Learning Unknown Intervention Targets in Structural Causal Models from Heterogeneous Data

Yuqin Yang*[1] Saber Salehkaleybar*[2] Negar Kiyavash[2]

[1]Georgia Tech [2]EPFL *Equal contribution

Motivation

Causal structure learning from **interventional data**

- We may not fully control the interv. target
- Intervention is done by an unknown source

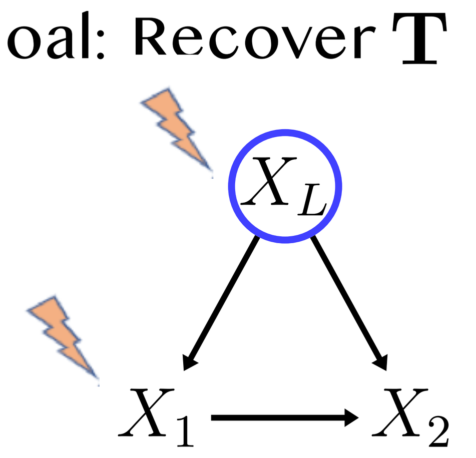
Task: Learn interv. targets from multi-domain data

Existing method: (possibly a byproduct)

- Limited to linear systems
- Requiring exponential CI / invariance tests
- Unable to handle latent confounders

Model Description

- SCM: $X_i = f_i(PA_i, N_i),\ \ X_i \in \mathbf{X}$ not given as input
  - Partitioned into $[\mathbf{O};\mathbf{L}]$ under latent conf.
- Soft intervention: $X_i = f_i(PA_i, N_i')$
- We collect data from D domains
  $\mathbf{T} := \{X_i | \exists d, d' \in [D], p_d(N_i) \neq p_{d'}(N_i)\}$
- Goal: Recover $\mathbf{T}$ $(\mathbf{T}\cap\mathbf{O})$



- Two obs. var., one latent
- Two noises change across environments

We propose **Locating Intervention Target (LIT)** algorithm, which includes Recovery phase and Matching phase.

Recovery Phase

Recover the noises $\mathbf{N_T} = \{N_i | X_i \in \mathbf{T}\}$ up to permutation and component-wise invertible transformations using contrastive learning approach.

- Mixing function: $\mathbf{X} = \mathbf{g}(\mathbf{N})$
- Auxiliary / domain variable U

Proposition 1: Assume $\min(D-1, |\mathbf{O}|) \geq |\mathbf{T}|$. Under certain conditions on $\mathbf{N}$, the recovery is possible when:
a) $\mathbf{L} = \emptyset$ and $\mathbf{g}$ is invertible;
b) $\mathbf{L} \neq \emptyset$ and $\exists$ invertible $\tilde{\mathbf{g}} : \mathbb{R}^{|\mathbf{O}|} \to \mathbb{R}^{|\mathbf{O}|}$ s.t. $\tilde{\mathbf{g}}(\mathbf{O}) = (\mathbf{N_T}; \mathbf{V})$, where $\mathbf{V} \perp\!\!\!\perp U$ and $\mathbf{V} \perp\!\!\!\perp \mathbf{N_T} | U$.

Invertibility holds when the model is a linear SCM, nonlinear ANM, or $\{f_i\}$ are MLPs with ReLU activation function and positive coefficients.

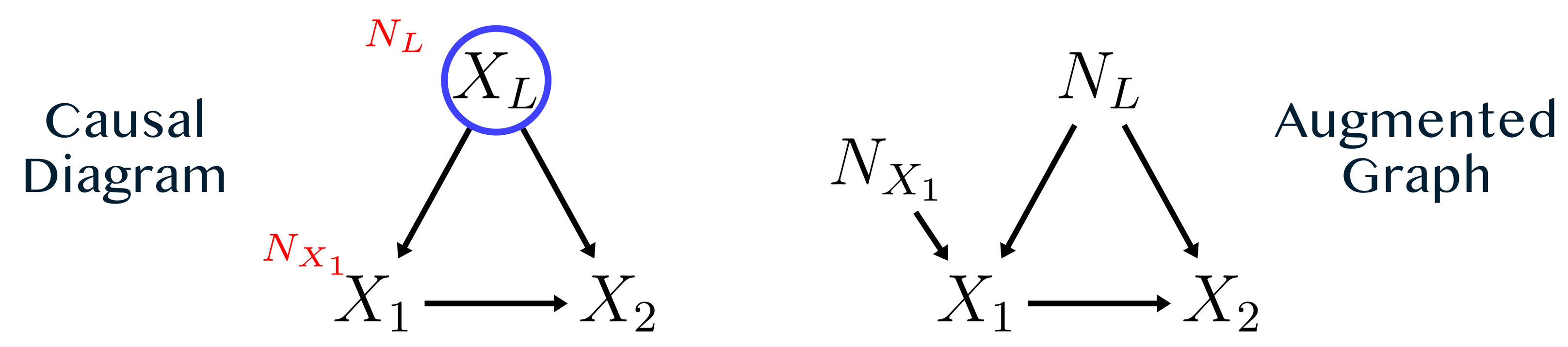
Matching Phase

Match the recovered noises in $\tilde{\mathbf{N}}_\mathbf{T}$ to $\mathbf{N_T}$ by comparing between $\tilde{\mathbf{N}}_\mathbf{T}$ and $\mathbf{X}$ (or $\mathbf{O}$)

T-faithfulness assumption

d-separation between noise and observed variable on the augmented graph is equivalent to independency.

Matching Phase (Cont'd)



- **Indicator set** $\mathcal{I} : \mathbf{I}_i = \{\tilde{N}_j | \tilde{N}_j \not\perp\!\!\!\perp X_i\}$
  - Includes all noises in $An(X_i) \cap \mathbf{N_T}$

Matching under causal sufficiency

Theorem 1: The intervention targets can be uniquely identified based on $\tilde{\mathbf{N}}_\mathbf{T}$, $\mathbf{X}$ and $\mathcal{I}$ using three conditions.

- LIT: Checking Cond (I) - (III) for all variables
- Requires quadratic CI tests: Bounded by $|\mathbf{T}| \cdot |\mathbf{X}|^2$

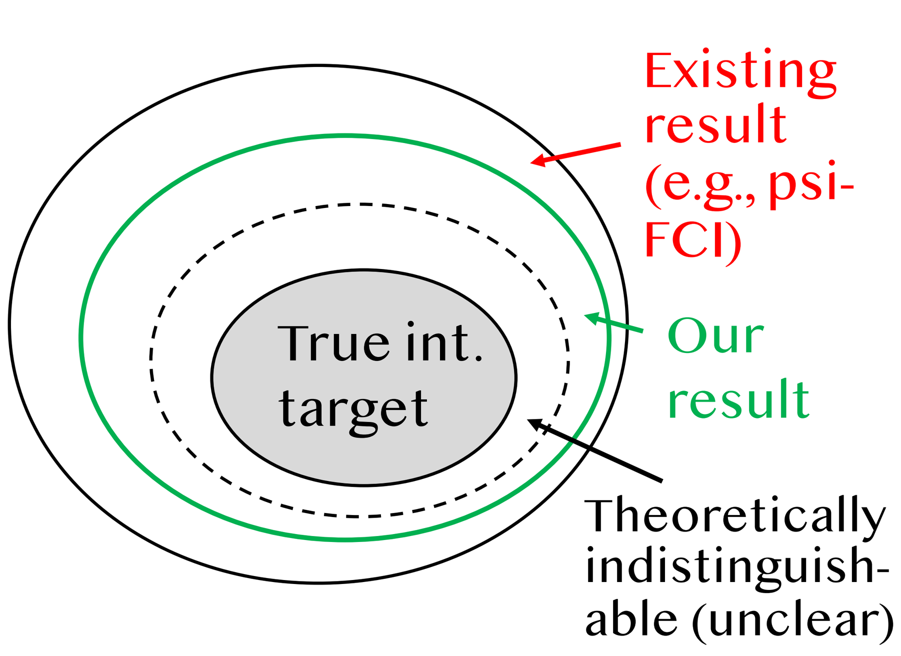
Algorithm 1: LIT algorithm

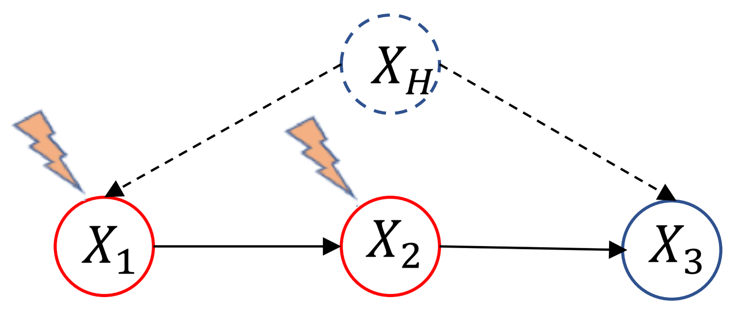
```
1 Obtain Ñ_T and I; U ← X; K ← ∅;
2 for X_i ∈ X do
3     if (I) holds then remove X_i (from U);
4     else if (IV) holds then remove X_i; // latent
5     else if (II) holds then add X_i to K, remove X_i;
6 Partition U into disjoint subsets U_1, ··· , U_r
    according to the indicator sets;
7 for U_i ∈ {U_1, ··· , U_r} do
8     Add X_{i_k} ∈ U_i satisfying (III) to K (resp.
      variables not satisfying (III-L));
9 return K
```

Matching under latent confounding

Theorem 2: By adding Cond. (IV) and changing (III), LIT can return a superset of the true intervention targets.

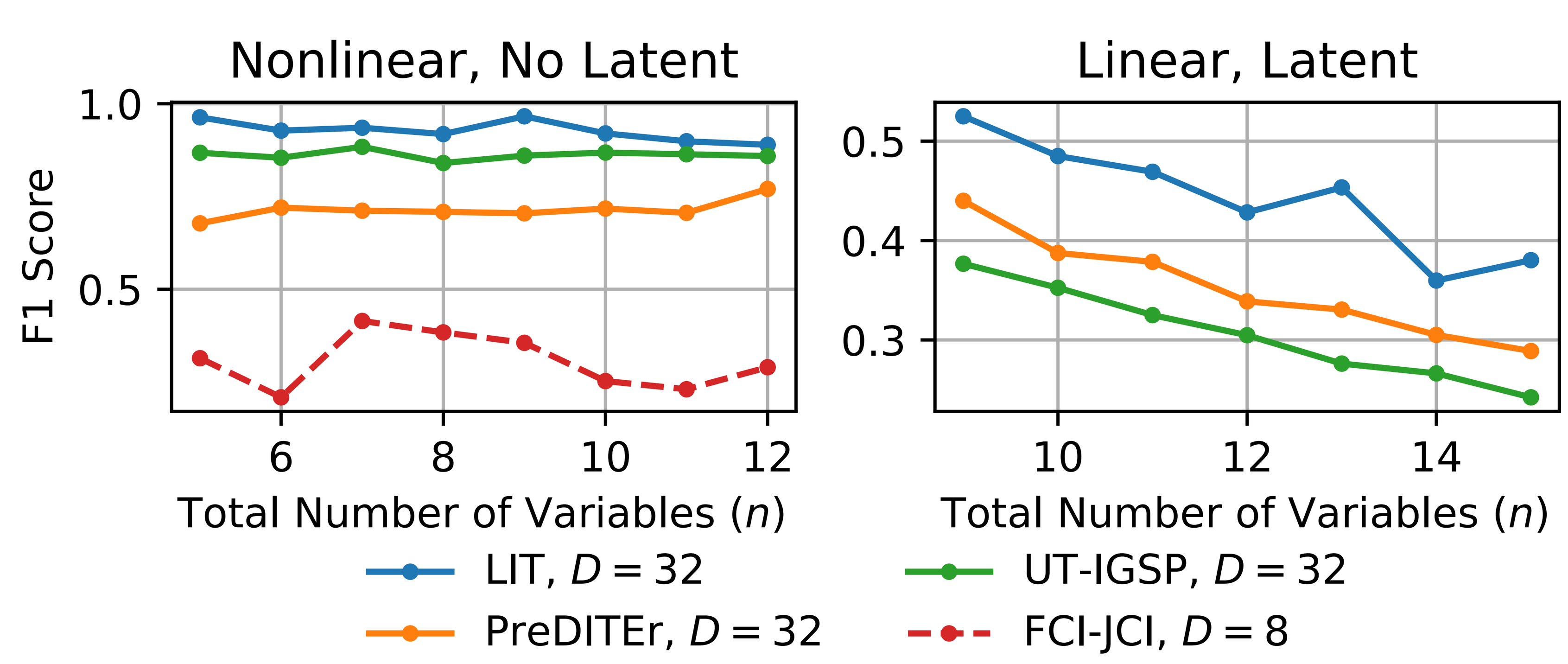
- Graphical characterization: Auxiliary graph
- Can handle latent intervention targets, i.e., $\mathbf{T}\cap\mathbf{L} \neq \emptyset$
- More informative than baselines when $\mathbf{T}\cap\mathbf{L} = \emptyset$





- LIT: Recovers the targets
- $\mathbf{I}_1 = \{\tilde{N}_1\}, \mathbf{I}_2 = \mathbf{I}_3 = \{\tilde{N}_1, \tilde{N}_2\}$
- $\tilde{N}_2 \perp\!\!\!\perp X_3 | X_1, X_2$
- Existing work (psi-FCI): Returns a superset

Simulations



- Compare the recovery of the intervention targets
- PreDITEr: linear-Gaussian; UT-IGSP: causal sufficiency
- # of CI tests: LIT ~80; PreDITEr ~30000