR	Temporal Pole Sup R
85	TML	Temporal Mid L	86	TMR	Temporal Mid R
87	TPML	Temporal Pole Mid L	88	TPMR	Temporal Pole Mid R
89	TIL	Temporal Inf L	90	TIR	Temporal Inf R

References

[1] Jiji Zhang. On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. *Artificial Intelligence*, 172(16-17):1873–1896, 2008.

[2] Mingming Gong, Kun Zhang, Bernhard Schoelkopf, Dacheng Tao, and Philipp Geiger. Discovering temporal causal relations from subsampled data. In *International Conference on Machine Learning*, pages 1898–1906. PMLR, 2015.

[3] Roxana Pamfil, Nisara Sriwattanaworachai, Shaan Desai, Philip Pilgerstorfer, Konstantinos Georgatzis, Paul Beaumont, and Bryon Aragam. Dynotears: Structure learning from time-series data. In *International Conference on Artificial Intelligence and Statistics*, pages 1595–1605. PMLR, 2020.