Estimating Multi-cause Treatment Effects via Single-cause Perturbation

Zhaozhi Qian, Alicia Curth, Mihaela van der Schaar

van_der_Schaar \ LAB

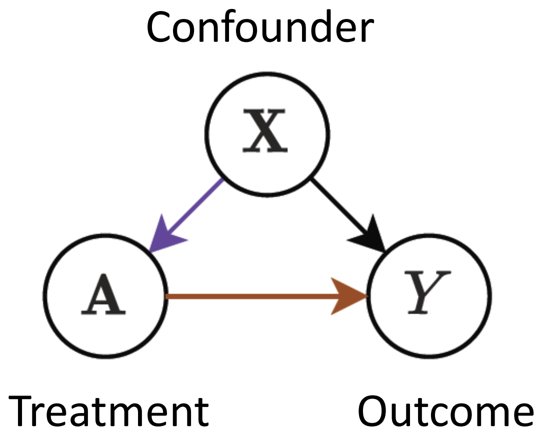
vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

Background: Treatment effect estimation

General Problem: Estimating the conditional average treatment effect (CATE) using observational data

Confounding bias: treatment allocation is not randomized but influenced by confounders.

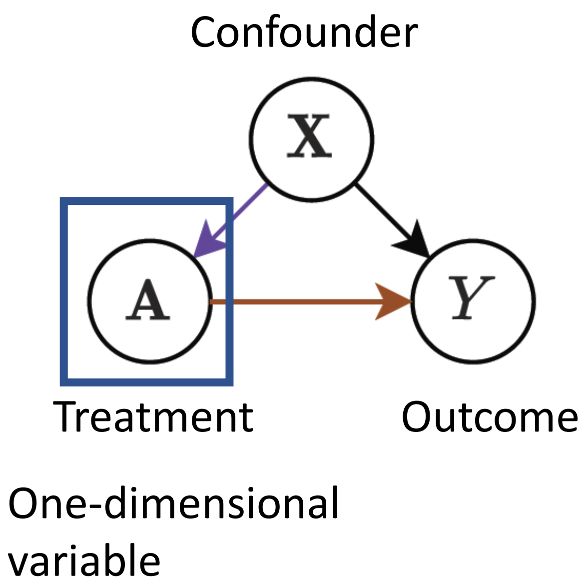


Background: Treatment effect estimation

General Problem: Estimating the conditional average treatment effect (CATE) using observational data

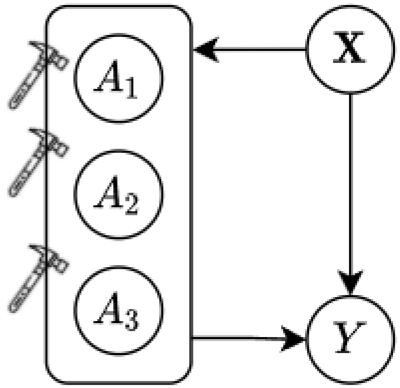
Confounding bias: treatment allocation is not randomized but influenced by confounders.

Limitation: *Single-cause intervention* – most existing methods assume intervention on a single variable



Multi-cause CATE estimation

Problem setup: extension of the single-cause setting



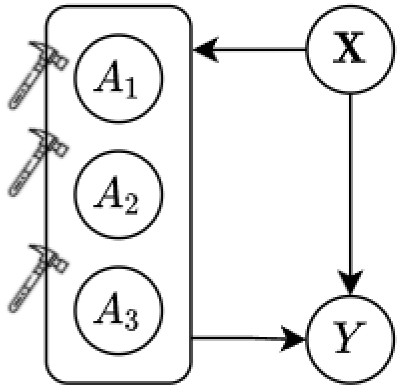
1. Causes: variables can be intervened on
   - Treatment: configuration of all causes
   - Causal structure between the causes

2. Confounders:
   - Influence *any* cause and the outcome

Importance

General

- Many decisions problems involve acting on multiple variables



Healthcare applications

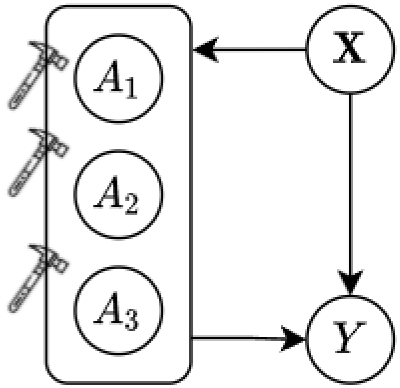
- Treating complex systemic disease
- Care for elder comorbid patients
- Polypharmacy – unnecessary medication

Challenge: combinatorial treatments with confounding

Combinatorial treatments

- K binary causes ➡ 2^K treatments
- Only observe *one* factual outcome

- How to overcome confounding bias?
  - Not just data sparsity in regression



ID		Y(a)								Y
	A_1 A_2 A_3	0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1	
1		?	2	?	?	?	?	?	?	2
2		?	?	?	1	?	?	?	?	1
3		?	?	?	?	4	?	?	?	4

Existing works

Methods for single-cause estimation

- Often fail to scale with combinatorial treatments
  - E.g., multi-head neural network (TARNET, Shalit et al. 2017), propensity score adjustment (Feng et al. 2012)

Methods for multi-cause estimation

- Parametric assumption: linear model (with low-order interaction)
- Latent variable assumption: performing dimensionality reduction
  - Deconfounder (Wang & Blei 2019), VSR (Zou et al. 2020)

Proposed solution: single-cause perturbation

Properties and features

- Leverage existing single-cause estimators as building blocks
- Address confounding by explicit data augmentation
- Easy to implement
- No functional form assumption

- Assumption: causal structure between the causes

Solution: single-cause perturbation

SCP is a two-step procedure

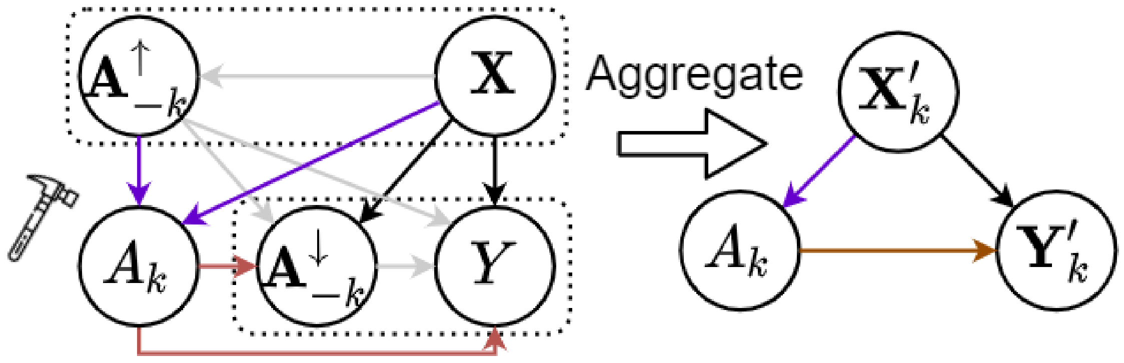
1. Data augmentation: model and predict the single-cause effect
2. Estimate the multi-cause effect on the augmented dataset

ID	A_1 A_2 A_3	Y(**a**) 0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1	Y
1		?	2	?	?	?	?	?	?	2
2		?	?	?	1	?	?	?	?	1
3		?	?	?	?	4	?	?	?	4

Green cells are "imputed" in the first step
As a result, each individual will have K+1 outcomes

Single-cause perturbation: (1) data augmentation

Illustrative causal diagrams:



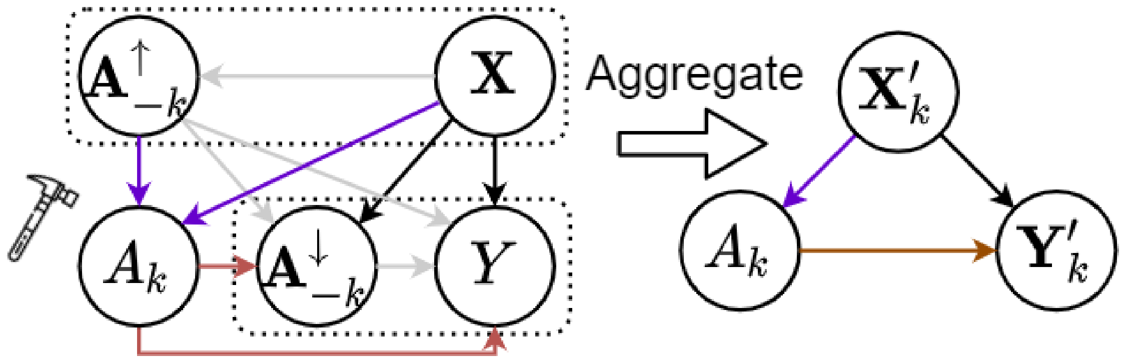
Intervention on A_k

A_k's Non-descendant causes $\mathbf{A}^{\uparrow}_{-k}$

A_k's Descendant causes $\mathbf{A}^{\downarrow}_{-k}$

Single-cause perturbation: (1) data augmentation

Illustrative causal diagrams:



Intervention on A_k

A_k's Non-descendant causes $\mathbf{A}^{\uparrow}_{-k}$

A_k's Descendant causes $\mathbf{A}^{\downarrow}_{-k}$

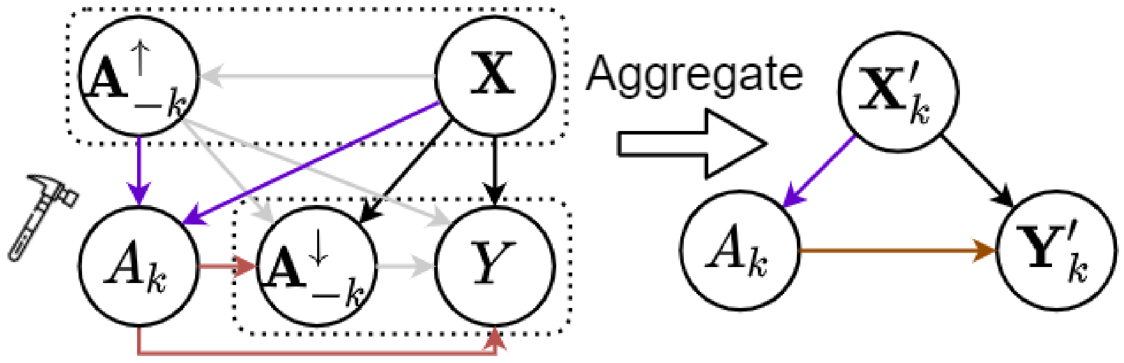
New confounders $\mathbf{X}'_k := (\mathbf{X}, \mathbf{A}^{\uparrow}_{-k})$

New outcomes $\mathbf{Y}'_k := (Y, \mathbf{A}^{\downarrow}_{-k})$

Estimate the counterfactual outcome of flipping A_k

Single-cause perturbation: (1) data augmentation

Illustrative causal diagrams:



Intervention on A_k

A_k's Non-descendant causes $\mathbf{A}^{\uparrow}_{-k}$

A_k's Descendant causes $\mathbf{A}^{\downarrow}_{-k}$

New confounders $\mathbf{X}'_k := (\mathbf{X}, \mathbf{A}^{\uparrow}_{-k})$

New outcomes $\mathbf{Y}'_k := (Y, \mathbf{A}^{\downarrow}_{-k})$

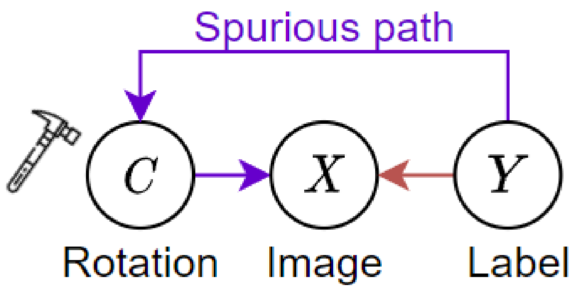
Estimate the counterfactual outcome of flipping A_k

This is a single-cause problem and can be solved by *any* existing ITE estimation algorithm

Comparison with traditional data augmentation

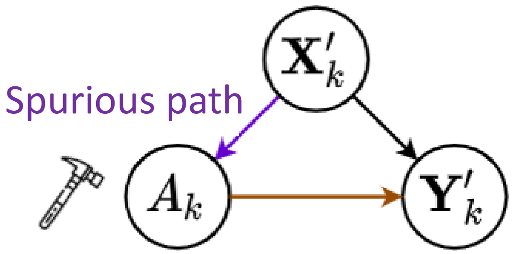
Traditional data augmentation

- Removes spurious correlation with non-causal variables
- The intervention effect is known, e.g. the rotating an image



SCP's data augmentation

- Remove confounding bias in treatment assignment
- The intervention effect is estimated by single-cause models – source of error



Single-cause perturbation: (2) estimation on augmented data

Covariate adjustment (S-learner):

Use *any* supervised learning algorithm to learn the conditional expectation

ID	A_1 A_2 A_3	0 0 0	0 0 1	0 1 0	0 1 1	1 0 0	1 0 1	1 1 0	1 1 1	Y
	Y(**a**)									
1		?	2	?	?	?	?	?	?	2
2		?	?	?	1	?	?	?	?	1
3		?	?	?	?	4	?	?	?	4

Green cells are “imputed” in the first step

Experiments

- Synthetic and semi-synthetic
- Turing various knobs to reflect different situations
  - Number of causes
  - Number of confounders
  - Sample size
  - Strengths of confounding
  - nonlinearity
  - Sparsity of interactions
  - High order interactions

- SCP's performance is consistently strong when the assumptions hold

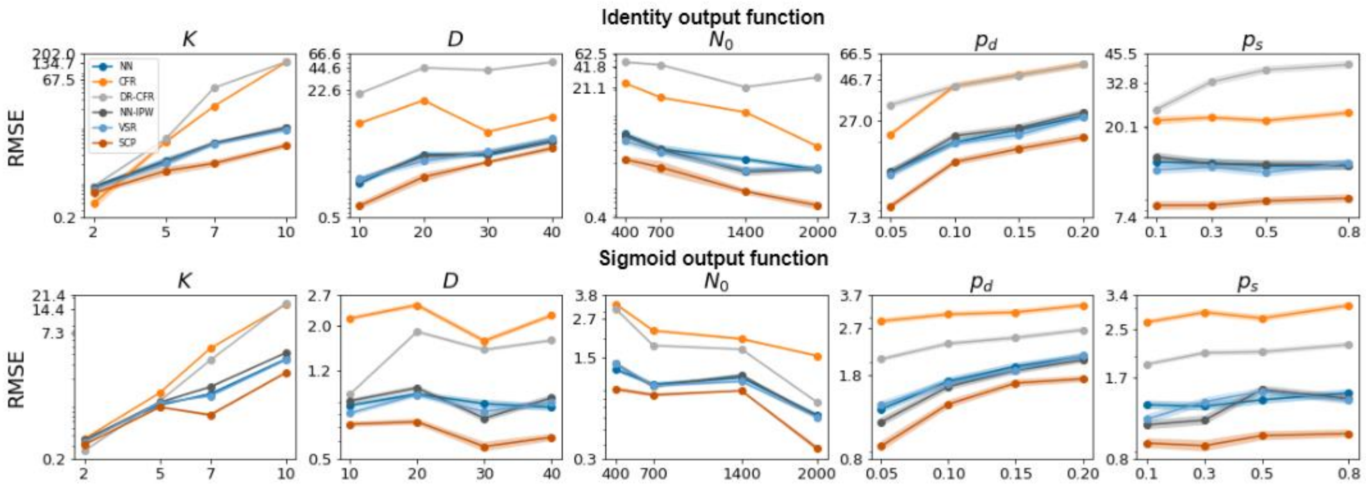


Figure 3. **Simulation Results** (best viewed in color). Y-axis is in *log scale*. RMSE is plotted with the 95% confidence interval shaded (the lower the better). Algorithms include NN, CFR, DR-CFR, NN-IPW, VSR and SCP. SCP consistently achieves the best performance.

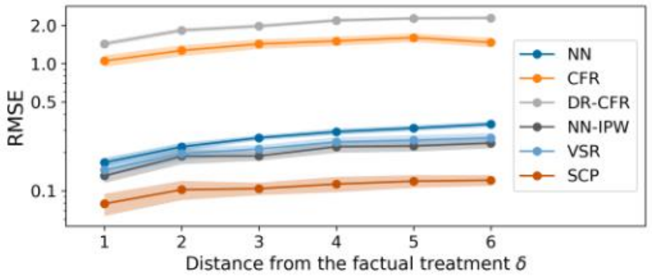


Figure 4. Counterfactual prediction accuracy as the target treatment a_i' moves farther away from the factual treatment a_i.

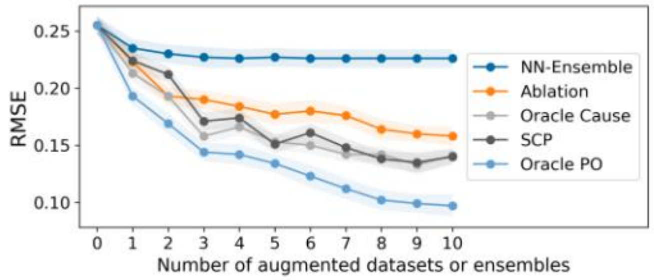


Figure 5. **Effect of data augmentation**. RMSE as more datasets are added to $\mathcal{D}^{Tr}$ or more models are added to the NN ensemble.

Reference:

Z. Qian, A. Curth, M. van der Schaar, Estimating Multi-cause Treatment Effects via Single-cause Perturbation, Neurips 2021

Code: https://github.com/ZhaozhiQIAN/Single-Cause-Perturbation-NeurIPS-2021

Lab website: https://www.vanderschaar-lab.com/

Personal profile: https://www.linkedin.com/in/qianzhaozhi/