A. Proofs from Section 3

A.1. Proofs from Section 3.2

The following lemma formalizes the claim that $\mathcal{M}_{\mathcal{I}}(\mathcal{G})$ as given in Definition 3.3 contains exactly the sets of interventional distributions that can be generated from a causal model with DAG $\mathcal{G}$ by intervening on $\mathcal{I}$.

Lemma A.1. *$\{f^{(I)}\}_{I\in\mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G})$ if and only if there exists $f^{(\emptyset)} \in \mathcal{M}(\mathcal{G})$ such that $\forall I \in \mathcal{I},\ f^{(I)}$ factorizes according to Equation (1) in Definition 3.2.*

Proof. Suppose there exists $f^{(\emptyset)} \in \mathcal{M}(\mathcal{G})$ such that $\forall I \in \mathcal{I},\ f^{(I)}$ factorizes according to Equation (1) in Definition 3.2. Then $f^{(I)} \in \mathcal{M}(\mathcal{G})$ is trivially satisfied for all $I \in \mathcal{I}$. Also, we have $f^{(I)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)}) = f^{(\emptyset)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)})\ \forall j \notin I$ and $I \in \mathcal{I}$. It follows that $f^{(\emptyset)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)}) = f^{(I)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)}) = f^{(J)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)}),\ \forall j \notin I \cup J$ and all $I, J \in \mathcal{I}$. Therefore, $\{f^{(I)}\}_{I\in\mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G})$.

Conversely, suppose $\{f^{(I)}\}_{I\in\mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G})$. We will prove that there exists $f^{(\emptyset)} \in \mathcal{M}(\mathcal{G})$ such that $\forall I \in \mathcal{I},\ f^{(I)}$ factorizes according to Equation (1. Since $f^{(\emptyset)} \in \mathcal{M}(\mathcal{G})$, $f^{(\emptyset)}$ must factorize as $f^{(\emptyset)}(X) = \prod_{j\in[p]} f^{(\emptyset)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)})$ For each $j \in [p]$, let $f^{(\emptyset)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)}) = f^{(I_j)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)})$ for some $I_j \in \mathcal{I}$ s.t. $j \notin I$. If such a choice of I_j does not exist, then let $f^{(\emptyset)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)})$ be an arbitrary strictly positive density. Then note that for any $I \in \mathcal{I}$, we have

$$
\begin{aligned}
f^{(I)}(X) &= \prod_{j\in[p]} f^{(I)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)}) \\
&= \prod_{i\in I} f^{(I)}(X_i|X_{\mathrm{pa}_\mathcal{G}(j)}) \prod_{j\notin I} f^{(I_j)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)}) \\
&= \prod_{i\in I} f^{(I)}(X_i|X_{\mathrm{pa}_\mathcal{G}(j)}) \prod_{j\notin I} f^{(\emptyset)}(X_j|X_{\mathrm{pa}_\mathcal{G}(j)}),
\end{aligned}
$$

which completes the proof. □

Proof of Proposition 3.8. To prove the "if" direction, choose any $I \in \mathcal{I}$ and use the chain rule to factorize $f^{(I)}$ according to a topological ordering π consistent with $\mathcal{G}$. Specifically, if we let $a_\pi(i)$ denote the nodes that precede i in this ordering, then $f^{(I)}(X) = \prod_i f^{(I)}(X_i|X_{a_\pi(i)})$. Since every node is d-separated from its non-descendents given its parents, using condition (1) of the $\mathcal{I}$-Markov property, we can reduce the factorizations to $f^{(I)}(X) = \prod_i f^{(I)}(X_i|X_{pa(i)})$. Furthermore, since any node $i \notin I$ is d-separated from I given its parents, using condition (2) of the $\mathcal{I}$-Markov property, we can substitute the interventional conditional distributions with the observational ones, resulting in $f^I(X) = \prod_{i\in I} f^I(X_i|X_{pa(i)}) \prod_{i\notin I} f^\emptyset(X_i|X_{pa(i)})$. Since this factorization holds for every $I \in \mathcal{I}$, $\{f^I\}_{I\in\mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G})$ by Lemma A.1.

To prove the "only if" part of the statement, suppose $\{f^I\}_{I\in\mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G})$. By Lemma A.1, $f^{(I)}$ factorizes according to Equation (1) and satisfies the Markov property with respect to $\mathcal{G}$ for all $I \in \mathcal{I}$. It follows that $f^{(I)}$ must also satisfy the Markov property based on d-separation with respect to $\mathcal{G}$ (Verma & Pearl, 1991). Therefore, condition (1) of the $\mathcal{I}$-Markov property is satisfied.

To prove the second condition, choose any disjoint $A, C \subset [p]$ and any $I \in \mathcal{I}$, and suppose $C \cup \zeta_{\mathcal{I}\setminus\{I\}}$ d-separates A from $\{\zeta_I\}$ in $\mathcal{G}^{\mathcal{I}}$. Let V_{An} be the ancestral set of A and C with respect to $\mathcal{G} = (V, E)$. Let $B' \subset V_{An}$ contain all nodes in V_{An} that are d-connected to $\{\zeta_I\}$ in $\mathcal{G}^{\mathcal{I}}$ given $C \cup \zeta_{\mathcal{I}\setminus\{I\}}$, and let $A' = V_{An}\setminus(B' \cup C)$. Since by Lemma A.1, $f^{(I)}$ factorizes over $\mathcal{G}$ according to Equation (1) for every $I \in \mathcal{I}$, then choosing $\hat{I} \in \{\emptyset, I\}$ yields

$$
\begin{aligned}
f^{(\hat{I})}(X) &= f^{(\hat{I})}(X_{A'}, X_{B'}, X_C, X_{V\setminus V_{An}}) \\
&= \prod_{i\in A'} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}}) \prod_{i\in C, pa_\mathcal{G}(i)\cap A'\neq\emptyset} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}}) \\
&\quad \prod_{i\in C, pa_\mathcal{G}(i)\cap A'=\emptyset} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}}) \prod_{i\in B'} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}}) \prod_{i\in V\setminus V_{An}} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}})
\end{aligned}
$$

$$= \prod_{i\in A'} f^{(\emptyset)}(X_i|X_{pa(i),\mathcal{G}}) \prod_{i\in C, pa_{\mathcal{G}}(i)\cap A'\neq\emptyset} f^{(\emptyset)}(X_i|X_{pa(i),\mathcal{G}})$$
$$\prod_{i\in C, pa_{\mathcal{G}}(i)\cap A'=\emptyset} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}}) \prod_{i\in B'} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}}) \prod_{i\in V\setminus V_{An}} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}})$$

The second equality holds by the factorization of Equation (1) because either $i \in A'$ or $i \in C \mid pa_{\mathcal{G}(i)} \cap A' \neq \emptyset$ implies that i is not targeted by the intervention on $\hat{I}$, i.e. $i \notin \hat{I}$. To see this, recall that A' is separated from $\zeta_{\hat{I}}$ in $\mathcal{G}^{\mathcal{I}}$, which implies that A' does not contain a child of $\zeta_{\hat{I}}$ in $\mathcal{G}^{\mathcal{I}}$ and is therefore not targeted by the intervention on $\hat{I}$. Likewise, $\{i \in C|pa_{\mathcal{G}(i)} \cap A' \neq \emptyset\}$ does not contain a child of $\zeta_{\hat{I}}$ in $\mathcal{G}^{\mathcal{I}}$ because otherwise A' and $\zeta_{\hat{I}}$ would be d-connected in $\mathcal{G}^{\mathcal{I}}$ by conditioning on this node.

Using similar reasoning, it is easy to see that the parent sets of A' and $\{i \in C|pa_{\mathcal{G}(i)} \cap A' \neq \emptyset\}$ with respect to $\mathcal{G}$ are subsets of $A' \cup C$; and the parent sets of $\{i \in C|pa_{\mathcal{G}(i)} \cap A' = \emptyset\}$ and B' are subsets of $B' \cup C$. Therefore, we can write

$$f^{(\hat{I})}(X) = g_1(X_{A'}, X_C)g_2(X_{B'}, X_C; \hat{I}) \prod_{i\in V\setminus V_{An}} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}})$$

where

$$g_1(X_{A'}, X_C) = \prod_{i\in A'} f^{(\emptyset)}(X_i|X_{pa(i),\mathcal{G}}) \prod_{i\in C, pa_{\mathcal{G}}(i)\cap A'\neq\emptyset} f^{(\emptyset)}(X_i|X_{pa(i),\mathcal{G}})$$

and

$$g_2(X_{B'}, X_C; \hat{I}) = \prod_{i\in C, pa_{\mathcal{G}}(i)\cap A'=\emptyset} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}}) \prod_{i\in B'} f^{(\hat{I})}(X_i|X_{pa(i),\mathcal{G}})$$

Marginalizing out $X_{A'\setminus A}, X_{B'}$ and $X_{V_{An}}$ yields

$$f^{(\hat{I})}(X_A, X_C) = \hat{g}_1(X_A, X_C)\hat{g}_2(X_C; \hat{I})$$

where $\hat{g}_1(X_A, X_C) = \int_{X_{A'\setminus A}} g_1(X_{A'}, X_C)$ and $\hat{g}_2(X_C; \hat{I}) = \int_{X_{B'}} g_2(X_{B'}, X_C; \hat{I})$. From here, it is easy to see that $f^{(\hat{I})}(X_A|X_C) = \frac{g^{(\emptyset)}(X_A,X_C)}{\int_{X_A} g^{(\emptyset)}(X_A,X_C)}$ is invariant to $\hat{I} \in \{\emptyset, I\}$. □

Proof of Lemma 3.10. Choose any disjoint $A, C \subset [p]$ and any $I \in \mathcal{I}$, and suppose $C \cup \zeta_{\mathcal{I}\setminus\{I\}}$ does not d-separate A from $\{\zeta_I\}$ in $\mathcal{G}^{\mathcal{I}}$. To prove this lemma, it is sufficient to construct $f^{(\emptyset)}$ and $f^{(I)}$ such that they satisfy the $\hat{\mathcal{I}}$-Markov properties with respect to $\mathcal{G}^{\hat{\mathcal{I}}}$, where $\hat{\mathcal{I}} = \{\emptyset, I\}$, and $f^{(\emptyset)}(X_A|X_C) \neq f^{(I)}(X_A|X_C)$.

To do this, we construct a subgraph $\mathcal{G}_{sub} = (V, E_{sub})$ that consists of a d-connected path, $P = \{p_1 \in I, p_2, \cdots, p_{k-1}, p_k \in A\}$ with $p_2, \cdots, p_{k-1} \notin I \cup A$, as well as the directed paths from colliders in P to their nearest descendents in C. All other nodes that are not part of these paths are part of the subgraph but have no edges. We parameterize the set of conditional probability distributions, $\{f^{(\emptyset)}(X_i|X_{pa_{G_{sub}}(i)})\}_{i\in V}$, using linear structural equations with non-zero coefficients and independent Gaussian noise. Consider $f^{(\emptyset)}(X_{p_1}|X_{pa_{G_{sub}}(p_1)}) = \mathcal{N}(\sum_{j\in pa_{G_{sub}}(p_1)} c_{j,p_1}X_j, \sigma^2_\emptyset)$. To construct $f^{(I)}$, let $f^{(I)}(X_{p_1}|X_{pa_{G_{sub}}(p_1)}) = \mathcal{N}(\sum_{j\in pa_{G_{sub}}(p_1)} c_{j,p_1}X_j, \sigma^2_I)$ for some $\sigma_I \neq \sigma_\emptyset$, and let $f^{(I)}(X_i|X_{pa_{G_{sub}}(i)}) = f^{(\emptyset)}(X_i|X_{pa_{G_{sub}}(i)})$ for all $i \neq p_1$.

Note that these distributions factor over $\mathcal{G}$ according to Definition 3.2, so by Proposition 3.8, they satisfy the $\hat{\mathcal{I}}$-Markov properties with respect to $\mathcal{G}^{\hat{\mathcal{I}}}$. Furthermore, it is straightforward to show that we can write $X_{p_k} = N_{(\emptyset)} + S(X_C)$ under the model corresponding to $f^{(\emptyset)}$ and $X_{p_k} = N_{(I)} + S(X_C)$ under the model corresponding to $f^{(I)}$, where $N_{(\emptyset)} \sim \mathcal{N}(0, c\sigma^2_\emptyset)$ and $N_{(I)} \sim \mathcal{N}(0, c\sigma^2_I)$ for some constant c, and $S(X_C)$ is a Gaussian random variable independent of $N_{(\emptyset)}$ and $N_{(I)}$. Since $\sigma^2_\emptyset \neq \sigma^2_I$, it follows that we have $f^{(\emptyset)}(X_A|X_C) \neq f^{(I)}(X_A|X_C)$, as desired. □

Proof of Corollary 3.12. If $\mathcal{G}_1$ and $\mathcal{G}_2$ are in the same perfect-$\mathcal{I}$-MEC, then $\mathcal{G}_1$ and $\mathcal{G}_2$ have the same skeleton and v-structures. Since $\mathcal{G}_1^{\mathcal{I}}$ and $\mathcal{G}_2^{\mathcal{I}}$ are constructed from $\mathcal{G}_1$ and $\mathcal{G}_2$ by adding the same set of vertices and edges, they must have the same skeleta, so we just need to show that they also have the same v-structures to prove they belong to the same $\mathcal{I}$-MEC. Suppose this is not the case. The only v-structures that can differ between $\mathcal{G}_1^{\mathcal{I}}$ and $\mathcal{G}_2^{\mathcal{I}}$ must involve $\mathcal{I}$-edges, since $\mathcal{G}_1$ and $\mathcal{G}_2$ have the same v-structures. Without loss of generality, suppose $\zeta_I \to i$ is part of a v-structure in $\mathcal{G}_1^{\mathcal{I}}$ but not in $\mathcal{G}_2^{\mathcal{I}}$. This could only occur if there were a neighbor $j \notin I$ with orientation $j \to i$ in $\mathcal{G}_1^{\mathcal{I}}$ and $i \to j$ in $\mathcal{G}_2^{\mathcal{I}}$. However, this contradicts the assumption that $\mathcal{G}_1$ and $\mathcal{G}_2$ belong to the same perfect-$\mathcal{I}$-MEC (Hauser & Bühlmann, 2012), since removing the incoming edges of i from $\mathcal{G}_1$ and $\mathcal{G}_2$ would result in graphs with different skeleta. Therefore, $\mathcal{G}_1^{\mathcal{I}}$ and $\mathcal{G}_2^{\mathcal{I}}$ must have the same v-structures.

Conversely, suppose that $\mathcal{G}_1$ and $\mathcal{G}_2$ are in the same $\mathcal{I}$-MEC. Then they must have the same skeleta and v-structures, and we just need to show that for any $I \in \mathcal{I}$, $\mathcal{G}_1$ and $\mathcal{G}_2$ have the same skeleton after removing the incoming edges of i for all $i \in I$ (Hauser & Bühlmann, 2012). Suppose this is not the case. This implies that for some $I \in \mathcal{I}$ and some $i \in I$, there is an edge between i and another vertex $j \notin I$ that is removed in $\mathcal{G}_1$ but not in $\mathcal{G}_2$. The orientation of this edge must be $j \to i$ in $\mathcal{G}_1$ and $i \to j$ in $\mathcal{G}_2$. But this would mean that $j \to i$ and $\zeta_I \to i$ form a v-structure in $\mathcal{G}_1^{\mathcal{I}}$ but not in $\mathcal{G}_2^{\mathcal{I}}$, which is a contradiction to Theorem 3.9. Therefore, $\mathcal{G}_1$ and $\mathcal{G}_2$ must belong to the same perfect-$\mathcal{I}$-MEC. □

A.2. Proofs from Section 3.3

The following definition formalizes the notion of relabeling the datasets and intervention targets:

Definition A.2. Let $\{f^{(I)}\}_{I\in\mathcal{I}}$ be a set of interventional distributions. Let $J \in \mathcal{I}$ be a particular intervention target. The corresponding *J-observation target set* is defined as $\tilde{\mathcal{I}}_J := \{\emptyset, \{I \cup J\}_{I\in\mathcal{I}, I\neq J}\}$. The relabeled set of interventional distributions is denoted $\{\tilde{f}_J^{(I)}\}_{I\in\tilde{\mathcal{I}}_J}$, with $\tilde{f}_J^{(\emptyset)} := f^{(J)}$ and $\tilde{f}_J^{(I\cup J)} := f^{(I)}$, $\forall I \in \mathcal{I}, I \neq J$.

Notice that $\{\tilde{f}_J^{(I)}\}_{I\in\tilde{\mathcal{I}}_J}$ contains the same distributions as $\{f^{(I)}\}_{I\in\mathcal{I}}$ but is reindexed to treat $f^{(J)}$ as the observational distribution and $\{f^{(I)}\}_{I\neq J}$ as distributions obtained under interventions on $I\cup J$. This relabeling is justified by the following lemma:

Lemma A.3. *$\{f^{(I)}\}_{I\in\mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G})$ if and only if $\{\tilde{f}_J^{(I)}\}_{I\in\tilde{\mathcal{I}}_J} \in \mathcal{M}_{\tilde{\mathcal{I}}_J}(\mathcal{G})$ for all $J \in \mathcal{I}$.*

Proof of Lemma A.3. To prove the "only if" direction, suppose $\{f^{(I)}\}_{I\in\mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G})$. It follows straight from Definition A.2 that $\tilde{f}_J^{(\emptyset)} = f^{(J)}$ for every $J \in \mathcal{I}$. Since $f^{(J)}$ is Markov with respect to $\mathcal{G}$, so is $\tilde{f}_J^{(\emptyset)}$, and hence $\tilde{f}_J^{(\emptyset)}$ can be factored according to Equation (1) with the observational distribution set to $\tilde{f}_J^{(\emptyset)}$. So it remains to show that for every $J \in \mathcal{I}$ and any $I \neq J$, $\tilde{f}_J^{(I\cup J)}$ factorizes according to Equation (1) with the observational distribution set to $\tilde{f}_J^{(\emptyset)}$. Then we have

$$
\begin{aligned}
\tilde{f}_J^{(I\cup J)}(X) &= f^{(I)}(X) && \text{(by Definition A.2)}\\
&= \prod_{i\notin I} f^{(\emptyset)}(X_i|X_{pa(i)}) \prod_{i\in I} f^{(I)}(X_i|X_{pa(i)}) && \text{(by Lemma A.1)}\\
&= \prod_{i\notin I,J} f^{(\emptyset)}(X_i|X_{pa(i)}) \prod_{i\notin I, i\in J} f^{(\emptyset)}(X_i|X_{pa(i)}) \prod_{i\in I} f^{(I)}(X_i|X_{pa(i)})\\
&= \prod_{i\notin I\cup J} \tilde{f}_J^{(\emptyset)}(X_i|X_{pa(i)}) \prod_{i\in I\cup J} \tilde{f}_J^{(I\cup J)}(X_i|X_{pa(i)})
\end{aligned}
$$

where the last equality holds because $\tilde{f}_J^{(\emptyset)}(X_i|X_{pa(i)}) = f^{(J)}(X_i|X_{pa(i)}) = f^{(\emptyset)}(X_i|X_{pa(i)})$ when $i \notin J$, and by relabeling the conditional distributions in the last two product terms as $\tilde{f}_J^{(I\cup J)}$. By Lemma A.1, it follows that $\{\tilde{f}_J^{(I)}\}_{I\in\tilde{\mathcal{I}}_J} \in \mathcal{M}_{\tilde{\mathcal{I}}_J}(\mathcal{G})$.

To prove the converse, we show how to construct the observational distribution $f^{(\emptyset)}$ such that $f^{(I)}$ can be factored over $\mathcal{G}$ according to Equation (1) for all $I \in \mathcal{I}$. For every $i \in \mathcal{V}$, let $f^{(\emptyset)}(X_i|X_{pa_{\mathcal{G}}(i)}) = \tilde{f}_I^{(\emptyset)}(X_i|X_{pa_{\mathcal{G}}(i)})$ for some $I \in \mathcal{I}$ such that $i \notin I$. The existence of such an I is guaranteed by the assumption that $\mathcal{I}$ is a conservative set of targets. Furthermore, $\tilde{f}_I^{(\emptyset)}(X_i|X_{pa_{\mathcal{G}}(i)})$ is unique; if there are multiple targets that satisfy this requirement (i.e. $\exists J \in \mathcal{I}\ s.t.\ i \notin J$ and $J \neq I$), we always have $\tilde{f}_I^{(\emptyset)}(X_i|X_{pa_{\mathcal{G}}(i)}) = \tilde{f}_J^{(\emptyset)}(X_i|X_{pa_{\mathcal{G}}(i)})$, since

$$
\tilde{f}_J^{(\emptyset)}(X_i|X_{pa_{\mathcal{G}}(i)}) = \tilde{f}_I^{(I\cup J)}(X_i|X_{pa_{\mathcal{G}}(i)}) = \tilde{f}_I^{(\emptyset)}(X_i|X_{pa_{\mathcal{G}}(i)})
$$

for $i \notin I \cup J$. The first equality follows by Definition A.2, and the second equality follows since by Proposition 3.8, $\{\tilde{f}_I^{(K)}\}_{K \in \tilde{\mathcal{I}}_I} \in \mathcal{M}_{\tilde{\mathcal{I}}_I}(\mathcal{G})$. Thus, we have defined $f^{(\emptyset)}$ such that $f^{(I)}$ can be factored over $\mathcal{G}$ according to Equation (1) for all $I \in \mathcal{I}$. This proves by Lemma A.1 that $\{f^{(I)}\}_{I \in \mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G})$. □

Proof of Theorem 3.14. We first prove the "only if" direction. If $\mathcal{G}_1$ and $\mathcal{G}_2$ belong to the same $\mathcal{I}$-MEC, then we have $\mathcal{M}_{\mathcal{I}}(\mathcal{G}_1) = \mathcal{M}_{\mathcal{I}}(\mathcal{G}_2)$. It follows from Lemma A.3 that $\mathcal{M}_{\tilde{\mathcal{I}}_J}(\mathcal{G}_1) = \mathcal{M}_{\tilde{\mathcal{I}}_J}(\mathcal{G}_2)$ for all $J \in \mathcal{I}$, and then directly from Theorem 3.9 that $\mathcal{G}_1^{\tilde{\mathcal{I}}_J}$ and $\mathcal{G}_2^{\tilde{\mathcal{I}}_J}$ have the same skeleta and v-structures for all $J \in \mathcal{I}$. The "if" direction follows by applying these same results in reverse: first we apply Theorem 3.9 to $\mathcal{G}_1^{\tilde{\mathcal{I}}_J}$ and $\mathcal{G}_2^{\tilde{\mathcal{I}}_J}$ for every $J \in \mathcal{I}$, followed by Lemma A.3. □

B. Proofs from Section 4

B.1. Proof of Theorem 4.6

In this section, we work up to the proof of Theorem 4.6. To do this, we first cover some basic results on the consistency of GSP. Let $\mathcal{G}$ be a DAG and let $\mathcal{H}$ be an independence map (I-map) of $\mathcal{G}$, meaning that all independences implied by $\mathcal{H}$ are satisfied by $\mathcal{G}$ (i.e. $\mathcal{G} \leq \mathcal{H}$). Chickering (2002) showed that there exists a sequence of covered edge reversals and edge additions resulting in a sequence of DAGs, $\mathcal{G}_0, \mathcal{G}_1, \cdots, \mathcal{G}_\tau$ such that

$$\mathcal{G} = \mathcal{G}_0 \leq \mathcal{G}_1 \leq \cdots \leq \mathcal{G}_\tau = \mathcal{H}$$

Furthermore, Solus et al. (2017) showed that for any $\mathcal{G}$ and $\mathcal{H}$, there exists such a Chickering sequence in which one sink node of $\mathcal{H}$ is fixed at a time. The following lemma connects this sequence over DAGs to a sequence over the topological orderings of the nodes.

Lemma B.1. *Let $\mathcal{G}_{i_1}, \cdots, \mathcal{G}_{i_p}$ be a subsequence of the Chickering sequence where one sink is fixed at a time, and let $\mathcal{G}_{i_j}$ be the first DAG in which the jth sink node is fixed, i.e. the sequence of DAGs from $\mathcal{G}_{i_{j-1}}$ to $\mathcal{G}_{i_j}$ involve covered edge reversals and edge additions required to resolve sink node j. Furthermore, let $\Pi(\mathcal{G})$ denote the set of topological orderings that are consistent with $\mathcal{G}$. Then for any $\pi_{i_{j-1}} \in \Pi(\mathcal{G}_{i_{j-1}})$ in which the last $j-1$ nodes correspond to the first $j-1$ fixed sink nodes, there exists a sequence of orderings $\pi_{i_{j-1}}, \cdots, \pi_{i_j}$ with $\pi_k \in \Pi(\mathcal{G}_k)$ such that the jth sink node moves only to the right, stopping in the jth position from the end, and the relative ordering of the other nodes remain unchanged.*

Proof. The correctness of this lemma follows directly from Lemma 13 of Solus et al. (2017). □

The following corollary is an immediate consequence of this lemma.

Corollary B.2. *For any DAG $\mathcal{G}$ over vertex set $[p]$ and any I-map $\mathcal{H}$, there exists a sequence of topological orderings*

$$\pi_0 \in \Pi(\mathcal{G}_0), \pi_1 \in \Pi(\mathcal{G}_1), \cdots, \pi_\tau \in \Pi(\mathcal{G}_\tau)$$

with $\mathcal{G}_0 = \mathcal{G}$ and $\mathcal{G}_\tau = \mathcal{H}$ corresponding to a Chickering sequence in which we fix the order of the nodes in reverse starting from the last node in π_τ. Specifically, the last node in π_τ is moved to the right until it is in the p-th position, then the second-last node in π_τ to the right until it is in the $(p-1)$-th position, etc. until all nodes are in the order given by π_τ.

Using this result, we now state the following lemma, which is useful in the proof of consistency of the algorithm.

Lemma B.3. *For any permutation π, there exists a list of covered arrow reversals from $\mathcal{G}_\pi$ to the true DAG $\mathcal{G}_{\pi^*}$ such that (1) the number of edges is weakly decreasing:*

$$\mathcal{G}_\pi = \mathcal{G}_{\pi^0} \geq \mathcal{G}_{\pi^1} \geq \cdots \geq \mathcal{G}_{\pi^{m-1}} \geq \mathcal{G}_{\pi^m} \geq \cdots \geq \mathcal{G}_{\pi^{M-1}} \geq \mathcal{G}_{\pi^M} = \mathcal{G}_{\pi^*}$$

and (2) if $i \to j$ is reversed from $\mathcal{G}_{\pi^{m-1}}$ to $\mathcal{G}_{\pi^m}$, then there is a directed path from j to i in $\mathcal{G}_{\pi^*}$ (i.e. j is an ancestor of i).

Proof. Suppose that j is not an ancestor of i in $\mathcal{G}_{\pi^*}$. Then there exists an ordering $\pi_0 \in \Pi(G_{pi^*})$ in which j comes after i. By Corollary B.2, there exists a Chickering sequence from $\mathcal{G}_{\pi^*}$ to $\mathcal{G}_\pi$ and a corresponding sequence of orderings such that j never moves from the left of i to the right of i. Specifically, either j is fixed before i and their relative ordering never changes, or i is fixed before j and moves from the left of j to the right of j once. It follows that $j \to i$ is never reversed in the Chickering sequence from $\mathcal{G}_{\pi^*}$ to $\mathcal{G}_\pi$, and thus $i \to j$ is never reversed when working backwards from $\mathcal{G}_\pi$ to $\mathcal{G}_{\pi^*}$. Therefore, if $i \to j$ is reversed from $\mathcal{G}_{\pi^{m-1}}$ to $\mathcal{G}_{\pi^m}$, then there must be a directed path from j to i in $\mathcal{G}_{\pi^*}$. □

In turn, Lemma B.3 allows us to prove the existance of a greedy path from $\mathcal{G}_\pi$ to the true DAG $\mathcal{G}_{\pi^*}$ by reversing $\mathcal{I}$-covered edges.

Lemma B.4. *For any permutation π, there exists a list of $\mathcal{I}$-covered arrow reversals from $\mathcal{G}_\pi$ to the true DAG $\mathcal{G}_{\pi^*}$ such that the number of edges is weakly decreasing and if $i \to j$ is reversed from $\mathcal{G}_{\pi^{m-1}}$ to $\mathcal{G}_{\pi^m}$, then there is a directed path from j to i in $\mathcal{G}_{\pi^*}$.*

Proof. From Lemma B.3, we know that there exists a sequence of covered arrow reversals from $\mathcal{G}_\pi$ to $\mathcal{G}_{\pi^*}$ in which the number of edges is weakly decreasing; and that this sequence has the property that if arrow $i \to j$ is reversed from $\mathcal{G}_{\pi^{m-1}}$ to $\mathcal{G}_{\pi^m}$, then there is a directed path from j to i in $\mathcal{G}_{\pi^*}$. It remains to be shown that all covered arrow reversals in this sequence are also $\mathcal{I}$-covered arrow reversals.

Suppose $\mathcal{I}_{j\setminus i} \in \mathcal{I}$ and let $\mathcal{G}^{(\mathcal{I})}_{\pi^*}$ denote the $\mathcal{I}$-DAG of $\mathcal{G}_{\pi^*}$ (Definition 3.5). Note that there is a directed path from $\{\zeta_{\mathcal{I}_{j\setminus i}}\}$ to i in $\mathcal{G}^{(\mathcal{I})}_{\pi^*}$ by the previous lemmas. Therefore, by Assumption 4.4, $f(X_i)$ is not invariant to $I \in \mathcal{I}_{j\setminus i} \cup \emptyset$. Suppose $\{i\} \in \mathcal{I}$ and let $\mathcal{G}^{(\mathcal{I})}_{\pi^*}$ denote the $\mathcal{I}$-DAG of $\mathcal{G}_{\pi^*}$ (Definition 3.5). Note that $\{\zeta_{\{i\}}\}$ is d-separated from j in $\mathcal{G}^{(\mathcal{I})}_{\pi^*}$. Therefore, $f(X_j)$ is invariant to $I \in \{\emptyset, \{i\}\}$ by the $\mathcal{I}$-Markov property (Definition 3.6) and Proposition 3.8. It follows that $i \to j$ is $\mathcal{I}$-covered in $\mathcal{G}_{\pi^{m-1}}$. If $\{i\} \notin \mathcal{I}$ and $\mathcal{I}_{j\setminus i}$ is empty, then the result is trivial as $i \to j$ is $\mathcal{I}$-covered as long as it is covered. ☐

The following lemma proves the correctness of using $\mathcal{I}$-contradictory arrows as the secondary search criterion; essentially, it states that when $\mathcal{G}_\pi$ is in the same MEC but not the same $\mathcal{I}$-MEC as $\mathcal{G}_{\pi^*}$, then $\mathcal{G}_\pi$ has more $\mathcal{I}$-contradictory arrows than $\mathcal{G}_{\pi^*}$.

Lemma B.5. *For any permutation π such that $\mathcal{G}_\pi$ and $\mathcal{G}_{\pi^*}$ are in the same MEC, there exists a list of $\mathcal{I}$-covered arrow reversals from $\mathcal{G}_\pi$ to the true DAG $\mathcal{G}_{\pi^*}$*

$$\mathcal{G}_\pi = \mathcal{G}_{\pi^0} \geq \mathcal{G}_{\pi^1} \geq \cdots \geq \mathcal{G}_{\pi^{m-1}} \geq \mathcal{G}_{\pi^m} \geq \cdots \geq \mathcal{G}_{\pi^{M-1}} \geq \mathcal{G}_{\pi^M} = \mathcal{G}_{\pi^*}$$

such that the number of arrows is non-increasing and for all m, if $\mathcal{G}_{\pi^{m-1}}$ and $\mathcal{G}_{\pi^m}$ are not in the same $\mathcal{I}$-MEC, then $\mathcal{G}_{\pi^m}$ is produced from $\mathcal{G}_{\pi^{m-1}}$ by the reversal of an $\mathcal{I}$-contradictory arrow.

Proof. From Lemma B.4, we know there exists a sequence of $\mathcal{I}$-covered arrow reversals from $\mathcal{G}_\pi$ to $\mathcal{G}_{\pi^*}$ in which the number of edges is weakly decreasing, with the property that if $i \to j$ is reversed from $\mathcal{G}_{\pi^{m-1}}$ to $\mathcal{G}_{\pi^m}$, then there is a directed path from j to i in $\mathcal{G}_{\pi^*}$.

Suppose the arrow $i \to j$ is reversed from $\mathcal{G}_{\pi^{m-1}}$ to $\mathcal{G}_{\pi^m}$. Since $\mathcal{G}_{\pi^{m-1}}$ and $\mathcal{G}_{\pi^*}$ are in the same MEC, this implies that $j \to i$ is in $\mathcal{G}_{\pi^*}$. Since $\mathcal{G}_{\pi^{m-1}}, \mathcal{G}_{\pi^m}$ are not in the same $\mathcal{I}$-MEC, then we must have $\mathcal{I}_{i\setminus j} \cup \mathcal{I}_{i\setminus j} \neq \emptyset$. Now, let $\mathcal{G}^{(\mathcal{I})}_{\pi^*}$ denote the $\mathcal{I}$-DAG of $\mathcal{G}_{\pi^*}$ (Definition 3.5), and consider the following cases:

(1) $\mathcal{I}_{i\setminus j} \neq \emptyset$. Then there exists a subset $S \subset ne_{\mathcal{G}}(j)\setminus\{i\}$ that d-separates $\zeta_{\mathcal{I}_{i\setminus j}}$ from j in $\mathcal{G}^{(\mathcal{I})}_{\pi^*}$. By the $\mathcal{I}$-Markov property (Definition 3.6) and Proposition 3.8, $f^{(\emptyset)}(X_j|X_S) = f^{(I)}(X_j|X_S)$ for all $I \in \mathcal{I}_{i\setminus j}$.

(2) $\mathcal{I}_{j\setminus i} \neq \emptyset$. Then for any subset $S \subset ne_{\mathcal{G}}(i)\setminus\{j\}$, $\zeta_{\mathcal{I}_{j\setminus i}}$ is d-connected to i in $\mathcal{G}^{(\mathcal{I})}_{\pi^*}$. By Assumption 4.5, $f^{(\emptyset)}(X_i|X_S) \neq f^{(I)}(X_i|X_S)$ for some $I \in \mathcal{I}_{j\setminus i}$.

(3) $\{i\} \in \mathcal{I}$. Then $\zeta_{\{i\}}$ is d-separated from j in $\mathcal{G}^{(\mathcal{I})}_{\pi^*}$. Therefore, $f^{\{i\}}(X_j) = f^{\emptyset}(X_j)$ by the $\mathcal{I}$-Markov property (Definition 3.6) and Proposition 3.8.

(4) $\{j\} \in \mathcal{I}$. Then $\zeta_{\{j\}}$ is not d-separated from i in $\mathcal{G}^{(\mathcal{I})}_{\pi^*}$. Therefore, $f^{\{j\}}(X_i) \neq f^{\emptyset}(X_i)$ by Assumption 4.4.

These are the defining properties of $\mathcal{I}$-contradictory edges. Therefore, the arrow $i \to j$ is $\mathcal{I}$-contradictory in $\mathcal{G}_{\pi^{m-1}}$. ☐

Proof of Theorem 4.6. This follows directly from Lemmas B.4 and B.5. ☐

B.2. Pooling Data for CI Testing

The following proposition gives sufficient conditions under which CI relations hold when the data come from a mixture of interventional distributions:

Proposition B.6. *Let $\{f^{(I)}\}_{I\in\mathcal{I}} \in \mathcal{M}_{\mathcal{I}}(\mathcal{G})$ for a DAG $\mathcal{G} = ([p], E)$ and intervention targets $\mathcal{I}$ s.t. $\emptyset \in \mathcal{I}$. For some $\mathcal{I}_s \subset \mathcal{I}$ and some disjoint $A, B, C \subset [p]$, suppose that $C \cup \zeta_{\mathcal{I}\setminus\mathcal{I}_s}$ d-separates A from $B \cup \zeta_{\mathcal{I}_s}$ in $\mathcal{G}^{\mathcal{I}}$. Then $X_A \perp\!\!\!\perp X_B \mid X_C$ under the distribution $X \sim \sum_{I\in\{\emptyset\}\cup\mathcal{I}_s} \alpha_I f^{(I)}$, for any $\alpha_I \in (0, 1)$ s.t. $\sum_{I\in\{\emptyset\}\cup\mathcal{I}_s} \alpha_I = 1$.*

Proposition B.6 can be used to derive a set of checkable conditions on $\mathcal{G}_\pi$ to determine whether each interventional dataset $I \in \mathcal{I}$ can be pooled with observational data to test $X_i \perp\!\!\!\perp X_k \mid X_{\mathrm{an}_{\mathcal{G}_\pi}(i)\setminus\{k\}}$ for $k \in \mathrm{pa}_{\mathcal{G}_\pi}(i)$.

Corollary B.7. *Suppose we want to test $X_i \perp\!\!\!\perp X_k \mid X_{\mathrm{an}_{\mathcal{G}_\pi}(i)\setminus\{k\}}$ for some $k \in \mathrm{pa}_{\mathcal{G}_\pi}(i)$. Let $\mathcal{I}_s \subset \mathcal{I}$ be interventional targets such that the following two conditions hold for every $j \in I \in \mathcal{I}_s$:*

(1) $j = i$ or j is neither a descendent nor an ancestor of i;

(2) $\pi(k) > \pi(j)$ and k is not a parent of j; or $\pi(j) > \pi(k)$ and j is not an ancestor of k,

where all relations are being considered with respect to $\mathcal{G}_\pi$, and $\pi(i)$ denotes the index of i in π. Then under the faithfulness assumption, $X_i \perp\!\!\!\perp X_k \mid X_{\mathrm{an}_{\mathcal{G}_\pi}(i)\setminus\{k\}}$ under $X \sim f^{\emptyset}$ if and only if this CI relation also holds under $X \sim \sum_{I\in\{\emptyset\}\cup\mathcal{I}_s} \alpha_I f^{(I)}$, where $\alpha_I \in (0, 1)$ and $\sum_{I\in\{\emptyset\}\cup\mathcal{I}_s} \alpha_I = 1$.

Proof. If $X_i \not\perp\!\!\!\perp X_k \mid X_{\mathrm{an}_{\mathcal{G}_\pi}(i)\setminus\{k\}}$ under $X \sim f^{\emptyset}$, then this CI relation will clearly not hold under $X \sim \sum_{I\in\{\emptyset\}\cup\mathcal{I}_s} \alpha_I f^{(I)}$, thereby implying the "if" direction. It remains to prove the "only if" direction, i.e. that $X_i \perp\!\!\!\perp X_k \mid X_{\mathrm{an}_{\mathcal{G}_\pi}(i)\setminus\{k\}}$ under $X \sim f^{\emptyset}$ implies conditional independence under $X \sim \sum_{I\in\{\emptyset\}\cup\mathcal{I}_s} \alpha_I f^{(I)}$.

We first consider the case where $j \neq i$ and j is neither a descendent nor an ancestor of i. By the faithfulness assumption, $X_i \perp\!\!\!\perp X_k \mid X_{\mathrm{an}_{\mathcal{G}_\pi}(i)\setminus\{k\}}$ implies that i and k are d-separated by $\mathrm{an}_{\mathcal{G}_\pi}(i) \setminus \{k\}$ in the true DAG $\mathcal{G}_*$. Since $\mathcal{G}_\pi$ is an independence map of $\mathcal{G}_*$, it follows from condition (2) that for any $j \in I \in \mathcal{I}_s$, j and k are d-separated by $\mathrm{an}_{\mathcal{G}_\pi}(j) \setminus \{k\}$ in $\mathcal{G}_*$. In addition, since j is neither a descendent nor an ancestor of i, then j and k are also d-separated by $\mathrm{an}_{\mathcal{G}_\pi}(i) \setminus \{k\}$ in $\mathcal{G}_*$.

If $i = j \in I \in \mathcal{I}_s$, then k and $\{i\} \cup \zeta_{\mathcal{I}_s}$ are d-separated in $\mathcal{G}_*^{\mathcal{I}}$ by $\zeta_{\mathcal{I}\setminus\mathcal{I}_s} \cup \mathrm{an}_{\mathcal{G}_\pi}(i) \setminus \{k\}$. It then follows from Proposition B.6 that $X_i \perp\!\!\!\perp X_k \mid X_{\mathrm{an}_{\mathcal{G}_\pi}(i)\setminus\{k\}}$ when $X \sim \sum_{I\in\{\emptyset\}\cup\mathcal{I}_s} \alpha_I f^{(I)}$. □

Proof of Proposition B.6. Similar to the proof of the second part of Proposition 3.8, it can be shown that for any disjoint $A, B, C \subset [p]$ and any $I \in \mathcal{I}$ such that $C \cup \zeta_{\mathcal{I}\setminus\{I\}}$ d-separates A from $\{\zeta_I\}$ in $\mathcal{G}^{\mathcal{I}}$, we have

$$f^{(I)}(X) = g_1(X_{A'}, X_C) g_2(X_{B'}, X_C; I) \prod_{i\in V\setminus V_{An}} f^{(I)}(X_i|X_{pa(i),\mathcal{G}})$$

where

$$g_1(X_{A'}, X_C) = \prod_{i\in A'} f^{(\emptyset)}(X_i|X_{pa(i),\mathcal{G}}) \prod_{i\in C, pa_{\mathcal{G}}(i)\cap A'\neq\emptyset} f^{(\emptyset)}(X_i|X_{pa(i),\mathcal{G}})$$

and

$$g_2(X_{B'}, X_C; I) = \prod_{i\in C, pa_{\mathcal{G}}(i)\cap A'=\emptyset} f^{(I)}(X_i|X_{pa(i),\mathcal{G}}) \prod_{i\in B'} f^{(I)}(X_i|X_{pa(i),\mathcal{G}})$$

where V_{An} is the ancestral set of $A \cup B \cup C$, A' is the largest subset of V_{An} that is d-separated from B and ζ_I given C, and $B' = V_{An} \setminus (A' \cup C)$. Noting that $B \subset B'$, we marginalize out $X_{A'\setminus A}$,$X_{B'\setminus B}$ and $X_{B'\setminus B}$, which yields

$$f^{(I)}(X_A, X_C) = \hat{g}_1(X_A, X_C)\hat{g}_2(X_B, X_C; I)$$

The mixture of distributions over all $I \in \mathcal{I}_s$ is therefore,

$$\sum_{I\in\mathcal{I}_s} \alpha_I f^{(I)}(X_A, X_C) = \hat{g}_1(X_A, X_C) \sum_{I\in\mathcal{I}_s} \alpha_I \hat{g}_2(X_B, X_C; I)$$

which factors into separate functions over X_A and X_B. Therefore, $X_A \perp\!\!\!\perp X_B|X_C$ when X is sampled from this mixture of distributions. □

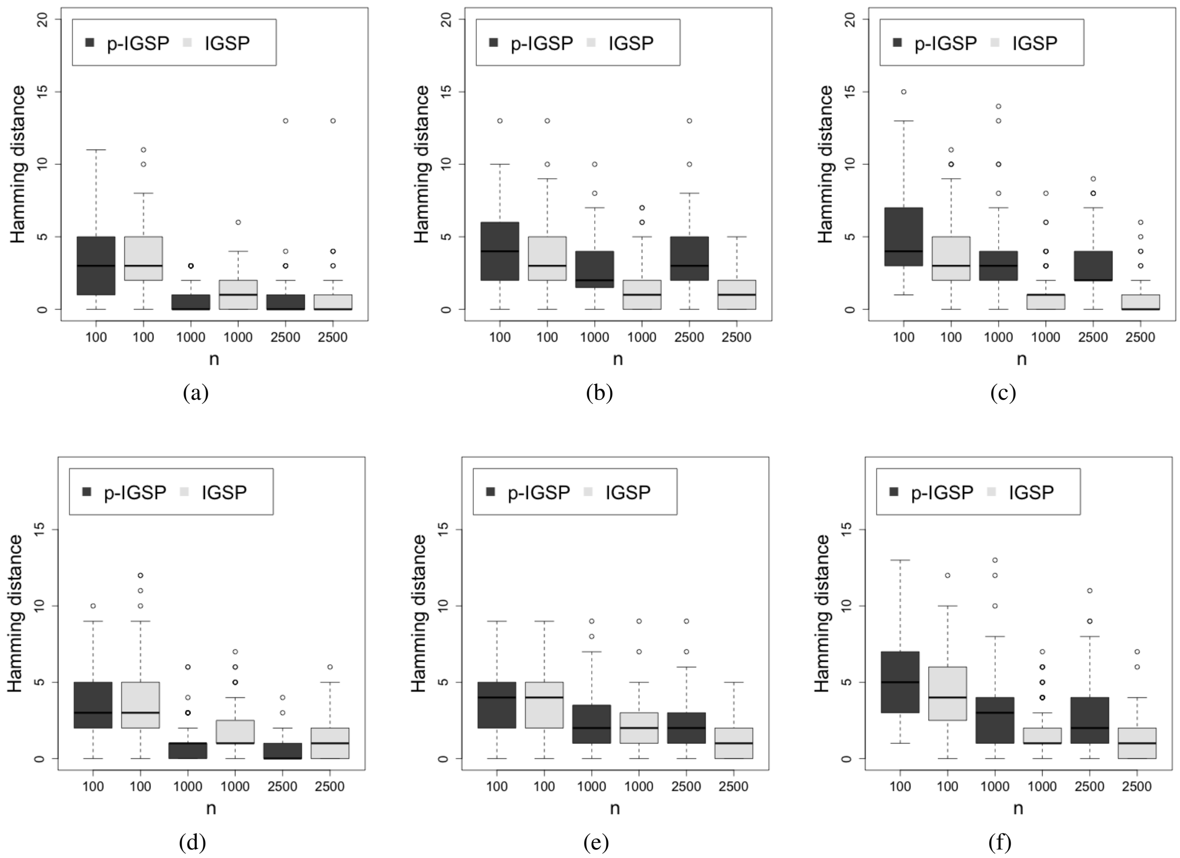

Figure 8. Distributions of Hamming distances of recovered DAGs using IGSP and perfect-IGSP (p-IGSP) for 10-node graphs under single-variable (a) perfect, (b) imperfect, and (c) inhibitory interventions and multi-variable (d) perfect, (e) imperfect, and (f) inhibitory interventions

C. Additional simulation results

C.1. IGSP vs. perfect-IGSP

As described in the main text, for each simulation, we sampled 100 DAGs from an Erdös-Renyi random graph model with an average neighborhood size of 1.5 and $p \in \{10, 20\}$ nodes. The data for each DAG $\mathcal{G}^*$ was generated using a linear structural equation model with independent Gaussian noise: $X = AX + \epsilon$, where A is an upper-triangular matrix with edge weights $A_{ij} \neq 0$ if and only if $i \to j$, and $\epsilon \sim \mathcal{N}(0, Id)$. For $A_{ij} \neq 0$, the edge weights were sampled uniformly from $[-1, -0.25] \cup [0.25, 1]$ to ensure that they are bounded away from zero. We simulated perfect interventions on i by setting the column $A_{,i} = 0$; inhibiting interventions by decreasing $A_{,i}$ by a factor of 10; and imperfect interventions with a success rate of $\alpha = 0.5$. Here, the results are shown for 10-node graphs in which interventions were performed on all single-variable targets (Figure 8), or all pairs of multiple-variable targets (Figure 8).

IGSP performed better on single-variable interventions than on multi-variable interventions (Figure 8). This is expected based on the discussion on Definition 4.2; IGSP requires fewer invariance tests when the data come from single-variable interventions. In contrast, perfect-IGSP (Wang et al., 2017) performs similarly between single-variable and multi-variable interventions; by assuming perfect interventions, perfect-IGSP avoids multiple hypothesis testing when there are multi-variable interventions.

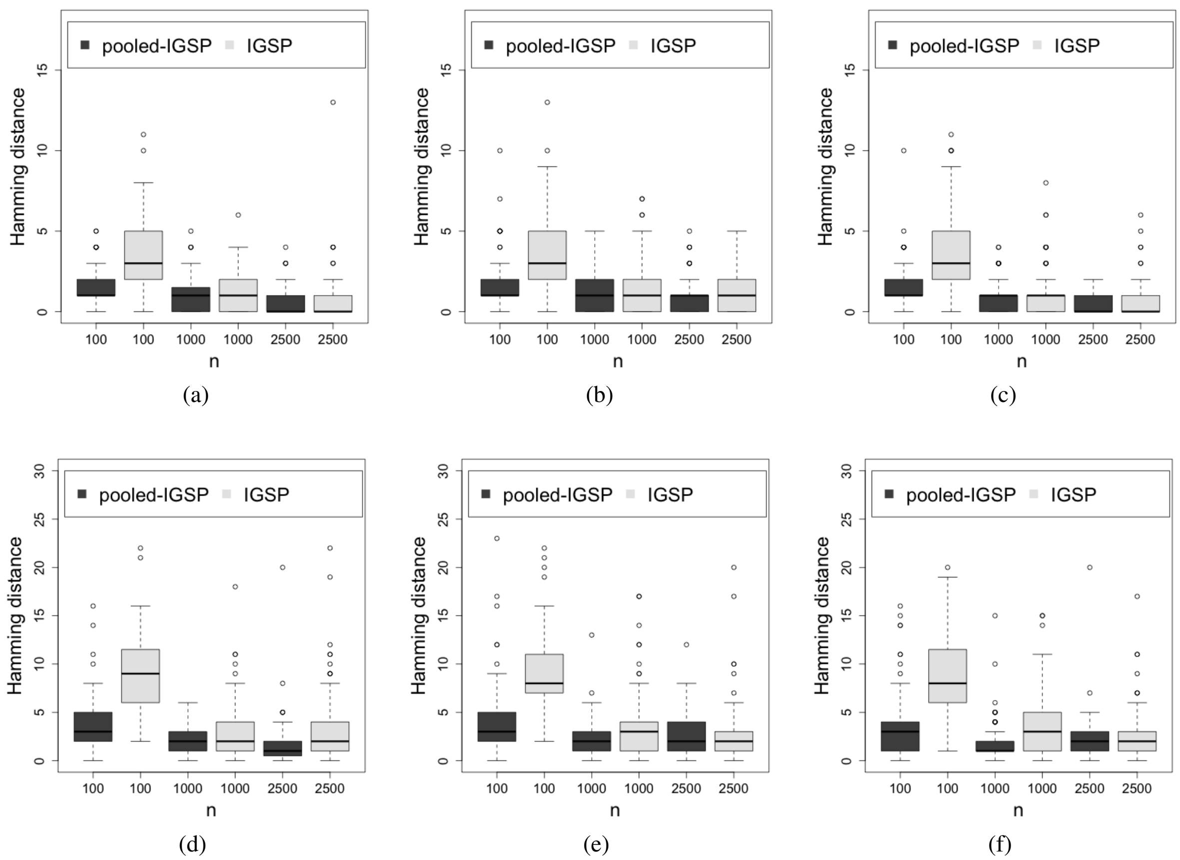

Figure 9. Distributions of Hamming distances of recovered DAGs using IGSP and a heuristic pooled-IGSP for 10-node graphs under (a) perfect, (b) imperfect, and (c) inhibitory interventions and 20-node graphs under (d) perfect, (e) imperfect, and (f) inhibitory interventions

C.2. Pooling

Corollary B.7 described testable conditions under which CI tests can be performed over pooled observational and interventional data in a provably correct way. Here we show that the simple heuristic of pooling all of the datasets for all the CI tests is also effective for improving the performance of IGSP, particularly when the sample sizes are limited. The simulations of Figure 9 compare IGSP to a heuristic version of IGSP, in which all of the data is pooled. However, the limitation of this method is that it is obviously not consistent in the limit of $n \to \infty$.