Adaptive Causal Dimensionality Reduction

Hedong Yan, CS, HKBU

Supervisor: Prof. Yiuming Cheung

Motivation

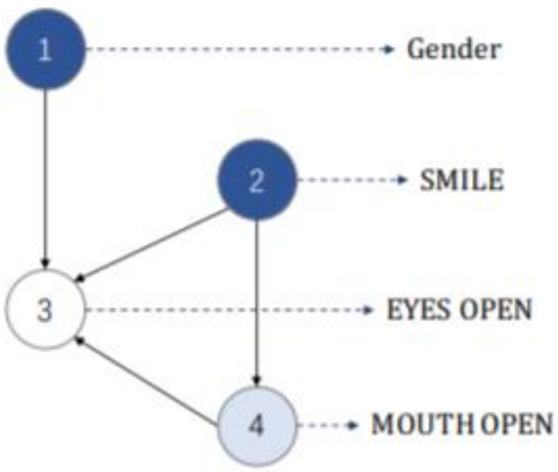
- Dimensionality reduction
  - Curse of dimensionality & visualization
- Causal learning and inference
  - Causal variable & structure
- What’s the most suitable scale for causal learning and inference?

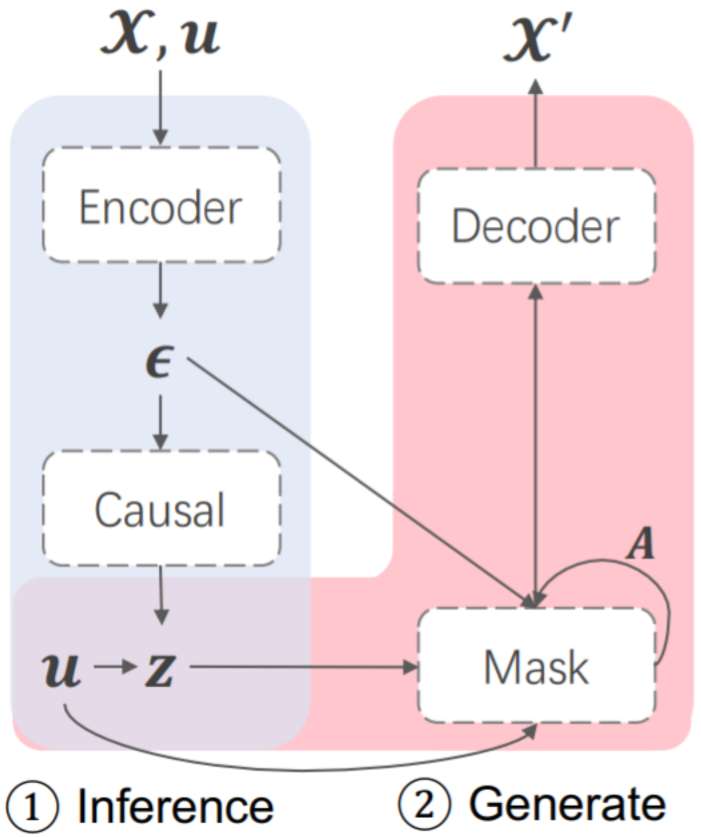
Motivation

- Higher dimension: more redundancy
- Lower dimension: lose more ‘causal’ information
- There is a trade-off

Literature survey

- Dimensionality reduction
  - PCA (linear, uncorrected)
  - VAE (nonlinear, reconstruction error)
  - Causal VAE (labels causal structure)





- **Weakness: can not tell us the proper dimension K***

Literature survey

SCM-based identification

1993 Pearl's do-calculus

2002 Tian's c-factorization

2006 Huang's ID and Shpitser's ID

2006 Shpitser's IDC

2012 zID, sID

2015 data fusion (ob, exp, biased, and dissimilar)

2019 ID in segregated graph

2019 gID

2020 graph with loop

from observation to intervention

from intervention to intervention

Others

1980 SUTVA

1984 strong ignorability

2019 single strong ignorability

2019 po-calculus and related algorithm

2020 sequential single strong ignorability

2021 matrix ID

2021 approximation

2021 neural ID

RCM-based identification

- **Weakness: diagram and distribution are known**

Literature survey

- Independence-based algorithms
  - PC, FCI
- Mechanism assumption
  - LiNGAM: Linear model with additive non-Gaussian noise
  - Post-NonLinear: $x_i = f_{i,2}\left(f_{i,1}(pa_i) + e_i\right)$ where $i = 1, \ldots, n$

invertible nonlinear

- **Weakness: the mechanism and noise assumptions may not be true**

Literature survey

- Benchmark and dataset for causal learning and reasoning

TABLE I: Causal Dataset

Type	Name	Introduction	website
Benchmark	Causeme [54]	time-series	https://causeme.uv.es/
Benchmark	JustCause [55]	support IHDP, ACIC etc.	https://justcause.readthedocs.io/en/latest/
Benchmark	e-CARE [56]	reasoning and explanation for NLP	https://scir-sp.github.io
Dataset	IHDP [49]	home visits and IQ testing	https://www.icpsr.umich.edu/web/HMCA/studies/9795
Dataset	Twins [57]	birth weight and mortality	\
Dataset	Jobs [58]	real world data	\
Dataset	ACIC2019	conference challenge	https://sites.google.com/view/acic2019datachallenge/home

Literature survey

- Packages for causal learning and reasoning

TABLE II: Causal Packages

Motivation	Toolbox	Support Team	Introduction
Causal Learning	causal-learn	CMU, DMIR, Gong Mingming team, Shouhei Shimizu team	python version of Tetrad
	Tetrad [59]	CMU	Java
	CausalDiscoveryToolbox [60]	FenTechSolutions	python, DAG/Pair, dataset, independence, structure learning, metrics
	gCastle	Huawei Noah	python, data generation and process, causal structure learning, metrics
	tigramite	Jakob Runge	python, learning from time-series data
Causal Reasoning	Ananke [61] [62] [63]	Ilya Shpitser team	python, support do-calculus
	EconML [64]	Microsoft	python, Econometrics
	dowhy [65]	Microsoft	python
	causalml [66]	Uber	python, campaign target optimization, personalized engagement
	CausalImpact	Google	R, time-series, adertisement and click
	WhyNot	John Miller	python, simulator and environment
	Causal-Curve [67]	Kobrosly, R.W.	python, continuous variable such as price, time and income
	grf [68]	grf-lab of Standford	R
	dosearch [69]	Santtu Tikka	R
	causaleffect [70]	Santtu Tikka	R
	dagitty [71]	\	R, support adjustment formula
End-to-End	causalnex	QuantumBlack	python, 0.11.0, structure learning, domain knowledge, estimation

Research questions

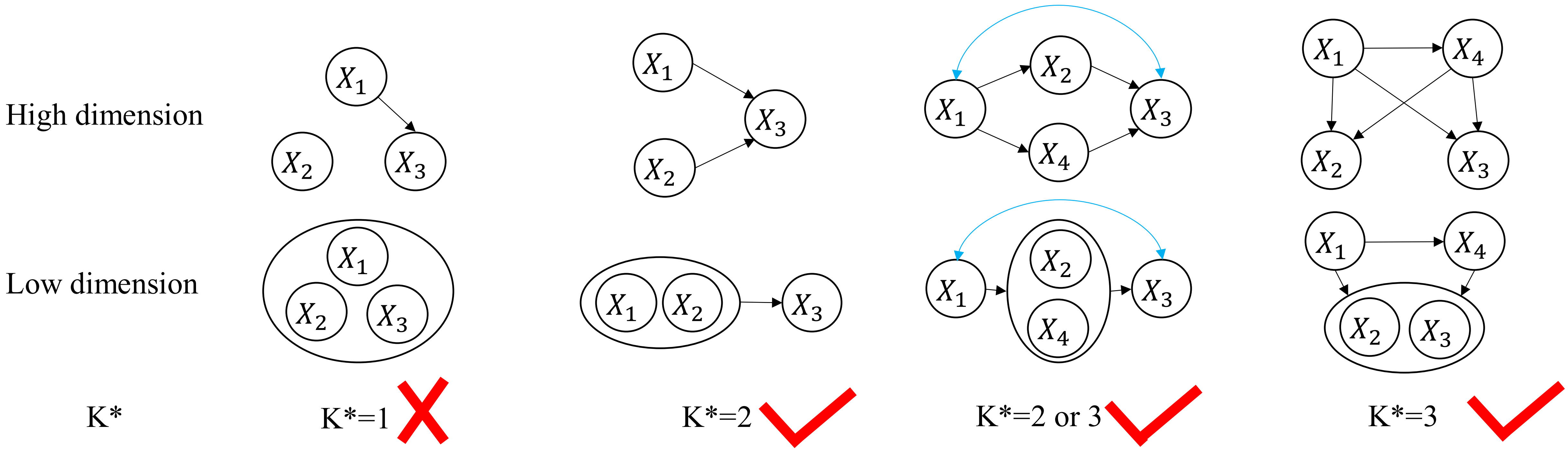
- How to efficiently compute ‘causal’ information and find the optimum scale **K*** of low dimensional representation?
  - What if no causal sufficiency, causal faithfulness, distribution, and causal diagrams
  - ADMG $\Theta\left(2^{n^2-n} * n!\, 1.3^{n^2}\right)$
- How to learn the encoding model for this most suitable scale **K***?
  - PCA
  - VAE

Proposal for adaptive methods

- Find the optimum scale **K***
  - Causal discovery + merge symmetric variables
  - All non-parametric causal models + identification + parametrization + testing
    - Symmetry: $I_{do(x_i)}(x_j) = I_{do(x_j)}(x_i)$ and $I_{do(x_k)}(x_i) = I_{do(x_k)}(x_j)$ and $I_{do(x_i)}(x_k) = I_{do(x_j)}(x_k)$ for any k

Proposal for adaptive methods

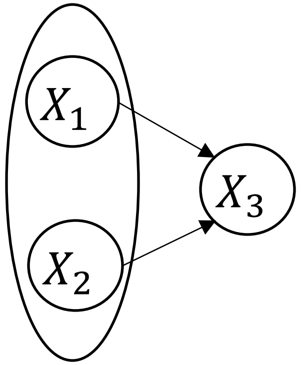
- Causal discovery + merge symmetric variables



Proposal for adaptive methods

- All non-parametric causal models + identification + parametrization + testing

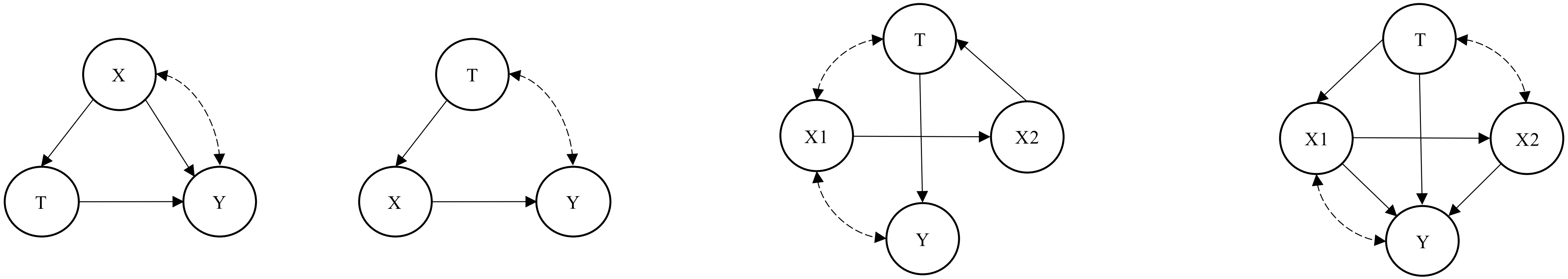
$M = 3, N_H = 5$
$p(X_j)$
$p(X_j \vert X_i)$
$\sum_{x_k} p(x_k) * p(X_j \vert X_i, x_k)$
$\sum_{x_k} p(x_k \vert X_i) \sum_{x_i} p(x_i)\, p(X_j \vert x_i, x_k)$
Not identifiable



for h=1, .., 5	$p_{do(x_1)}(x_j)$	$p_{do(x_2)}(x_j)$	$p_{do(x_3)}(x_j)$
$p_{do(x_i)}(x_1)$	1//1//1//1//1	?//?//?//?//?	?//?//?//?//?
$p_{do(x_i)}(x_2)$	?//?//?//?//?	1//1//1//1//1	?//?//?//?//?
$p_{do(x_i)}(x_3)$	?//?//?//?//?	?//?//?//?//?	1//1//1//1//1

h	$I_{do(x_1)}(x_j)$	$I_{do(x_2)}(x_j)$	$I_{do(x_3)}(x_j)$
$I_{do(x_i)}(x_1)$	0	1	1
$I_{do(x_i)}(x_2)$	1	0	1
$I_{do(x_i)}(x_3)$	2	2	0

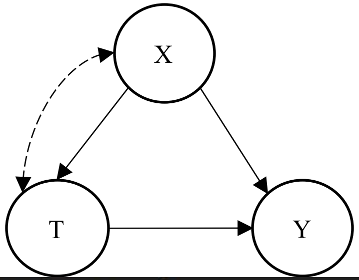
Results of identification

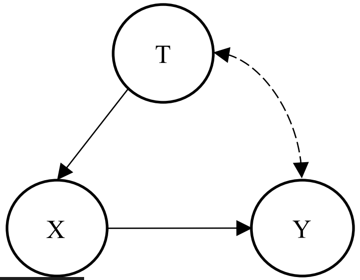


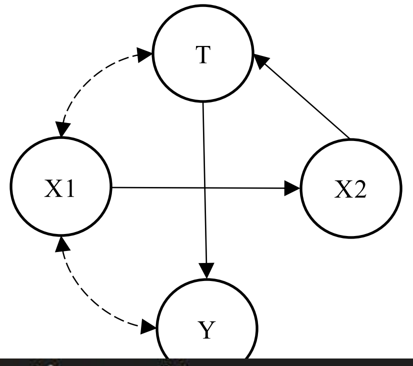
We did not find **correct open-source** codes (including causaleffect, ananker, dowhy, cee, dagitty) to provide the identified latex **expression** of Shpitser's **complete** ID algorithm. We implement the algorithm in python. Some non-trivial running results will be given.

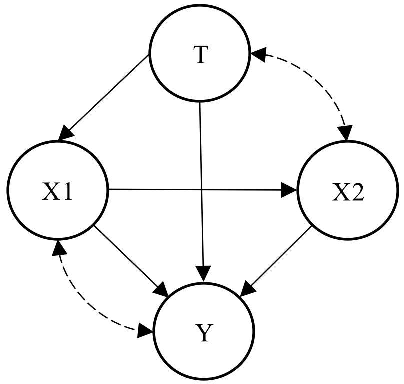
Results of identification

$P(Y|do(T))$










```
identification result:
 \sum_{X}{\sum_{T,Y}{p(X,T,Y)}*\frac{p(X,T,Y)}{\sum_{Y}{p(X,T,Y)}}}
```

We did not find … dagitty) to provide the final identified **expr**… python. Here are some non-trivial running results.

```
identification result:
 \sum_{X}{\frac{\sum_{Y}{p(X,T,Y)}}{\sum_{X}{\sum_{Y}{p(X,T,Y)}}}*\sum_{T}{\frac{\sum_{X,Y}{p(X,T,Y)}*p(X,T,Y)}{\sum_{Y}{p(X,T,Y)}}}}
```

```
identification result:
 \frac{\sum_{X1}{\frac{\sum_{X2,T,Y}{p(X1,X2,T,Y)}*p(X1,X2,T,Y)}{\sum_{T,Y}{p(X1,X2,T,Y)}}}}{\sum_{Y}{\sum_{X1}{\frac{\sum_{X2,T,Y}{p(X1,X2,T,Y)}*p(X1,X2,T,Y)}{\sum_{T,Y}{p(X1,X2,T,Y)}}}}}
```

```
identification result:
 \sum_{X1,X2}{\frac{\sum_{Y,X2}{p(X1,X2,T,Y)}*p(X1,X2,T,Y)}{\sum_{X1,Y,X2}{p(X1,X2,T,Y)}*\sum_{Y}{p(X1,X2,T,Y)}}*\sum_{T}{\frac{\sum_{X2,X1}{\sum_{Y}{p(X1,X2,T,Y)}}*\sum_{Y}{p(X1,X2,T,Y)}}{\sum_{X2,X1,T}{\sum_{Y}{p(X1,X2,T,Y)}}*\sum_{X2}{\sum_{Y}{p(X1,X2,T,Y)}}}}}
```

$$\sum_X \frac{\sum_Y p(X,T,Y)}{\sum_X \sum_Y p(X,T,Y)} * \sum_T \frac{\sum_{X,Y} p(X,T,Y) * p(X,T,Y)}{\sum_Y p(X,T,Y)}$$

$$\frac{\sum_{X1} \frac{\sum_{X2,T,Y} p(X1,X2,T,Y) * p(X1,X2,T,Y)}{\sum_{T,Y} p(X1,X2,T,Y)}}{\sum_Y \sum_{X1} \frac{\sum_{X2,T,Y} p(X1,X2,T,Y) * p(X1,X2,T,Y)}{\sum_{T,Y} p(X1,X2,T,Y)}}$$

$$\sum_X \sum_{T,Y} p(X,T,Y) * \frac{p(X,T,Y)}{\sum_Y p(X,T,Y)}$$

$$\sum_{X1,X2} \frac{\sum_{Y,X2} p(V) * p(V)}{\sum_{X1,Y,X2} p(V) * \sum_Y p(V)} * \sum_T \frac{\sum_{X2,X1} \sum_Y p(V) * \sum_Y p(V)}{\sum_{X2,X1,T} \sum_Y p(V) * \sum_{X2} \sum_Y p(V)}$$

$$\frac{\sum_{x_1} p(x_1) p(T,Y|x_1,x_2)}{\sum_y \sum_{x_1} p(x_1) p(T,y|x_1,x_2)}$$

$$\sum\nolimits_x p(x) p(Y|x,T) \qquad \sum\nolimits_x p(x|T) \sum\nolimits_t p(t)\, p(Y|t,x)$$

$$\sum\nolimits_{x_1,x_2} [p(x_1|T) p(Y|x_1,x_2,T) \sum\nolimits_t p(t) p(x_2|x_1,t)]$$

Results of parametrization and bias testing

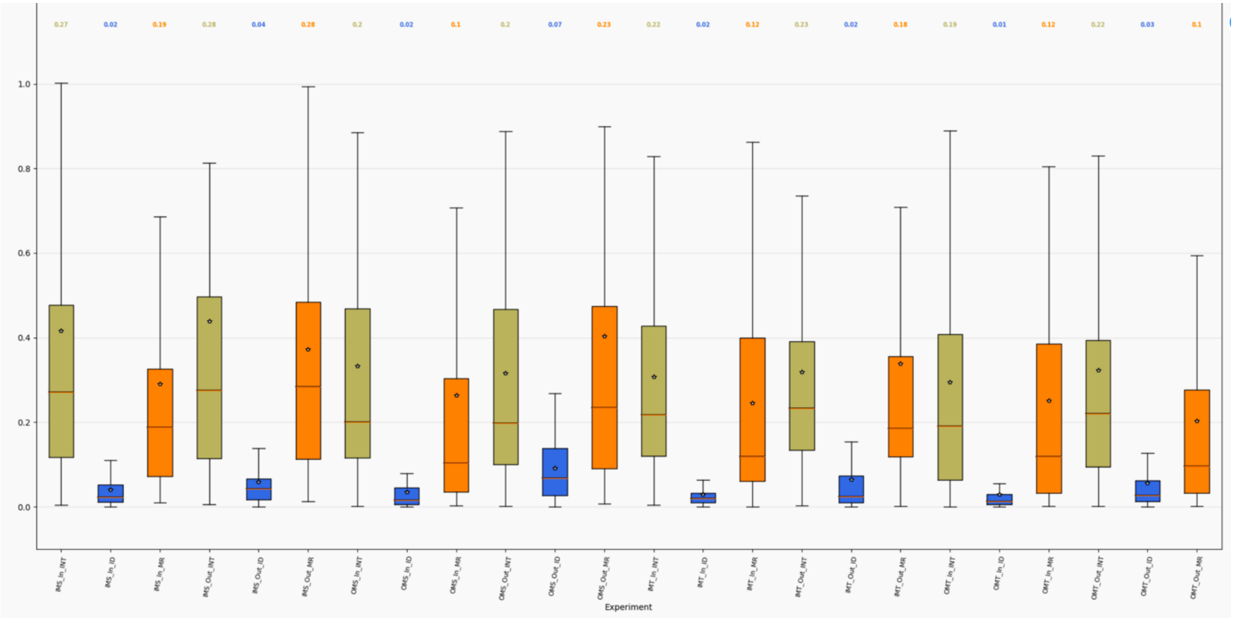


Fig. 3: Experiment error for ATE estimation where X1 is continuous. Star is median value. Red line is average value. 'I' means inner mechanisms, and 'O' means outer mechanisms. 'S' means the parametric intervention is mechanism shifting, and 'T' means the parametric intervention is random transformation of mechanism.

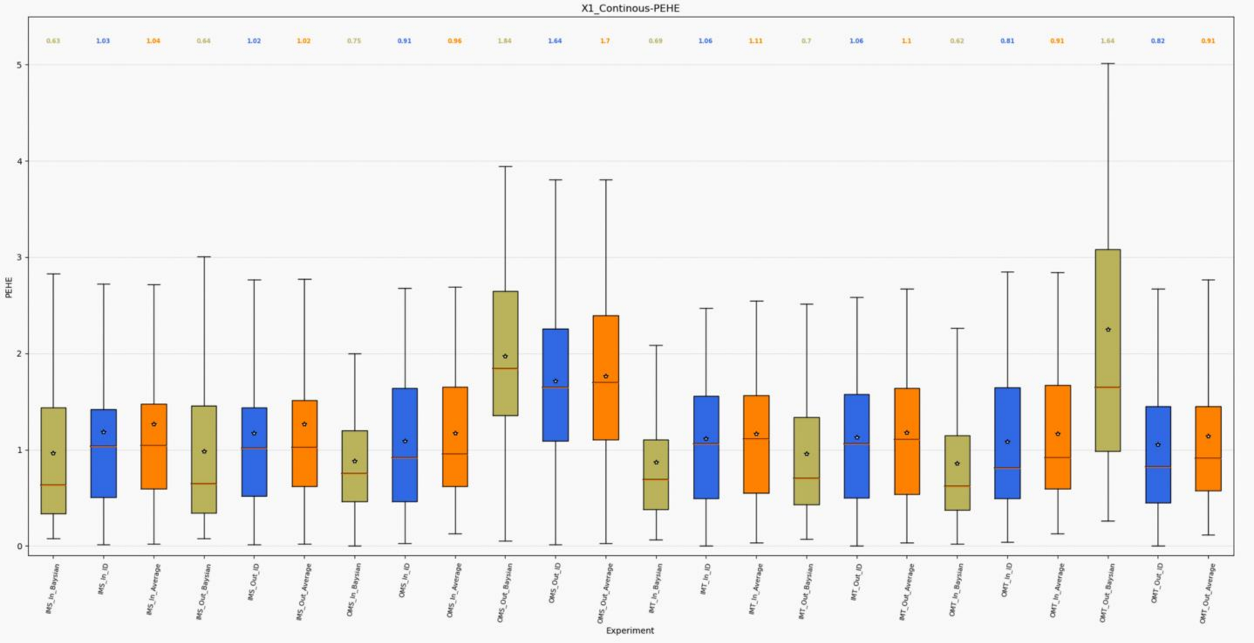
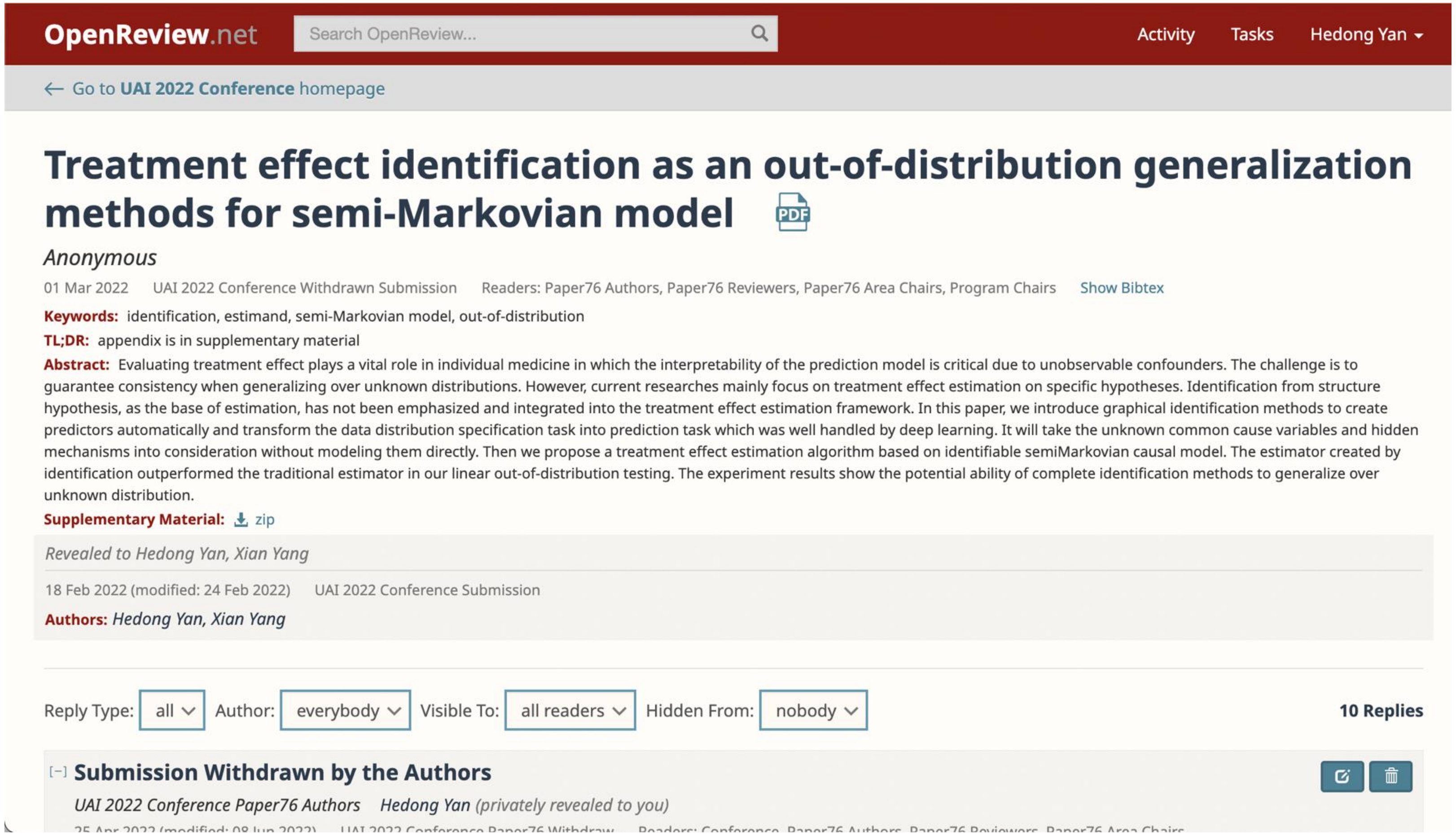


Fig. 5: Experiment error for PEHE estimation where X1 is continuous. Star is median value. Red line is average value. 'I' means inner mechanisms, and 'O' means outer mechanisms. 'S' means the parametric intervention is mechanism shifting, and 'T' means the parametric intervention is random transformation of mechanism.

Identification will not be obviously helpful to reduce MSE of estimation, but it can help you to reduce the bias of causal effect estimation if your diagram is correct.

Papers

- 1 rejected paper



OpenReview.net Search OpenReview... Activity Tasks Hedong Yan

← Go to **UAI 2022 Conference** homepage

Treatment effect identification as an out-of-distribution generalization methods for semi-Markovian model

Anonymous

01 Mar 2022 UAI 2022 Conference Withdrawn Submission Readers: Paper76 Authors, Paper76 Reviewers, Paper76 Area Chairs, Program Chairs Show Bibtex

Keywords: identification, estimand, semi-Markovian model, out-of-distribution

TL;DR: appendix is in supplementary material

Abstract: Evaluating treatment effect plays a vital role in individual medicine in which the interpretability of the prediction model is critical due to unobservable confounders. The challenge is to guarantee consistency when generalizing over unknown distributions. However, current researches mainly focus on treatment effect estimation on specific hypotheses. Identification from structure hypothesis, as the base of estimation, has not been emphasized and integrated into the treatment effect estimation framework. In this paper, we introduce graphical identification methods to create predictors automatically and transform the data distribution specification task into prediction task which was well handled by deep learning. It will take the unknown common cause variables and hidden mechanisms into consideration without modeling them directly. Then we propose a treatment effect estimation algorithm based on identifiable semiMarkovian causal model. The estimator created by identification outperformed the traditional estimator in our linear out-of-distribution testing. The experiment results show the potential ability of complete identification methods to generalize over unknown distribution.

Supplementary Material: zip

Revealed to Hedong Yan, Xian Yang

18 Feb 2022 (modified: 24 Feb 2022) UAI 2022 Conference Submission

Authors: *Hedong Yan, Xian Yang*

Reply Type: all Author: everybody Visible To: all readers Hidden From: nobody **10 Replies**

[–] **Submission Withdrawn by the Authors**

UAI 2022 Conference Paper76 Authors *Hedong Yan* *(privately revealed to you)*

Thanks!