FIrst version: Mar 2019
Last version: Apr 2020

Generalizability in Causal Inference: Theory and Algorithms
Errata

Elias Bareinboim

March 2019

The following is a revision of some results presented in [1]. Section, definition, theorem and lemma numbers match the manuscript in mention. Changes are highlighted in red. I thank Professors Juan D. Correa and Sanghack Lee for raising some of the issues discussed here and helping to fix them. An updated and more general version of some of these results appeared in [2, 3].

4 Transportability from Multiple Studies with Limited Experiments

4.1 Characterizing mz-Transportable Relations

The following is a revised definition of mz^*-shedge, a graphical structure that witnesses the non-transportability of a causal distribution. The removal of condition 3 of the original definition is not strictly needed but since it's entailed by conditions 1 and 2, we prefer to phrase in this way for the sake of clarity.

Definition 18 (mz^*-shedge)**.** *Let $\mathcal{D} = (D^{(1)}, \ldots, D^{(n)})$ be a collection of selection diagrams relative to source domains $\Pi = (\pi_1, \ldots, \pi_n)$ and target domain π^*, respectively, $\mathbf{S_i}$ represents the collection of S-variables in the selection diagram $D^{(i)}$, and let $D^{(*)}$ be the causal diagram of π^*. Let $\{\langle P^i, I^i_z \rangle\}$ be the collection of pairs of observational and interventional distributions of $\{\pi_i\}$, where $I^i_z = \bigcup_{\mathbf{Z}' \subseteq \mathbf{Z}_i} P^i(\mathbf{v}|do(\mathbf{z}'))$, and in an analogous manner, $\langle P^*, I^*_z \rangle$ be the observational and interventional distributions of π^*, for $\mathbf{Z_i}$ the set of experimental variables in π_i. Consider a pair of $\mathbf{R}$-rooted C-~~forests~~components $\mathcal{F} = \langle F, F' \rangle$ such that $F' \subset F$, $F' \cap \mathbf{X} = \emptyset$, $F \cap \mathbf{X} \neq \emptyset$, and $\mathbf{R} \subseteq An(\mathbf{Y})_{G_{\overline{\mathbf{X}}}}$ ~~(called hedge)~~. We say that ~~the induced~~ a collection of pairs of $\mathbf{R}$-rooted C-forests over each diagram, $\langle \mathcal{F}^{(*)}, \mathcal{F}^{(1)}, \ldots, \mathcal{F}^{(n)} \rangle$, with $\mathcal{F}^{(i)} = \langle F^{(i)}, F'^{(i)} \rangle$, $F^{(i)} \subseteq F$, $i = \{*, 1, \ldots, n\}$, $\bigcup_i F'^{(i)} = F'$, is an mz^*-shedge for $P^*_{\mathbf{x}}(\mathbf{y})$ relative to experiments $(I^*_z, I^1_z, \ldots, I^n_z)$ if they are all hedges for $P_{\mathbf{x}}(\mathbf{y})$, and one of the following conditions hold for each domain π_i, $i = \{*, 1, \ldots, n\}$:*

1. *There exists at least one variable of* $\mathbf{S_i}$ *pointing to the induced diagram* $F'^{(i)}$*, or*

2. $(F^{(i)} \setminus F'^{(i)}) \cap \mathbf{Z_i}$ *is an empty set.*

3. *The collection of pairs of C-forests induced over diagrams,* $\langle \mathcal{F}^{(*)}, \mathcal{F}^{(1)}, \ldots, F^{(i)} \setminus \mathbf{Z_i^*}, \ldots, \mathcal{F}^{(n)} \rangle$*, is also an mz-shedge relative to* $(I_z^*, I_z^1, \ldots, I^i_{z \setminus z_i^*}, \ldots, I_z^n)$*, where* $\mathbf{Z_i^*} = (F^{(i)} \setminus F'^{(i)}) \cap \mathbf{Z_i}$*.*

We call mz^**-shedge the mz-shedge in which there exist one directed path from* $\mathbf{R} \setminus (\mathbf{R} \cap De(\mathbf{X})_F)$ *to* $(\mathbf{R} \cap De(\mathbf{X})_F)$ *not passing through* $\mathbf{X}$*.*

With a revised definition, we provide a new proof for Theorem 17 and related Lemmas 17, 18 and 19.

Theorem 17. *Let* $\mathcal{D} = \{D^{(1)}, \ldots, D^{(n)}\}$ *be a collection of selection diagrams relative to source domains* $\Pi = \{\pi_1, \ldots, \pi_n\}$*, and target domain* π^**, respectively, and* $\{I_z^i\}$*, for* $i = \{*, 1, \ldots, n\}$ *defined appropriately. If there is an* mz^**-shedge for the effect* $R = P^*_{\mathbf{x}}(\mathbf{y})$ *relative to experiments* $(I_z^*, I_z^1, \ldots, I_z^n)$ *in* $\mathcal{D}$*,* R *is not mz-transportable from* Π *to* π^* *in* $\mathcal{D}$ *(relative to all experiments* I_z^i*).*

Proof sketch. Let F be the $\mathbf{R}$-rooted C-component (basis). Without loss of generality, we will consider a structure with a maximal root-set. That is, one that when subjected to the following procedure remains unchanged:

1. let $\mathbf{B} = An(\mathbf{Y})_{G_{\overline{\mathbf{X}}}} \cap (F \setminus \mathbf{X})$,

2. consider the subgraph $F \setminus \mathbf{X}$ and let $\mathbf{R}'$ be the set of variables in $\mathbf{B}$ that are also in the same C-component as any element of $\mathbf{R}$ in that subgraph.

3. Then, remove from $\mathcal{F}$ the edges outgoing from $\mathbf{R}'$ and let $\mathbf{R} = \mathbf{R} \cup \mathbf{R}'$.

After the previous steps, we obtain a new mz^*-shedge with a maximal root-set, where the variables in F' are exactly those in the root-set $\mathbf{R}$. To witness, assume for the sake of contradiction there exists a variable V in F' not in $\mathbf{R}$, by definition F' is an $\mathbf{R}$-rooted C-component containing no variables in $\mathbf{X}$, and since V belongs to F' it must fall into $\mathbf{B}$ in step one and also satisfied step two. Hence, V can be added to the root-set as in step three, contradicting the fact that F had maximal root-set.

Let $\mathbf{T} = \mathbf{F} \setminus \mathbf{R}$ be the observable variables in F that are not in $\mathbf{R}$. Let $\mathbf{U}'$ be the set of unobservable variables in F and partition it into the sets:

- $\mathbf{U_T} = \{U \in \mathbf{U}' \mid T_1 \leftarrow U \rightarrow T_2 \text{ and } T_1, T_2 \in \mathbf{T}\}$,
- $\mathbf{U_R} = \{U \in \mathbf{U}' \mid R_1 \leftarrow U \rightarrow R_2 \text{ and } R_1, R_2 \in \mathbf{R}\}$, and
- $\mathbf{U}_\times = \{U \in \mathbf{U}' \mid T \leftarrow U \rightarrow R \text{ and } T \in \mathbf{T}, R \in \mathbf{R}\}$.

Let $\mathbf{U^i_T} = \mathbf{U_T} \cap F^{(i)}, \mathbf{U^i_R} = \mathbf{U_R} \cap F^{(i)}$ and $\mathbf{U}^i_\times = \mathbf{U}_\times \cap F^{(i)}$.

We construct two causal models M_1 and M_2 that will agree on the collection of distributions $\{\langle P^i, I_z^i \rangle\}$, $\langle P^*, I_z^* \rangle$, but disagree on the interventional distribution $P^*_{\mathbf{x}}(\mathbf{y})$.

Let k_t be the number of $\mathcal{F}^{(i)}$s in which a variable $T \in \mathbf{T}$ appears. Then, we will parametrize T as a k_t-bit variable with $T_{[i]}$ representing the bit in T corresponding to $\mathcal{F}^{(i)}$. Similarly, define k_u for $U \in \mathbf{U_T} \cup \mathbf{U}_\times$, then U is a k_u-bit variable where $U_{[i]}$ stands for the bit associated with $\mathcal{F}^{(i)}$.

Call $\mathbf{W}$ the set of variables pointed by S-nodes in F' and consider the following encoding for the domains: let S_i be the index variable corresponding to the source domain $\pi_i \in \Pi$, and let the tuple $\langle S_1 = 0, \ldots, S_i = 1, \ldots, S_n = 0\rangle$ represent the index for the functional model relative to this domain. Let the tuple $\langle S_1 = 0, S_2 = 0, \ldots, S_n = 0\rangle$ represent the index for functional model relative to the target domain π^*.

Let $\mathbf{Pa}_v$ stand for the set of observable and unobservable parents of variable V in F and $\mathbf{Pa}_v^i$ for the set of parents of the same variable in $F^{(i)}$. For a set of variables $\mathbf{V}$, let $\mathbf{Pa_v} = \bigcup_{V\in\mathbf{V}} \mathbf{Pa}_v$ and $\mathbf{Pa_v}^i = \bigcup_{V\in\mathbf{V}} \mathbf{Pa}_v^i$.

In both models, let each bit $T_{[i]}$ of $T \in \mathbf{T}$ be governed by the function

$$f_{t_{[i]}} = \bigoplus_{A\in\mathbf{Pa}_t^i} A_{[i]}. \tag{1}$$

Variables in $\mathbf{R} \cup \mathbf{U_R}$ are binary. Pick an arbitrary variable $R^* \in \mathbf{R}$. For any $R \in \mathbf{R} \setminus \mathbf{W}$ in model 1 and 2 except for R^* in model 2, let

$$f_r = \left(\bigwedge_{\pi_i\in\Pi, T\in\mathbf{Pa}_r^i\cap\mathbf{T}} g_i(T) \quad \wedge \bigwedge_{\pi_i\in\Pi, U\in\mathbf{Pa}_r^i\cap\mathbf{U}_\times} U_{[i]} \right) \wedge \left(\bigoplus_{U\in\mathbf{Pa}_r\cap\mathbf{U_R}} U \right); \tag{2}$$

where $g_i(.)$ is defined as follows:

$$g_i(T) = \begin{cases} \overline{T_{[i]}} & \text{if } |\mathbf{Pa_r}^i \cap \mathbf{T}| \text{ is odd and } |\mathbf{U}_\times^i| \text{ is odd, or} \\ & \text{if } |\mathbf{Pa_r}^i \cap \mathbf{T}| \text{ is even and } |\mathbf{U}_\times^i| \text{ is even and } T = T^{(i)}. \\ T_{[i]} & \text{otherwise.} \end{cases} \tag{3}$$

Where $T^{(i)}$ is any variable chosen from the set $\mathbf{Pa_r}^i \cap \mathbf{T}$ for each domain π_i.

For R^* in model 2:

$$f_{r^*} = \left(\bigwedge_{T\in\mathbf{Pa}_{r^*}^i\cap\mathbf{T}} g_i(T) \quad \wedge \bigwedge_{U\in\mathbf{Pa}_{r^*}^i\cap\mathbf{U}_\times} U_{[i]} \right) \wedge \overline{\left(\bigoplus_{U\in\mathbf{Pa}_{r^*}\cap\mathbf{U_R}} U \right)}. \tag{4}$$

For $R \in (\mathbf{R} \cap \mathbf{W})$ let

$$R \leftarrow f_r \wedge \bigwedge_{S_i|(S_i\to R)\in\mathcal{D}} \overline{S_i}, \tag{5}$$

where f_r is constructed as in the previous case and S_i is an S-node pointing to R, relative to domain π_i.

Every bit of the $\mathbf{U}$-variables is set to behave as a fair coin.

Lemma 17. *The two models M_1 and M_2 are compatible with the selection diagrams $\mathcal{D}$.*

Proof. The result is immediate. Consider the functional model that generates any domain π_i, in both models M_1 and M_2. By construction, the index tuple is set to $\langle S_1 = 0, \ldots, S_i = 1, \ldots, S_n = 0\rangle$ in π_i, and $\langle S_1 = 0, \ldots, S_i = 0, \ldots, S_n = 0\rangle$ in π^*. So, it is obvious that in both models, the only structural differences between π_i and π^* are the equations of $W \in \mathbf{W}$ in which S_i appears. □

Lemma 18. *The two models agree in the distribution of $P^i(\mathbf{t}, \mathbf{r}), i = \{*, 1, \ldots, n\}$ and there exists an assignment for $\mathbf{X}$ and $\mathbf{Y}$ such that $P^*_{M_1}(\mathbf{Y}|do(\mathbf{X})) \neq P^*_{M_2}(\mathbf{Y}|do(\mathbf{X}))$.*

Proof. (*Matching observational distributions*)
First consider any particular domain π_i and a particular assignment $\mathbf{u}$ of the variables in $\mathbf{U}$. We have that in both models the value of $\mathbf{T}$ has to be the same since the functions are the same in those models (with fixed π_i all S_i have the same value).

Let $\mathbf{R}_0$ be the set of nodes in $\mathbf{R}$ for which the expression in the first parenthesis of equation (2) evaluates to 0 in both models. Note that the set $\mathbf{R}_0$ is determined by the variables in $\mathbf{U}_\mathbf{T} \cup \mathbf{U}_\times$, because those determine the values of $\mathbf{T}$, and the variables in $\mathbf{U}_\mathbf{R}$ only appear in the second part of equation (2) which is not taken into account in the definition of $\mathbf{R}_0$. We will show that $\mathbf{R}_0$ is not empty in the context of the non-intervened models corresponding to π_i. Consider any $U \in \mathbf{U}^i_\times$ such that $U_{[i]} = 0$, then any R that is pointed by U will have value 0 in both models due to the construction of f_r, and we are done. We continue with the situation where all such U have $U_{[i]} = 1$. Consider the quantity C_i defined as

$$C_i = \bigoplus_{T \in \mathbf{Pa}^i_\mathbf{r} \cap \mathbf{T}} T_{[i]}, \tag{6}$$

and note that due to the forestness of $F^{(i)}$ and the parametrization; C_i computes the xor of all the unobservable variables in $\mathbf{U}_\mathbf{T}$ and $\mathbf{U}_\times$, having those in $\mathbf{U}_\mathbf{T}$ accounted twice. Together with the fact that for any $U \in \mathbf{U}^i_\times$, $U_{[i]} = 1$, it follows that

$$C_i = \bigoplus_{U \in \mathbf{U}^i_\times} U_{[i]} = |\mathbf{U}^i_\times| \ mod\ 2. \tag{7}$$

Note that the set of parents of variables in $\mathbf{R}$ in $\mathbf{T}$ (i.e. $\mathbf{Pa}^i_\mathbf{r} \cap \mathbf{T}$) must be non-empty for any given hedge, then consider each one of the following four scenarios:

1. $|\mathbf{Pa}^i_\mathbf{r} \cap \mathbf{T}|$ is odd and $|\mathbf{U}^i_\times|$ is odd: We have $C_i = 1$ which implies that at least one of $T_{[i]}$ has to be 1. Since g_i negates all $T_{[i]}$ in this case, we have that at least one R (with T as a parent) will have 0 as value.

2. $|\mathbf{Pa}^i_\mathbf{r} \cap \mathbf{T}|$ is odd and $|\mathbf{U}^i_\times|$ is even: We have $C_i = 0$ which implies that at least one of $T_{[i]}$ has to be 0. Since g_i leaves each $T_{[i]}$ the same, we have that at least one R will have 0 as value.

3. $|\mathbf{Pa}^i_\mathbf{r} \cap \mathbf{T}|$ is even and $|\mathbf{U}^i_\times|$ is odd: We have $C_i = 1$ which implies that at least one of $T_{[i]}$ has to be 0. As in the previous case g_i leaves $T_{[i]}$ the same so at least one $R = 0$ in both models.

4. $|\mathbf{Pa}^i_\mathbf{r} \cap \mathbf{T}|$ is even and $|\mathbf{U}^i_\times|$ is even: We have $C_i = 0$, so for all combinations but for all $T_{[i]} = 1$, there are always at least two $T_{[i]} = 0$. Since g_i negates only one $T_{[i]}$, it follows that there is always at least one $g_i(T) = 0$ and any R pointed by T will have value 0.

Due to the previous analysis we have that $\mathbf{R}_0$ is non-empty. Pick $\hat{R} \in \mathbf{R}_0$ that is closest to R^* in terms of the length of the bidirected path $\overline{p}$ made of edges in $\mathbf{U_R}$ between them (the length of the path is 0 if $R^* \in \mathbf{R}_0$. Make $\mathbf{u}^1 = \mathbf{u}$ (the considered assignment) and $\mathbf{u}^2$ equal to $\mathbf{u}$ for all $\mathbf{U}$ except those in $\overline{p}$ for which their negation is taken. By definition $\overline{p}$ intersects with $\mathbf{R}_0$ only at the endpoints. Also, for every intermediate node R of $\overline{p}$, there are two parents in $\mathbf{U_R}$ being negated; from the parametrization of f_r we can tell that the value of R remains the same because this change does not affect the parity being computed by the xor. We have then that $\mathbf{u}^1$ corresponds to $\mathbf{u}^2$ and repeating the reasoning for any other assignment of the $\mathbf{u}$ we get a bijective relationship between assignments producing the same observation in both models, hence the distributions over the observed variables is the same.

(*Different interventional distribution*)
For the second part of the claim, consider the distribution $P(\mathbf{r}|do(\mathbf{X} = \hat{\mathbf{x}}))$, where $\hat{\mathbf{x}}$ is an assignment where each bit of $X \in \mathbf{X}$ is given by

$$\hat{x}_{[i]} = \begin{cases} 0 & \text{if } X \in F^{(i)} \text{ and } X \notin \mathbf{Pa}^i_\mathbf{r}, \\ g_i(1) & \text{if } X \in \mathbf{Pa}^i_\mathbf{r}, \end{cases} \tag{8}$$

Start by noting when this intervention on $\mathbf{X}$ is performed, every $F^{(i)}$ is affected (because by definition every $\mathcal{F}^{(i)}$ intersects $\mathbf{X}$). We want to show that under this circumstance, there exists at least one assignment $\mathbf{u}$ such that $\mathbf{R}_0$ is empty. Start with an assignment where $\mathbf{u}_\times = 1$, if every $g_i(T) = 1$ for $i = \{*, 1, \ldots, n\}$, $T \in \mathbf{Pa}^i_\mathbf{r}$ we are done. Otherwise, for every i, T such that $g_i(T) = 0$ find a path $\overline{p}$, in $F^{(i)}$, between T and a variable in $An(\mathbf{X})_{F^{(i)}}$ (that includes $\mathbf{X}$) made of bidirected edges corresponding to variables in $\mathbf{U_T}$. Such path must exists due to the fact that the mz^*-shedge under consideration has a maximal root-set and T is in $An(\mathbf{Y})_{G_{\overline{\mathbf{X}}}}$ (because it is a parent of some $R \in \mathbf{R}$).

We can flip the bit associated with π_i for all $\mathbf{u}$s in $\overline{p}$, which preserve the parity (hence the bit value) of every intermediate observable, while the value of the observable in the endpoint is either fixed by intervention (if the path ends in some $X \in \mathbf{X}$) or can change without affecting any variable in $\mathbf{Pa}^i_\mathbf{r} \setminus \{T\}$ (and hence R as well) because it is an ancestor of $\mathbf{X}$ that has been intervened and $F^{(i)}$ is a forest. Changing the unobservables in the path also changes the parity of T and since $g_i(\overline{T}) = \overline{g_i(T)}$ we have that now $g_i(T) = 1$.

This process only affects bits associated with π_i, by repeating it for every other T, i such that $g_i(T) = 0$, we get an assignment where $\mathbf{R}_0$ is empty. Under these circumstances, the value of variables in $\mathbf{R}$ is determined by the xor in the second parenthesis of equation (2) that depends only on variables in $\mathbf{U_R}$, that free so far. Then, in M_1 we have

$$\bigoplus_{R\in\mathbf{R}} R = \bigoplus_{R\in\mathbf{R}} \bigoplus_{U\in\mathbf{Pa}_r\cap\mathbf{U_R}} U = \bigoplus_{U\in\mathbf{U_R}} (U \oplus U) = 0, \tag{9}$$

since every $U \in \mathbf{U_R}$ appears exactly twice. As for M_2

$$\bigoplus_{R\in\mathbf{R}} R = \left(\bigoplus_{R\in\mathbf{R}, R\neq R^*} \bigoplus_{U\in\mathbf{Pa}_r\cap\mathbf{U_R}} U\right) \oplus \overline{\left(\bigoplus_{U\in\mathbf{Pa}_{r^*}\cap\mathbf{U_R}} U\right)} \tag{10}$$

$$= \left(\bigoplus_{R\in\mathbf{R}, R\neq R^*} \bigoplus_{U\in\mathbf{Pa}_r\cap\mathbf{U_R}} U\right) \oplus \left(1 \oplus \bigoplus_{U\in\mathbf{Pa}_{r^*}\cap\mathbf{U_R}} U\right) \tag{11}$$

$$= 1 \oplus \bigoplus_{R\in\mathbf{R}} \bigoplus_{U\in\mathbf{Pa}_r\cap\mathbf{U_R}} U \tag{12}$$

$$= 1 \oplus \left(\bigoplus_{U\in\mathbf{U_R}} (U \oplus U)\right) \tag{13}$$

$$= 1. \tag{14}$$

Then, from the first part of this proof we have that for any $\mathbf{u}$ for which $\mathbf{R}_0$ the distributions both models produce the same observations, however, for the intervention $do(\mathbf{X} = \hat{\mathbf{x}})$ there are $\mathbf{u}$ for which $\mathbf{R}_0$ is empty and we have that model 2 produces more observations where $\bigoplus \mathbf{R} = 1$ hence the different observations which implies $P^*_{M_1}(\bigoplus \mathbf{r} = 1 | do(\hat{\mathbf{x}})) \neq P^*_{M_2}(\bigoplus \mathbf{r} = 1 | do(\hat{\mathbf{x}}))$.

(Mapping $\mathbf{R}$ to $\mathbf{Y}$)
By definition, there is a directed path in G from every $R \in \mathbf{R}$ to $\mathbf{Y}$ (could be zero-length) not intersecting $\mathbf{X}$. Augment M_1 and M_2 such that for any non-zero-length path $\overline{q}$ from R to $Y \in \mathbf{Y}$, each variable except for R let the function be an xor of its parents. If the path contains an intermediate variable $R' \in \mathbf{R}$ in $\overline{q}$ add an extra bit to it, such that the original bit computes the original function and the new one the xor of its parents.

In this new models $\bigoplus \mathbf{Y} = \bigoplus \mathbf{R}$, then the second part of the lemma follows.

Lemma 19. *The two models agree in the collection of interventional distributions ($\{I^i_z\}$) in the respective source domains π_i, $i = 1, \ldots, n$, and target domain π^*.*

Proof. Consider a domain π_i and a set $\mathbf{Z} \subseteq \mathbf{Z_i}$. From the definition of mz^*-shedge we have that either condition 1 or 2 are true for $F^{(i)}$. In the former case we have some indicator in $\mathbf{S_i}$ pointing to a variable $R \in \mathbf{R}$ that will be set to 0 in both models in domain π_i. In the latter case, and by the same argument used in the proof for lemma 18, we have that any variable $R \in \mathbf{R}$ that is a child of a variable $T \in \mathbf{T} \cap F^{(i)}$ belongs to $\mathbf{R}_0$.

In any case $\mathbf{R}_0$ is not empty and as in the proof for lemma 18 this implies that the observed distributions match in both models for the variables in F. The mapping described in lemma 18 modifies both models in the same way and may only change the functions of $\mathbf{R}$ by adding extra bits without changing the fact that the observations match between distributions. □

6 Causal Inference by Surrogate Experiments

6.1 Characterizing *zID* Relations

Theorem 31. *Let* $\mathbf{X}$*,* $\mathbf{Y}$*,* $\mathbf{Z}$ *be disjoint sets of variables and let* G *be the causal diagram. The causal effect* $Q = P(\mathbf{y}|do(\mathbf{x}))$ *is zID in* G *if and only if one of the following conditions hold:*

a. *Q is identifiable in G; or,*

b. *There is no hedge $\mathcal{F} = \langle F, F' \rangle$ for Q in G such that $(F \setminus F') \cap \mathbf{Z}$ is empty.*

(i) ~~$\mathbf{X}$ *intercepts all directed paths from* $\mathbf{Z}'$ *to* $\mathbf{Y}$*, and*~~

(ii) ~~*Q is identifiable in* $G_{\overline{\mathbf{Z}'}}$.[6]~~

Proof. (only if) Suppose there exists a hedge as described in condition b, Suppose Q is not identifiable in G (condition a) and there is a hedge $\mathcal{F}$ as described in condition b. Note that $\mathcal{F}$ satisfies the definition for mz^*-shedge, hence by Theorem 17 it follows that Q is not identifiable from $P(\mathbf{v})$, $\{P_{\mathbf{z}'}(\mathbf{v}|do(\mathbf{z}'))\}_{\mathbf{Z}' \subseteq \mathbf{Z}}$, which equates to Q not being *zID*.

(if) Suppose Q is not *zID*, then it easy to see that Q is not identifiable from $P(\mathbf{v})$ (which is considered by *zID*) therefore condition a is not satisfied. Let $\mathcal{F} = \langle F, F' \rangle$ be the hedge in G witnessing that some factor Q' associated with Q is not identifiable from $P(\mathbf{v})$. Let $\mathbf{Z}' = (F \setminus F') \cap \mathbf{Z}$, if $\mathbf{Z}' = \emptyset$, condition b does not hold and we are done. Otherwise, we can consider the distribution $P(\mathbf{v}|do(\mathbf{z}'))$ associated with $G_{\overline{\mathbf{z}'}}$ where $\mathcal{F}$ cannot be a hedge (every variable in $\mathbf{Z}'$ belongs to a different C-component in that graph). Then, Q' is identifiable from $P(\mathbf{v}|do(\mathbf{z}'))$ and there has to be another Q'' that is not identifiable from $P(\mathbf{v})$ else Q is *zID*. Let $\mathcal{F}'$ be the hedge associated with Q'' and by repeating the reasoning above, we have that either we end up with a hedge as forbidden by condition b or a contradiction. Therefore, Q being not *zID* implies that both conditions a and b are false; which entails the forward direction of this theorem. □

The corollary below followed from the original condition in Theorem 31.

[3]This condition can be rephrased graphically as "There exists no hedge for Q as an edge subgraph in $G_{\overline{\mathbf{Z}'}}$."

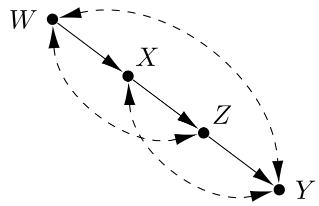

Figure 1: Graph in which $P(y|do(x))$ is not ID from $P(\mathbf{v})$ and G, but it is $z\mathit{ID}$ with experiments on Z, that is in $De(X)_{G_{An(Y)}}$.

Corollary 22. *Let G be the causal diagram, $\mathbf{X}$, $\mathbf{Y} \subset \mathbf{V}$ be disjoint sets of variables, and $\mathbf{Z} \subseteq De(\mathbf{X})_{G_{An(\mathbf{Y})}}$. The causal effect $Q = P(\mathbf{y}|do(\mathbf{x}))$ is not $z\mathit{ID}$ from P and $do(\mathbf{Z})$ in G, if Q is not ID from P in G.*

This corollary is not valid. To understand the subtlety with this statement, consider the graph in Fig. 1 where the query to be z-identified is $Q = P(y|do(x))$ and the available distributions are $P(\mathbf{v})$ and $P(y, x, w|do(z))$. According with the corollary, if Q is not identifiable from G and $P(\mathbf{v})$, it would not be identifiable even with experiments on Z because $\{Z\} \subseteq De(X)_{G_{An(Y)}}$. However, the following derivation follows:

$$P(y|do(x)) = \sum_z P(y|do(x), z)P(z|do(x)) \tag{15}$$

$$= \sum_z P(y|do(x), do(z))P(z|do(x)) \tag{16}$$

$$= \sum_z P(y|do(z))P(z|do(x)) \tag{17}$$

$$= \sum_z P(y|do(z)) \sum_w P(z|do(x), w)P(w|do(x)) \tag{18}$$

$$= \sum_z P(y|do(z)) \sum_w P(z|x, w)P(w|do(x)) \tag{19}$$

$$= \sum_z P(y|do(z)) \sum_w P(z|x, w)P(w), \tag{20}$$

that certifies that Q is $z\mathit{ID}$. The key point missed in Corollary 22 is that if Q is decomposable into more than one factor, some of them could be identified from the observational distribution and others from experimental distributions, fact that cannot be captured in a non-recursive condition.

6.2 A Complete Algorithm for $z\mathit{ID}$

Below the algorithm $\mathbf{ID^z}$ is reestated to make some recursive calls more explicit.

function $\mathbf{ID^z}(\mathbf{y}, \mathbf{x}, \mathbf{Z}, \mathcal{I}, \mathcal{J}, P, G)$

INPUT: $\mathbf{x}, \mathbf{y}$: value assignments; $\mathbf{Z}$: variables with interventions available; $\mathcal{I}, \mathcal{J}$: see caption; P: current probability distribution $do(\mathcal{I}, \mathcal{J}, x)$ (observational when $\mathcal{I} = \mathcal{J} = \emptyset$); G: causal graph.

OUTPUT: Expression for $P_{\mathbf{x}}(\mathbf{y})$ in terms of $P, P_{\mathbf{z}}$ or $\mathbf{FAIL}(F, F')$.

1 **if** $\mathbf{x} = \emptyset$, **return** $\sum_{\mathbf{v} \setminus \mathbf{y}} P(\mathbf{v})$.

2 **if** $\mathbf{V} \setminus An(\mathbf{Y})_G \neq \emptyset$,
  return $\mathbf{ID^z}(\mathbf{y}, \mathbf{x} \cap An(\mathbf{Y})_G, \mathbf{Z},$
  $\mathcal{I}, \mathcal{J}, \sum_{\mathbf{v} \setminus An(\mathbf{Y})_G} P, An(\mathbf{Y})_G)$.

3 Set $\mathbf{Z_w} = ((\mathbf{V} \setminus (\mathbf{X} \cup \mathcal{I} \cup \mathcal{J})) \setminus An(\mathbf{Y})_{G_{\overline{\mathbf{X} \cup \mathcal{I} \cup \mathcal{J}}}}) \cap \mathbf{Z}$.
  Set $\mathbf{W} = ((\mathbf{V} \setminus (\mathbf{X} \cup \mathcal{I} \cup \mathcal{J})) \setminus An(\mathbf{Y})_{G_{\overline{\mathbf{X} \cup \mathcal{I} \cup \mathcal{J}}}}) \setminus \mathbf{Z}$.
  if $(\mathbf{Z_w} \cup \mathbf{W}) \neq \emptyset$,
  return $\mathbf{ID^z}(\mathbf{y}, \mathbf{x} \cup \mathbf{w}, \mathbf{Z} \setminus \mathbf{Z_w}, \mathcal{I} \cup \mathbf{z_w}, \mathcal{J}, P_{\mathcal{I}, \mathbf{z_w}, \mathcal{J}}, G \setminus \mathbf{Z_w})$.

4 **if** $\mathcal{C}(G \setminus (\mathbf{X} \cup \mathcal{I} \cup \mathcal{J})) = \{S_0, S_1, ..., S_k\}$,
  return $\sum_{\mathbf{v} \setminus \{\mathbf{y}, \mathbf{x}, \mathcal{I}\}} \prod_i \mathbf{ID^z}(s_i, (\mathbf{v} \setminus s_i) \setminus \mathbf{Z},$
  $\mathbf{Z} \setminus (\mathbf{V} \setminus S_i), \mathcal{I}, \mathcal{J} \cup (\mathbf{Z} \cap (\mathbf{v} \setminus \mathbf{s_i})), P_{\mathcal{I}, \mathcal{J}, \mathbf{Z} \cap (\mathbf{V} \setminus \mathbf{S_i})}, G \setminus (\mathbf{Z} \cap (\mathbf{V} \setminus \mathbf{S_i})))$.
  if $\mathcal{C}(G \setminus (\mathbf{X} \cup \mathcal{I} \cup \mathcal{J})) = \{S\}$,

5 **if** $\mathcal{C}(G) = \{G\}$, $\mathbf{FAIL}(G, S)$.

6 **if** $S \in \mathcal{C}(G)$,
  return $\sum_{s \setminus \mathbf{y}} \prod_{i | V_i \in S} P(v_i | v_G^{(i-1)} \setminus (\mathcal{I} \cup \mathcal{J}))$.

7 **if** $(\exists S') S \subset S' \in \mathcal{C}(G)$,
  return $\mathbf{ID^z}(\mathbf{y}, \mathbf{x} \cap S', \mathbf{Z}, \mathcal{I}, \mathcal{J},$
  $\prod_{i | V_i \in C'} P(V_i | V_G^{(i-1)} \cap S', v_G^{(i-1)} \setminus (S' \cup \mathcal{I} \cup \mathcal{J})), S')$.

Figure 2: $\mathbf{ID^z}$: Algorithm capable of recognizing zID; The variables $\mathcal{I}, \mathcal{J}$ represent indices for currently active Z-interventions introduced respectively by steps 3 or 4. Note that P is sensitive to current instantiations of $\mathcal{I}, \mathcal{J}$.

References

[1] Elias Bareinboim. *Generalizability in Causal Inference: Theory and Algorithms*. PhD thesis, UCLA, 2014.

[2] S. Lee, J. Correa, and E. Bareinboim. General identifiability with arbitrary surrogate experiments. In *Proceedings of the 35th Conference on Uncertainty in Artificial Intelligence*, Tel Aviv, Israel, 2019. AUAI Press.

[3] S. Lee, J. Correa, and E. Bareinboim. Generalized transportability: Synthesis of experiments from heterogeneous domains. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, New York, NY, 2020. AAAI Press.