A Proofs

A.1 Proof of Proposition 3.2

Given infinite samples per intervention $I \in \mathcal{I}$, G^* is recovered up to its $\mathcal{I}$-Markov equivalence class. Hence, the resulting entropy after placing an infinite number of samples at each intervention is equal to $\log_2 |\text{Ess}^{\mathcal{I}}(G)|$ when the true DAG is G. Since the true DAG is unknown, this entropy must be averaged over our prior distribution on $\mathcal{G}$, which is uniform. Hence, the entropy after observing an infinite number of samples per intervention in $\mathcal{I}$ equals $\frac{1}{|\mathcal{G}|}\sum_{G\in\mathcal{G}} \log_2 |\text{Ess}^{\mathcal{I}}(G)|$. Minimizing this entropy over all possible interventions sets of size at most K completes the proof.

A.2 Proof of Theorem 3.4

Let

$$\mathcal{I}^\infty := \{I \in \mathcal{I}^* : \sum_{b=1}^{\infty} |\tilde{I} \in \xi_b : \tilde{I} = I| = \infty \ \mu^* a.s.\},$$

where ξ_b denotes the interventions selected at batch b by $U^f_{\text{M.I}}$. Since $\mathcal{I}^*$ is finite, $\mathcal{I}^\infty$ is non-empty. When $|\mathcal{I}^\infty| > 1$, $\mathcal{I}^\infty$ is a conservative family of targets since $\mathcal{I}^*$ is a family of single-node interventions. Hence, we identify the $\mathcal{I}^\infty$-MEC of G^* in the limit of an infinite number of batches and samples (Hauser and Bühlmann, 2012). Assume $|\mathcal{I}^\infty| > 1$. If $f(G)$ is identifiable in $\text{Ess}^{\mathcal{I}^\infty}(G^*)$, then

$$\mathbb{P}\left(f(G) \mid D_B\right) \xrightarrow{\mu^* \text{ a.s.}} \mathbb{1}(f(G) = f(G^*)).$$

Hence, it suffices to show that the interventions $U^f_{\text{M.I}}$ selects infinitely often identifies $f(G)$ in the limiting interventional essential graph $\text{Ess}^{\mathcal{I}^\infty}(G^*)$. Suppose towards a contradiction that $f(G)$ were not fully identifiable in $\text{Ess}^{\mathcal{I}^\infty}(G^*)$. By definition of almost sure convergence, there exists some $b^* < \infty$ such that any $\tilde{I} \in \mathcal{I}^* \setminus \mathcal{I}^\infty$ is never selected again after batch b^* with probability one since $\mathcal{I}^*$ is finite. Maximizing $U^f_{\text{M.I.}}$ is equivalent to minimizing the conditional entropy,

$$H^b_\xi(f \mid Y_\xi) := \mathbb{E}_{y\sim\mathbb{P}(y|D_b,\xi)}\ H(f \mid D_b, Y = y). \quad (16)$$

If $b > b^*$, then

$$\underset{\xi\in\mathbb{Z}^{\mathcal{I}^*}\cap C_b}{\arg\min}\ H^b_\xi(f \mid Y_\xi) = \underset{\xi\in\mathbb{Z}^{\mathcal{I}^\infty}\cap C_b}{\arg\min}\ H^b_\xi(f \mid Y_\xi) \quad (17)$$

since any batch b after b^* never selects an intervention in $\tilde{I} \in \mathcal{I}^* \setminus \mathcal{I}^\infty$. Since f is not identifiable in $\text{Ess}^{\mathcal{I}^\infty}(G^*)$, that implies

$$\lim_{b\to\infty} H^b_{\xi_\infty}(f \mid Y_{\xi_\infty}) \to L > 0.$$

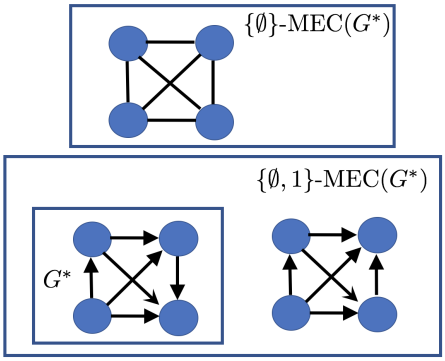

Figure 5: Each box represents the members of the interventional Markov equivalence classes. For G^* given in the bottom left box, the observational Markov equivalence class has no edges oriented. The top box represents the essential graph of the observational Markov equivalence class. The interventional Markov equivalence class for an intervention at node one consists of two DAGs given in the bottom box.

Since $\mathcal{I}^*$ consists of all single-node interventions, $\mathcal{I}^*$ can identify $f(G)$ (Hauser and Bühlmann, 2012). Hence, there must be some $\tilde{I} \in \mathcal{I}^* \setminus \mathcal{I}^\infty$ and $\epsilon > 0$ such that

$$\lim_{b\to\infty} H^b_{\xi_\infty \cup \tilde{I}_\infty}(f) < L - \epsilon, \quad (18)$$

where $\tilde{I}_\infty$ denotes selecting $\tilde{I}$ infinitely many times. But Eq. (18) implies that there must exist some batch $b > b^*$ such that the conditional entropy of the design $\tilde{\xi} = \{\tilde{I}\}$ is uniformly smaller than the conditional entropy of any $\xi \in \mathbb{Z}^{\mathcal{I}^\infty}$. But this is a contradiction because then $\tilde{I}$ would be selected again after some batch $b > b^*$ and Eq. (17) would no longer hold.

For $|\mathcal{I}^\infty| = 1$, we no longer have a conservative family of targets. However, a nearly identical argument works by noting that, in the limit, we learn the observational equivalence class of the $\mathcal{I}^\infty$ mutilated graph of G^*.

A.3 Consistency Counterexample

Suppose we know the Markov equivalence class of G^* and the goal is to fully recover G^*. Suppose $C_b = \{\xi : \|\xi\|_0 = K\}$, where $\|\cdot\|_0$ counts the number of unique interventions in ξ. Since there is no constraint on the number of samples, only on the number of unique interventions, we may allocate an infinite number of samples per intervention within each batch. This constraint is equivalent to the one examined in Ghassami et al. (2018). The scores in both Ness et al. (2018) and Ghassami et al. (2018) select interventions by maximizing the expected number of oriented edges in the interventional Markov equivalence classes. In particular, the utility function in Ness et al. (2018) is

equivalent to maximizing,

$$U(\mathcal{I}; D) = \sum_{G \in \mathcal{G}} A(\text{Ess}^{\mathcal{I}}(G))\mathbb{P}(G), \tag{19}$$

where $A(\text{Ess}^{\mathcal{I}}(G))$ equals the additional number of edges oriented relative to the observational Markov equivalence class. Suppose G^* equals the graph in Fig. 5 and that $K = 1$ unique interventions are allowed within each batch. Assume that $\mathcal{I}^* = \{\{1\}, \cdots, \{4\}\}$ and that we start with a uniform prior over $\mathcal{G}$. Then, since all arrows are undirected in the observational Markov equivalence class, symmetry implies $U(\{j\}; \emptyset) = U(\{j\}; \emptyset)$ for all $i, j \in 1, \cdots, 4$. Without any loss of generality suppose intervention one is selected in batch one. We show that every subsequent batch will select intervention $\{1\}$. If only $\{1\}$ were selected, $U(\mathcal{I}; D)$ would not be consistent since the $\{\emptyset, \{1\}\}$-MEC(G^*) contains two graphs, as shown at the bottom of Fig. 5. After batch one, the posterior is supported on these two graphs since an infinite number of samples are allocated to the intervention at node one.

The utility function in Eq. (19) scores interventions relative to the observational equivalence class, which causes the consistency issue. In particular, the posterior in batch two is only supported on the two DAGs given in the bottom box of Fig. 5. The score of $\{1\}$ equals 5 in batch two while the scores of interventions $\{2\}, \{3\}, \{4\}$ equal $4, 3, 4$, respectively. Hence, in batch two, intervention $\{1\}$ will be selected again, but the posterior will remain the same since the $\{\emptyset, \{1\}\}$ interventional Markov equivalence class of G^* is already known.

An easy way to fix Eq. (19) (for this given counterexample) would be to only select interventions not selected in previous batches. This modification would fix the issue with the counterexample, namely prevent intervention one from being selecting infinitely often. However, when one can only allocate a finite number of samples per batch, this modification would not lead to a consistent estimator. In particular, if a certain intervention is done in some batch, and that intervention must be conducted in order to identify f, then only placing finitely many samples to that intervention in that batch and never placing any more samples in subsequent batches will not lead to a consistent method.

A.4 Proof of Theorem 4.1

Definition A.1. (Soma and Yoshida, 2016) Let E be a finite set. A function $f : \mathbb{Z}^E \to \mathbb{R}$ is *diminishing returns submodular* (DR-submodular) if for $x \leq y$

$$f(x + \chi_e) - f(x) \geq f(y + \chi_e) - f(y), \; x, y \in \mathbb{Z}^E \tag{20}$$

where $e \in E$ and χ_e is the ith unit vector.

Lemma A.2. *$\tilde{U}^f_{M.I.}(\xi; D)$ is DR-submodular.*

Proof. $f(G) = G$ so we omit f in $\tilde{U}^f_{\text{M.I.}}$ to simplify notation. Since the sum of submodular functions is submodular, it suffices to show

$$\begin{aligned} \mathbb{E}_{y|G,\hat{\theta}^G_{\text{MLE}},\xi} \tilde{U}_{\text{M.I.}}(y, \xi; D) &= H(G) - H(G \mid Y_\xi) \\ &= I((G, \hat{\theta}^G_{\text{MLE}}), Y_\xi) \end{aligned} \tag{21}$$

is DR-submodular, where I is the mutual information. Consider an $A \subseteq B \in \mathbb{Z}^{\mathcal{I}^*}$. Take any $C \in \mathcal{I}^*$. Since entropy decreases with more conditioning,

$$\begin{aligned} H(Y_C \mid Y_A) - H(Y_C \mid (G, \hat{\theta}^G_{\text{MLE}})) \geq \\ H(Y_C \mid Y_B) - H(Y_C \mid (G, \hat{\theta}^G_{\text{MLE}})). \end{aligned} \tag{22}$$

By conditional independence,

$$\begin{aligned} H(Y_C \mid (G, \hat{\theta}^G_{\text{MLE}})) &= H(Y_C \mid (G, \hat{\theta}^G_{\text{MLE}}), Y_A) \\ &= H(Y_C \mid (G, \hat{\theta}^G_{\text{MLE}}), Y_B). \end{aligned} \tag{23}$$

Hence, Eq. (22) may be rewritten as,

$$\begin{aligned} I((G, \hat{\theta}^G_{\text{MLE}}), Y_C \mid Y_A) = \\ H(Y_C \mid Y_A) - H(Y_C \mid (G, \hat{\theta}^G_{\text{MLE}}), Y_A) \geq \\ H(Y_C \mid Y_B) - H(Y_C \mid (G, \hat{\theta}^G_{\text{MLE}}), Y_B) = \\ I((G, \hat{\theta}^G_{\text{MLE}}), Y_C \mid Y_B). \end{aligned} \tag{24}$$

Eq. (24) implies

$$\begin{aligned} I((G, \hat{\theta}^G_{\text{MLE}}), Y_A + Y_C) - I((G, \hat{\theta}^G_{\text{MLE}}), Y_A) \\ \geq I((G, \hat{\theta}^G_{\text{MLE}}), Y_B + Y_C) - I((G, \hat{\theta}^G_{\text{MLE}}), Y_B) \end{aligned} \tag{25}$$

as desired. □

The proof of Theorem 4.1 then follows directly from Lemma A.2 and Soma et al. (2014, Theorem 2.4).

A.5 Proof of Proposition 4.2

For each graph $G \in \mathcal{G}_T$, compute the associated edge weights $\hat{\theta}^G_{\text{MLE}}$. Computing each $\hat{\theta}^G_{\text{MLE}}$ takes $O(p\kappa^3)$ time using the formula given in Hauser and Bühlmann (2012, pg. 17). Since there are T DAGs, the total time to compute the MLE estimates of the edge weights of each DAG is $O(Tp\kappa^3)$. Sampling from a multivariate Gaussian with bounded indegree with known adjacency matrix takes $O(p\kappa)$ time. $\hat{U}^f_{\text{M.I.}}$ requires a total of $|\mathcal{I}^*|MN_bT^2$ samples. Hence, the total computation time of sampling all the y_{mt} in Eq. (14) is $O(|\mathcal{I}^*|MN_b\kappa pT^2)$. Evaluating $\hat{U}^f_{\text{M.I.}}$ takes $O(MT^2)$ time using these samples, which is of lower computational complexity than computing $\hat{U}^f_{\text{M.I.}}$. Hence, the total runtime is $O(p\kappa^3 + |\mathcal{I}^*|MN_b\kappa pT^2)$.

A.6 Constraint on the Number of Unique Interventions

If we are only allowed to allocate at most K unique interventions per batch, we modify Algorithm 3 by allocating $\frac{N_b}{K}$ samples per intervention in Algorithm 2. Once an intervention is selected, that intervention is removed from I^* and another one is greedily selected from the remaining set. With this strategy, Algorithm 2 will terminate after K iterations. Hence, there will be at most K unique interventions as desired.

A.7 DREAM4 Supplementary Figures

We applied our targeted experimental design strategy towards learning the *downstream pathways* of select genes from a 10-node network from the DREAM4 challenge. We observed a modest improvement over the random strategy for some central genes in the network (Fig. 6, top). However, the results are subject to high variations (Fig. 6, bottom), which we surmise to be due to the small size of the observational dataset. Nevertheless, these preliminary results illustrate the promise of applying targeted experimental design to real, large-scale biological datasets.

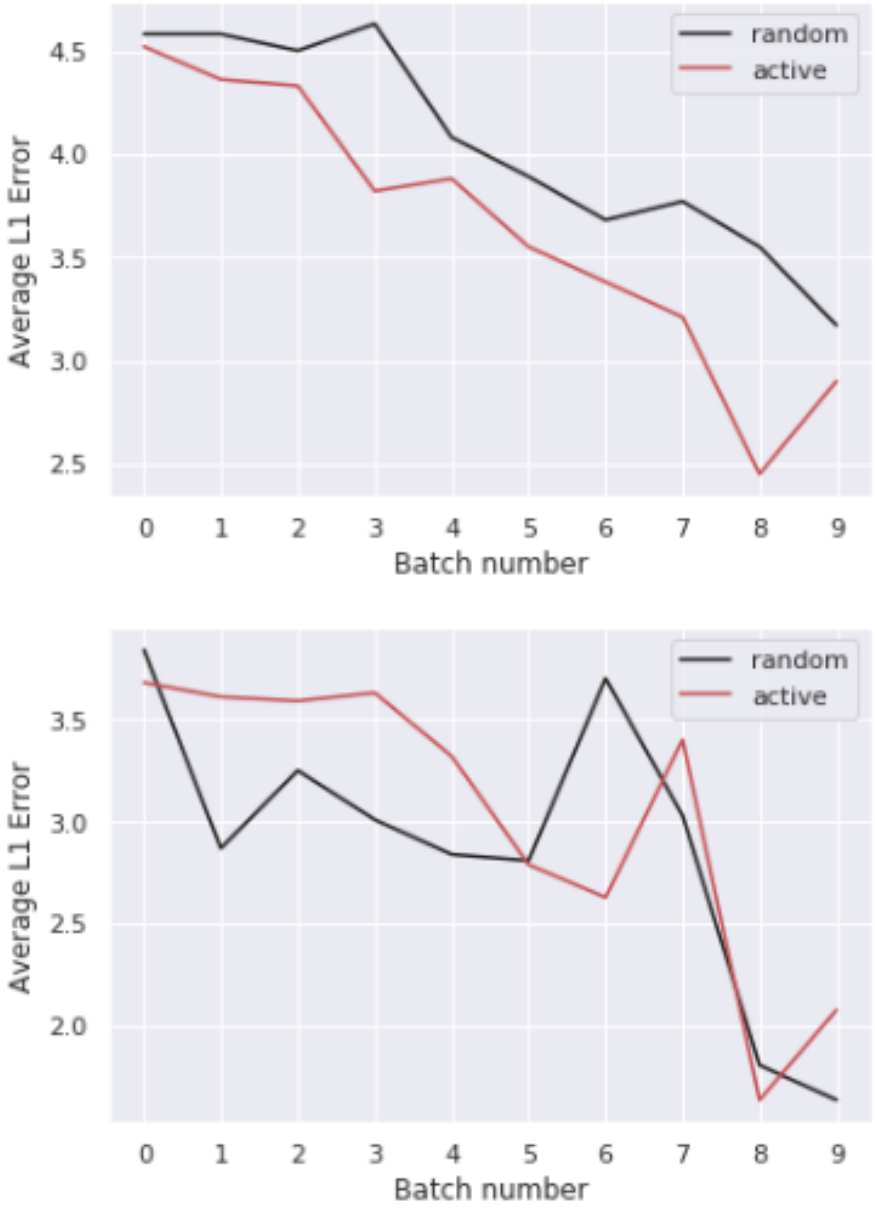

Figure 6: Performance of intervention strategies on predicting the descendants of genes 6 (top) and 8 (bottom).