A Example and Analysis for Algorithm 1

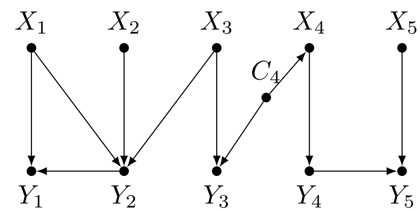

Figure 11: An interaction network of 5 individuals.

Example 2. *Input: $t = 3$, interaction network in Figure 11.*

Iteration 1: Units = [1, 5, 4, 2, 3]. $B = \{5, 2\}$.

Iteration 2: Units = [5, 1, 4, 3, 2]. $B = \{5, 3\}$.

Iteration 3: Units = [5, 2, 3, 1, 4]. $B = \{5, 2\}$.

The three choices of B all have the same size 2. So the output is any of the three choices of B.

Note that the subnetwork formed by 5 and 3 contains a bidirected path between Y_3 and Y_5 (due to the path $Y_3 \leftarrow C_4 \rightarrow X_4 \rightarrow Y_4 \rightarrow Y_5$), and this does not constitute a bias structure.

Complexity Analysis The time complexity is $O(tn^2d^p)$. d is the maximum degree of each node (how many other nodes a node is directly connected to), and p is the length (number of edges) of the longest simple path. This is polynomial if the degree is bounded.

Lemma 2. *The following two statements are equivalent. The first statement is used in this algorithm for simpler computation, and the second statement is used in the main text for easier understanding.*

1. *For each individual i in B, i has no deflecting bias structure in G^* with another individual j in B.*
2. *For each individual i in B, i has no deflecting bias structure in the latent projection of G^* on B.*

The definition of latent projection is by Pearl [2009], as follows.

Definition 9 (Projection[Pearl, 2009])**.** *A latent structure $L_{[O]} =< D_{[O]}, O >$ is a projection of another latent structure L if and only if:*

1. *every unobservable variable of $D_{[O]}$ is a parentless common cause of exactly two nonadjacent observable variables; and*
2. *for every stable distribution P generated by L, there exists a stable distribution P' generated by $L_{[O]}$ such that $I(P_{[O]}) = I(P'_{[O]})$.*

Proof of Lemma 2.

Proof. If statement 1 is false, then there exists an open path between X_i and Y_j in G^*, where $i, j \in B$. The latent projection contains both i and j so the open path still exists, which imply a deflecting bias structure in the latent projection.

If statement 2 is false, then there exists an open path between X_i and Y_j in the latent projection. This implies a deflecting bias structure in G^*. □

B An Additional Simulation

Experiment: Subset Size of THM-2 We use same parameter settings as the previous experiment, except that we let $dRate$ and $rRate$ vary in $0.01, 0.1, 0.3, 0.5$. The subset sizes selected by THM-2 are in Table 1. Observe that as the graph gets denser (larger $dRate$ and $rRate$), THM-2 is unable to use most of the input samples. However, for the tests with samples ≥ 3, THM-2 yields very accurate estimates. Given that the ground truth is 100, **the estimates of THM-2 range between** 99.96 **and** 100.06**.**

		$dRate$			
		0.01	0.1	0.3	0.5
$rRate$	0.01	155	147	131	115
	0.1	26	24	23	23
	0.3	9	8	8	8
	0.5	5	4	3	0

Table 1: Each cell denotes the subset size selected using THM-2.

C Proof of the Theorems

All lemmas and proofs are attached in Section D of the appendix.

Theorem 1. *Let $M^*(G^*, S^*)$ be a balanced interaction model in which treatment variable X_i and outcome variable Y_i are not confounded by any variable in $\mathcal{V}_i$, $\forall i$. Let D be the available data generated by M^* and let $G^\dagger$ be the approximate graph constructed using D. Let $TACE_{XY}$ be identifiable in $G^\dagger$ and be given by β_{YX}, the regression coefficient of Y on X. Let α denote the true value of $TACE_{X,Y}$ in M^*. If X satisfies ASDC then the interaction bias is given by,*

$$\left|E[\hat{\beta_{YX}}] - \alpha\right| = \left|\frac{1}{n} \sum_{1\leq i\leq n} \sum_{p\in P[iji]} Val(p)\frac{\sigma^2_{R_p}}{\sigma^2_X} - \frac{1}{n(n-1)} \sum_{1\leq i\leq n} \sum_{p\in P[ji]} Val(p)\frac{\sigma^2_{R_p}}{\sigma^2_X}\right|,$$

where $P[iji]$ is the set of reflecting bias structures between X_i and Y_i through any explicit variable W_j of unit j with $i \neq j$, $P[ji]$ is the set of deflecting bias structures between X_j and Y_i with $i \neq j$, and R_p is the root of path p.

Proof. By Lemma 9,

$$\begin{aligned}
&E[\hat{\beta_{YX}}] \\
=&\alpha \\
&+ \frac{1}{n}\Big(\sum_{p\in\mathcal{P}} Val(p) + \sum_{1\leq i\leq n} \sum_{R\in(\mathcal{R}[iji]\setminus\{X_i\})} c_R\beta_{RX}\Big) \\
&- \frac{1}{n(n-1)} \sum_{1\leq i\leq n} \sum_{R\in\mathcal{R}[ji]} c_R\beta_{RX},
\end{aligned}$$

where $\mathcal{P}$ is the set of directed paths from X_i to Y_i for any i passing through an intermediate node $W_j \in \mathcal{V}_{(j)}$, $i \neq j$, $\mathcal{R}[iji]$ is the set of roots of the open paths between X_i and Y_i through some W_j with $j \neq i$, $\mathcal{R}[ji]$ is the set of roots of the open paths between X_j and Y_i for $j \neq i$, and c_R is the sum of values of the directed paths from a variable R ($\in (\mathcal{R}[iji] \setminus \{X_i\})$ or $\in \mathcal{R}[ji]$) to Y_i not passing through any variable in $\mathcal{R}[iji] \cup \mathcal{R}[ji]$ for any $j \neq i$.

We prove this is equivalent to

$$\begin{aligned}
&E[\hat{\beta_{YX}}] \\
=&\alpha \\
&+ \frac{1}{n} \sum_{1\leq i\leq n} \sum_{p\in P[iji]} Val(p)\frac{\sigma^2_{R_p}}{\sigma^2_X} \\
&- \frac{1}{n(n-1)} \sum_{1\leq i\leq n} \sum_{p\in P[ji]} Val(p)\frac{\sigma^2_{R_p}}{\sigma^2_X},
\end{aligned}$$

where $P[iji]$ is the set of open paths between X_i and Y_i through any $W_j \in \mathcal{V}_{(j)}$ with $i \neq j$, $P[ji]$ is the set of open paths between X_j and Y_i through any $W_j \in \mathcal{V}_{(j)}$ with $i \neq j$, and R_p is the root of path p.

We first check the term $\sum_{R \in \mathcal{R}[ji]} c_R \beta_{RX}$. For an R that is the root of a path between X_j and Y_i, since X satisfies ASDC, we must have $R \in \mathcal{V}_{(j)}$. Rename it as R_j. We also have $\beta_{RX} = \sigma_{RX}/\sigma_X^2$. By Wright's Rules, σ_{RX} is equal to the sum of open path values between R and X times the variance of the root of that path. Recall that $R \in Anc(X)$, X satisfies ASDC, so R satisfies ASDC. So σ_{RX} is equal to the sum of open path values between R_j and X_j times the variance of the root of that path. We prove that each term that appears in $A = \sum\limits_{p \in P[ji]} Val(p) \frac{\sigma_{R_p}^2}{\sigma_X^2}$ also appears in $B = \sum_{R \in \mathcal{R}[ji]} c_R \beta_{RX}$, and there is no extra term.

Each R_p in A is a root between X_j and Y_i for some $j \neq i$, and must be included if it is a root. So we just have to check all the roots between X_j and Y_i for some $j \neq i$. For each root R_p, we check where in B will $\sigma_{R_p}^2/\sigma_X^2$ exist. When R in B is R_p, the term containing $\sigma_{R_p}^2/\sigma_X^2$ in β_{RX} is the sum of paths from R_p to X_j where R_p is the root, so is the sum of directed paths from R_p to X_j. So the term containing $\sigma_{R_p}^2/\sigma_X^2$ in $c_R \beta_{RX}$ is the sum of paths between Y_i and X_j through R_p with 1) R_p being the root and 2) the sub-path from R_p to Y_i does not go through any variable in $\mathcal{R}[iki] \cup \mathcal{R}[ki]$ for any $k \neq i$.

The terms that are left in $Val(p) \frac{\sigma_{R_p}^2}{\sigma_X^2}$ to cover in B are the $X_j - R_p - Y_i$ paths whose sub-path from R_p to Y_i go through some variable in $\mathcal{R}[iki] \cup \mathcal{R}[ki]$ for any $k \neq i$. We just have to go over all types of R in B, and see which ones contain $\sigma_{R_p}^2/\sigma_X^2$.

Case 1: $R \in Anc(R_p)$**.** There is no such a path in c_R or β_{RX}. c_R does not go through R since $R \in \mathcal{R}[ji] c_R \beta_{RX}$. β_{RX} also does not contain $\sigma_{R_p}^2$ since $R \in Anc(R_p)$, so R_p is never a root on any paths between R and X_j. Hence $c_R \beta_{RX}$ does not contain such a path.

Case 2: $R \in Desc(R_p)$**.** Again, c_R does not contain R_p. However β_{RX} contains $\sigma_{R_p}^2$. R_p can be a root on some paths between R and X_j. Those paths are from R_p to R and R_p to X_j. Recall that c_R denotes directed paths from R to Y_i. The term that contains $\sigma_{R_p}^2$ in $c_R \beta_{RX}$ are the paths between X_j and Y_i, that pass through some variable in $\mathcal{R}[iki] \cup \mathcal{R}[ki]$ (R), with R_p being the root. As a result, this case completely cover the missing term.

Case 3: $R \perp\!\!\!\perp R_p$**.** It is easy to derive that in this case, $c_R \beta_{RX}$ does not contain a path that goes through R_p. Otherwise R and R_p would be dependent.

Case 4: R **and** R_p **are only connected through common ancestors.** In this case, in any path that contains both R and R_p, R_p will not be the root. Their common ancestors will be the roots. So this case also does not provide any term containing $\sigma_{R_p}^2/\sigma_X^2$.

We have proved that for every R_p in A, the coefficient of $\sigma_{R_p}^2/\sigma_X^2$ (equal to a sum of those paths in $P[ji]$ with R_p being the root) is equal to the the coefficient of $\sigma_{R_p}^2/\sigma_X^2$ in B. As stated before, A and B have the same set of roots, so they have the same $\sigma_{R_p}^2/\sigma_X^2$ terms. So the sum of those terms are equal.

Next, we prove the reflecting bias terms are also equal. Observe that $\bigcup_{1 \leq i \leq n} P[iji] = \mathcal{P}$, so we just have to prove that $\sum_{p \in P[iji]} Val(p) + \sum_{R \in (\mathcal{R}[iji] \setminus \{X_i\})} c_R \beta_{RX}$ is equivalent to $\sum_{p \in P[iji]} Val(p) \frac{\sigma_{R_p}^2}{\sigma_X^2}$. This can be proven using the exact same reasoning above, so we omit the proof.

Thus, the two expressions for $E[\hat{\beta_{YX}}]$ are equivalent. $\square$

Corollary 1. *Let $M^{**}(G^{**}, S)$ be a balanced interaction model in which X satisfies ASDC and TACE is identified as $\beta_{YX} = \alpha$ in the approximate graph, then interaction bias exists iff G^{**} contains a reflecting or deflecting bias structure.*

Proof. (if part) Follows from theorem 1.There are two terms that cause bias in theorem 1 and they can be attributed to the two bias structures.
(only if part) Had there been additional structures that caused bias, then theorem 1 would have had additional terms to account for it. Since theorem 1 has only two bias terms fully accounted for by the two structures, there exist no other structure that creates bias. □

Theorem 2. *Let G^* be an interaction network. Given the conditions in Theorem 1 and 'B' a bias-free subset for G^*, $TACE_{XY} = E[\hat{\beta_{YX}}]$ where the regression coefficient is calculated using only samples in set B.*

Proof. We check the interaction network G_S^* formed by B, by treating any variable from $\mathcal{V}_{(j)}$ where $j \notin S$ as unobserved. Next, we calculate $E[\hat{\beta_{YX}}]$ for G_S^*.

By Theorem 1,

$$E[\hat{\beta_{YX}}] = \alpha + \frac{1}{n}Term_2 - \frac{1}{n(n-1)}Term_3.$$

The second term is obtained by summing over paths of the form: $X_i - \cdots - W_j - \cdots - Y_i$, and the third term is obtained by summing over paths of the form: $X_i - \cdots - Y_j$. These paths do not exist in G_S^*. Hence, the two bias terms are 0, and $E[\hat{\beta_{YX}}] = \alpha$. □

D Lemmas

Lemma 1. *If W satisfies ASDC, then any two explicit variables W_i and W_j are IID (Independent and Identically Distributed.)*

Proof. If W satisfies ASDC, and W_i is the root for some i, then from the third property of ASDC, W_i must be the root for all i. The roots are only caused by their error terms, the error terms are IID (identically distributed and independent), so W is IID.

If W_i is not the root for any i, W satisfies ASDC, and all its parents are IID, then we have for any i

$$W_i = \sum_{V_i \in Pa(W)} c_{V_i} V_i + U_{W_i},$$

where c_{V_i} is the coefficient of the variable V_i on the edge $V_i \to W_i$. Each term is IID for any $i \neq j$. So W_i and W_j are IID.

If W_i is not the root for any i, W satisfies ASDC, and there exists a parent of W, V such that V_i and V_j are not IID. Then from our previous derivation, there exists a parent of V, V', such that V_i' and V_j' are not IID. Keep tracing up until a root variable R, such that R_i and R_j are not IID. However, this violates our derivation in the beginning, that if a variable is the root and satisfies ASDC, it must be IID. We reach a contradiction. Hence, if W_i is not the root for any i, W satisfies ASDC, then all its parents are IID, and W is thus IID. □

Lemma 3. *Let $\mathcal{X} = \{X_1, \ldots, X_n\}$ be n IID random variables where the $\sigma_X^2 > 0$, and a random variable W_i. Among $\mathcal{X}$, W_i is dependent of X_i only, and $W_i = aX_i + b$ where a and b are constants. Then the following expectation exists.*

$$E\left[\frac{(X_i - \bar{X})W_i}{\sum\limits_{1 \le i \le n}(X_i - \bar{X})^2}\right].$$

Proof. We have to prove that the function $f(X_1, \ldots, X_n, W_i)$ inside of the expectation is bounded. For convenience, rewrite it by plugging in $W_i = aX_i + b$.

$$E\left[\frac{(X_i - \bar{X})(aX_i + b)}{\sum\limits_{1 \le i \le n}(X_i - \bar{X})^2}\right].$$

For any X_j with $j \neq i$, the denominator is a quadratic function on X_j, and the numerator is a linear function of X_j from the term $\bar{X}$. For X_i, the denominator is a quadratic function on X_i, and the numerator is a quadratic function on X_i. Since $\sigma_X^2 \neq 0$, $X_1, \ldots, X_n$ cannot take on the same value, so the denominator is always positive. When considering X_i as the variable, f might only go to infinity when X_i goes to infinity or negative infinity, and same with X_j.

When considering X_i as the variable, and X_j for all other j as constants, the denominator can be written in the form of $AX_i^2 + BX_i + C$, with A, B, C being constants. Hence, the order (of the polynomial) of the denominator is 2, and the order of the numerator is 2. So the limit of f when X_i goes to ∞ or $-\infty$ is a finite value equal to the ratio of the coefficient of X_i^2 in the numerator divided by the coefficient of X_i^2 in the denominator.

When considering X_j as the variable, the order of the denominator is 2, and the order of the numerator is 1. So the limit of f when X_i goes to ∞ or $-\infty$ is 0. Hence, f is bounded. $\square$

Lemma 4. *Given a balanced interaction model $M^{**}(G^{**}, S^{**})$, if generic variables V and X both satisfy ASDC, and $dSep(V_i, X_i|\emptyset)$ for all i in G^{**}, then*

$$E\left[\frac{\sum\limits_{1\leq i\leq n}(X_i - \bar{X})V_i}{\sum\limits_{1\leq i\leq n}(X_i - \bar{X})^2}\right] = 0.$$

Proof. The d-separation condition implies $X_i \perp\!\!\!\perp V_i$. V and X are IID implies that we can treat all X_i's as the same variable X, and treat all V_i's as the same variable V. Hence, $X \perp\!\!\!\perp V$ and $\sigma_{XV} = 0$, which gives $\beta_{VX} = \sigma_{XV}\sigma_X^2 = 0$. Also note that

$$\hat{\beta_{VX}} = \frac{\sum\limits_{1\leq i\leq n}(X_i - \bar{X})V_i}{\sum\limits_{1\leq i\leq n}(X_i - \bar{X})^2}.$$

Since the ordinary least squares estimator is unbiased, we have $E[\hat{\beta_{VX}}] = \beta_{VX} = 0$. $\square$

Lemma 5. *Given a balanced interaction model, with the following conditions: 1) X_i and Y_i are not confounded by a path containing only variables in $\mathcal{V}_i$, $\forall i$, and 2) X_i satisfies ASDC. Then there exists a set $\mathcal{S}$ consisting of the following three subsets of explicit variables:*

1. *$\mathcal{S}_1$: X_i,*
2. *$\mathcal{S}_2$: the root variables (excluding X_i) of each open path between X_j and Y_i (j can be the same as i),*
3. *$\mathcal{S}_3$: the root variables of this interaction network that are in $Anc(Y_i)$ and d-separated (by an empty set) from X_j for all j,*

such that Y_i can be expressed as a linear function of the variables in S i.e.,

$$Y_i = \sum_{W_t \in \mathcal{S}} c_{W_t} W_t,$$

where c_{W_t} is equal to the sum of the values of the directed paths from W_t to Y_i that do not go through any variable in $\mathcal{S}$.

Proof. Consider the following protocol.

- Start from the initial structural equation of Y_i, $Y_i = f(Pa(Y_i))$, denoted $SE(Y_i)$.
- For each variable A_q in the r.h.s. of $SE(Y_i)$,
  - if $A_q \in \mathcal{S}$, keep it.
  - if $A_q \notin \mathcal{S}$ and not a root of the network, replace it with its structural equation, $A_q = g(Pa(A_q))$ and plug it into $SE(Y_i)$.

– if $A_q \notin \mathcal{S}$ and is a root of the network, keep it.

- Keep replacing until no more replacement can be done in the r.h.s. of $SE(Y_i)$.
- Denote the final $SE(Y_i)$ as $SE_f(Y_i)$.

We prove $SE_f(Y_i)$ is

$$Y_i = \sum_{W_t \in \mathcal{S}} c_{W_t} W_t,$$

where c_{W_t} is equal to the sum of the product of path coefficients of the directed paths from W_t to Y_i that do not go through any variable in $\mathcal{S}$.

First, we prove that the r.h.s. of $SE_f(Y_i)$ contains only variables in $\mathcal{S}$. If it contains a variable, $A_r \notin \mathcal{S}$, then A_r must be a root variable of the network. Otherwise it would have been replaced by its parents according to the protocol. $A_r \notin \mathcal{S}$, so $A_r \notin \mathcal{S}_3$, hence A_r must be d-connected (given an empty set) to at least one X_j for some j. Since A_r is a root of the network, A_r must be the ancestor of X_j. We next discuss if it is X_j for $j = i$ or $j \neq i$.

- $j = i$, i.e., A_r is an ancestor of X_i. Since X is ASDC, X_i cannot be caused by a variable belonging to another unit. Hence, we have $r = i$. If all directed paths from A_r to Y_i pass through variables in $\mathcal{S}$, then A_r cannot be replaced into the r.h.s. of $SE_f(Y_i)$. Hence, there exists at least one directed path from A_r to Y_i that does not pass through any variable in $\mathcal{S}$, which we denote as p_d. Since A_r is an ancestor of X_i and A_r to Y_i is a directed path not through $\mathcal{S}$(including X_i), there exists a confounding path between X_i and Y_i through A_r. Since X_i and Y_i are not confounded by only variables of i, p_d must go through a variable of a different unit, and is the root of that confounding path. However, then $A_r \in \mathcal{S}_2$ by definition, which contradicts the assumption that $A_r \notin \mathcal{S}$.
- $j \neq i$, i.e., A_r is an ancestor of X_j for some $j \neq i$. Again, there exists at least one directed path from A_r to Y_i that does not pass through any variable in $\mathcal{S}$, which we denote as p_d. Since A_r is ancestor to both X_j and Y_i, there is a confounding path between X_j and Y_i through A_r. A_r is the root on this path, which implies $A_r \in \mathcal{S}_3$, and contradicts the assumption that $A_r \notin \mathcal{S}$.

Thus, our counterproof assumption is wrong, which means the r.h.s. of $SE_f(Y_i)$ generated by the above protocol contains only variables in $\mathcal{S}$. Next we prove that the coefficients C_{W_t} for each $W_t\mathcal{S}$ in the linear combination is equal to the sum of the values of the directed paths from W_t to Y_i that do not go through any variable in $\mathcal{S}$. In the protocol above, every time a variable is replaced by its parents, there is a multiplier equal to the directed edge between each parent and the variable. For example, in SE_{Y_i}, a term is γC_i. If C_i is replaced by its parents, D_j and E_k, where $C_i = \delta D_j + \theta E_k$, then the term in SE_{Y_i} becomes $\gamma(\delta D_j + \theta E_k)$. So the coefficient of D_j is C_i's coefficient γ multiplied by δ, the edge $D_j \to C_i$. Since replacements of a variable stops if it is in $\mathcal{S}$, we have that the final coefficient of a variable is equal to the sum of all directed paths from that variable to Y_i, which do not pass through any other variable in $\mathcal{S}$. □

Lemma 6. *Given n IID random variables $X_1, \dots, X_n$, and n IID random variables $R_1, \dots, R_n$. For each i, R_i is not independent of X_i only. Then we have*

$$E\left[\frac{(X_i - \bar{X})R_i}{\sum\limits_{1 \leq k \leq n} (X_k - \bar{X})^2}\right] = \frac{\beta_{RX}}{n},$$

and β_{RX} is the OLS regression coefficient of R on X, treating $X_1, \dots, X_n$ as a single variable X, and $R_1, \dots, R_n$ as a single variable R.

Proof. The above expression only depends on i, and from the property of IID, it is the same for any i. We sum over i for that expression, and get

$$\begin{aligned}
&nE\left[\frac{(X_i-\bar{X})R_i}{\sum\limits_{1\le k\le n}(X_k-\bar{X})^2}\right]\\
=&\sum_{1\le i\le n}E\left[\frac{(X_i-\bar{X})R_i}{\sum\limits_{1\le k\le n}(X_k-\bar{X})^2}\right]\\
=&E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})R_i}{\sum\limits_{1\le k\le n}(X_k-\bar{X})^2}\right]\\
=&E\left[\hat{\beta_{RX}}\right]\\
=&\beta_{RX}.
\end{aligned}$$

Divided by n on both sides, we have the equation in the lemma. $\square$

Lemma 7. *Given n IID random variables $X_1,\dots,X_n$, and n IID random variables $R_1,\dots,R_n$. For each i, R_i is not independent of X_i only. Then we have*

$$E\left[\frac{(X_i-\bar{X})R_j}{\sum\limits_{1\le k\le n}(X_k-\bar{X})^2}\right]=-\frac{\beta_{RX}}{n(n-1)},$$

for $i\neq j$, and β_{RX} is the OLS regression coefficient of R on X, treating $X_1,\dots,X_n$ as a single variable X, and $R_1,\dots,R_n$ as a single variable R.

Proof. Denote the expectation of interest as E_{ij}. X and R are both IID regarding different units, and X_i and R_j are independent for $i\neq j$. Thus, $E_{ij}=E_{i'j}$, for any $i'\neq j$. Below when the sum is over $i\neq j$, it means summing over $i\in\{1,\dots,n\}\setminus\{j\}$. We have

$$\begin{aligned}
(n-1)E_{ij}=&\sum_{i\neq j}E_{ij}\\
=&E\left[\frac{\sum\limits_{i\neq j}(X_i-\bar{X})R_j}{\sum\limits_{1\le k\le n}(X_k-\bar{X})^2}\right]\\
=&E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})R_j-(X_j-\bar{X})R_j}{\sum\limits_{1\le k\le n}(X_k-\bar{X})^2}\right]\\
=&E\left[\frac{(\sum\limits_{1\le i\le n}(X_i-\bar{X})-(X_j-\bar{X}))R_j}{\sum\limits_{1\le k\le n}(X_k-\bar{X})^2}\right]\\
=&E\left[\frac{(0-(X_j-\bar{X}))R_j}{\sum\limits_{1\le k\le n}(X_k-\bar{X})^2}\right]\\
=&-E\left[\frac{(X_j-\bar{X})R_j}{\sum\limits_{1\le k\le n}(X_k-\bar{X})^2}\right].
\end{aligned}$$

By Lemma 6, we have

$$(n-1)E_{ij} = -\frac{\beta_{RX}}{n}.$$

Divided by $(n-1)$ on both sides, we get the equation we wanted to prove. $\square$

Lemma 8. *Given n IID random variables $X_1, \ldots, X_n$, and a variable L_t independent of $X_1, \ldots, X_n$. Then we have*

$$E\left[\frac{(X_i - \bar{X})L_t}{\sum\limits_{1\le k\le n}(X_k - \bar{X})^2}\right] = 0.$$

Proof. Denote the expectation of interest as E_i, then $E_i = E_j$ for any i, j, since X_i and X_j are IID. So we have

$$\begin{aligned}
nE_i = &\sum_{1\le i\le n} E_i \\
=&E\left[\frac{\sum\limits_{1\le i\le n}(X_i - \bar{X})L_t}{\sum\limits_{1\le k\le n}(X_k - \bar{X})^2}\right] \\
=&0.
\end{aligned}$$

$\square$

To prove Theorem 1, we first prove a slightly different version of it, Lemma 9.

Lemma 9. *Given the interaction network G^* of a balanced linear interaction model, with X_i and Y_i not confounded by any variable in $\mathcal{V}_i$, $\forall i$. Given that X satisfies ASDC, then the expected value of the OLS estimator $\hat{\beta_{YX}}$ is given by*

$$\begin{aligned}
&E[\hat{\beta_{YX}}] \\
=&\alpha \\
&+\frac{1}{n}\Big(\sum_{p\in\mathcal{P}} Val(p) + \sum_{1\le i\le n}\sum_{R\in(\mathcal{R}[iji]\setminus\{X_i\})} c_R\beta_{RX}\Big) \\
&-\frac{1}{n(n-1)}\sum_{1\le i\le n}\sum_{R\in\mathcal{R}[ji]} c_R\beta_{RX},
\end{aligned}$$

where $\mathcal{P}$ is the set of directed paths from X_i to Y_i for all i through any $W_j \in \mathcal{V}_{(j)}$ with $i \neq j$, $\mathcal{R}[iji]$ is the set of roots of the open paths between X_i and Y_i through some W_j with $j \neq i$, $\mathcal{R}[ji]$ is the set of roots of the open paths between X_j and Y_i for $j \neq i$, and c_R is the sum of values of the directed paths from a variable R ($\in (\mathcal{R}[iji] \setminus \{X_i\})$ or $\in \mathcal{R}[ji]$) to Y_i not passing through any variable in $\mathcal{R}[iji] \cup \mathcal{R}[ji]$ for any $j \neq i$.

Proof.

$$
\begin{aligned}
E[\hat{\beta_{YX}}] =& E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right] \\
=& E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})Y_i}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right] - E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})\bar{Y}}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right] \\
=& E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})Y_i}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right] - E\left[\frac{(\sum\limits_{1\le i\le n}X_i-n\bar{X})\bar{Y}}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right] \\
=& E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})Y_i}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right] - E\left[\frac{(n\bar{X}-n\bar{X})\bar{Y}}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right] \\
=& E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})Y_i}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right]
\end{aligned}
$$

Y_i is can be written as a linear combination of the set in Lemma 5, $\mathcal{S}$. By Lemma 5, $\mathcal{S}$ is composed of

1. X_i,
2. the root variables (excluding X_i) of each open path between X_j and Y_i, and
3. the root variables of this interaction network that are in $Anc(Y_i)$ and d-separated (by an empty set) from X_j for all j, denoted by $\mathcal{L}_i$.

The second component can be further divided into two sub-components as follows.

1. $\mathcal{R}[iji] \setminus \{X_i\}$, the set of roots of the open paths between X_i and Y_i through some W_j with $j \neq i$, with X_i excluded, and
2. $\mathcal{R}[ji]$, the set of roots of the open paths between X_j and Y_i for $i \neq j$.

We have

$$Y_i = c_i X_i + \sum_{R\in(\mathcal{R}[iji]\setminus\{X_i\})} c_R R + \sum_{R\in\mathcal{R}[ji]} c_R R + \sum_{L\in\mathcal{L}_i} c_L L,$$

where c_i, c_R, and c_L denote coefficients for the linear combination. The variables in the above expression are $\mathcal{S}$, i.e., $\mathcal{S} = \mathcal{R}[iji] \cup \mathcal{R}[ji] \cup \mathcal{L}_i$. Next, we compute the coefficients c_i, c_R, c_L.

c_i is the sum of the directed path values from X_i to Y_i not passing through any variable in $\mathcal{S}$. There are three types of directed paths from X_i to Y_i:

1. the directed edge $X_i \to Y_i$,
2. directed paths $X_i \to \cdots \to V_i \to \cdots \to Y_i$, and
3. directed paths $X_i \to \cdots \to V_j \to \cdots \to Y_i$ for $j \neq i$.

The first two types belong to TACE by definition. So $c_i = \alpha + c_{i3}$, where c_{i3} is the coefficient contributed by the third type of directed paths. Note that V_j cannot be a root of another path between X_k and Y_l for some $k \neq l$. This is because V_j is caused by X_i, so V cannot be ASDC, so X cannot be ASDC since X_k is caused by V_j, which violates the assumption that X is ASDC. Hence, c_{i3} is equal to the sum of all directed paths from X_i to Y_i through some variable V_j for any j, which is equal to $\sum_{p\in\mathcal{P}}$ in the lemma statement.

For the second and third components in Y_i, each c_R is the sum of the directed paths (multiplications of edge coefficients) from R to Y_i not through variables in $\mathcal{S}$. This follows from Lemma 5.

We have

$$
\begin{aligned}
&E[\hat{\beta_{YX}}]\\
=&E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})Y_i}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right]\\
=&E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})(c_iX_i+\sum_{R\in(\mathcal{R}[iji]\setminus\{X_i\})}c_RR+\sum_{R\in\mathcal{R}[ji]}c_RR+\sum_{L\in\mathcal{L}_i}c_LL)}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right]\\
=&\alpha E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})X_i}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right]+E\left[\frac{\sum\limits_{1\le i\le n}(X_i-\bar{X})c_{i3}X_i}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right]\\
&+\sum_{1\le i\le n}\sum_{R\in(\mathcal{R}[iji]\setminus\{X_i\})}c_RE\left[\frac{(X_i-\bar{X})R}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right]\\
&+\sum_{1\le i\le n}\sum_{R\in\mathcal{R}[ji]}c_RE\left[\frac{(X_i-\bar{X})R}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right]+\sum_{1\le i\le n}\sum_{L\in\mathcal{L}_i}c_LE\left[\frac{(X_i-\bar{X})L}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right].
\end{aligned}
$$

For the first term: similar to the way $\bar{Y}$ is removed before, in the first term, we can change X_i to $X_i-\bar{X}$. The numerator and the denominator are the same in the expectation. So the first term is α.

The second term is equal to

$$\sum_{1\le i\le n}c_{i3}E\left[\frac{(X_i-\bar{X})X_i}{\sum\limits_{1\le i\le n}(X_i-\bar{X})^2}\right].$$

By Lemma 6, it becomes

$$\sum_{1\le i\le n}c_{i3}\frac{\beta_{XX}}{n},$$

where c_{i3} is the sum of directed paths from X_i to Y_i through V_j for any $j\neq i$ and any V.

For the third term: we look at one single R first. R is the root variable of an open path between X_i and Y_i through some W_j with $j\neq i$, so R causes X_i. Then R must belong to unit i since X satisfies ASDC. Since R is the root, $R\in Anc(X)$, so R satisfies ASDC, and is IID for different units. So we relabel this R as R_i, and we have IID $R_1,\dots,R_n$. Applying Lemma 6, we have the expectation term is equal to β_{RX}/n. c_R is the sum of the directed paths from R_i to Y_i, not through variables in $\mathcal{S}$. So the third term is equal to

$$\frac{1}{n}\sum_{1\le i\le n}\sum_{R\in(\mathcal{R}[iji]\setminus\{X_i\})}c_R\beta_{RX}.$$

For the fourth term: we look at one single R first. R is the root variable of an open path between X_j and Y_i, for some $j\neq i$, so either R causes X_j or $R=X_j$. If R causes X_j, then R must belong to unit j, because X satisfies ASDC. So either case R belongs to unit j. Since R is the root, $R\in Anc(X)$, so R satisfies ASDC, and is IID for different units. So we relabel this R as R_j, and we have IID $R_1,\dots,R_n$. Applying Lemma 7, we have the expectation term is equal to $-\beta_{RX}/(n(n-1))$. c_R is the sum of the directed paths from R_j to Y_i, not through variables in $\mathcal{S}$. So the fourth term is equal to

$$-\frac{1}{n(n-1)}\sum_{1\le i\le n}\sum_{R\in\mathcal{R}[ji]}c_R\beta_{RX}.$$

The fifth term is 0 by Lemma 4.

Finally, recall that $Val(p)$ denotes the value of an open path p. Plugging the above values back into the expression for $E[\hat{\beta_{YX}}]$, we have the results as in Lemma 9. $\square$