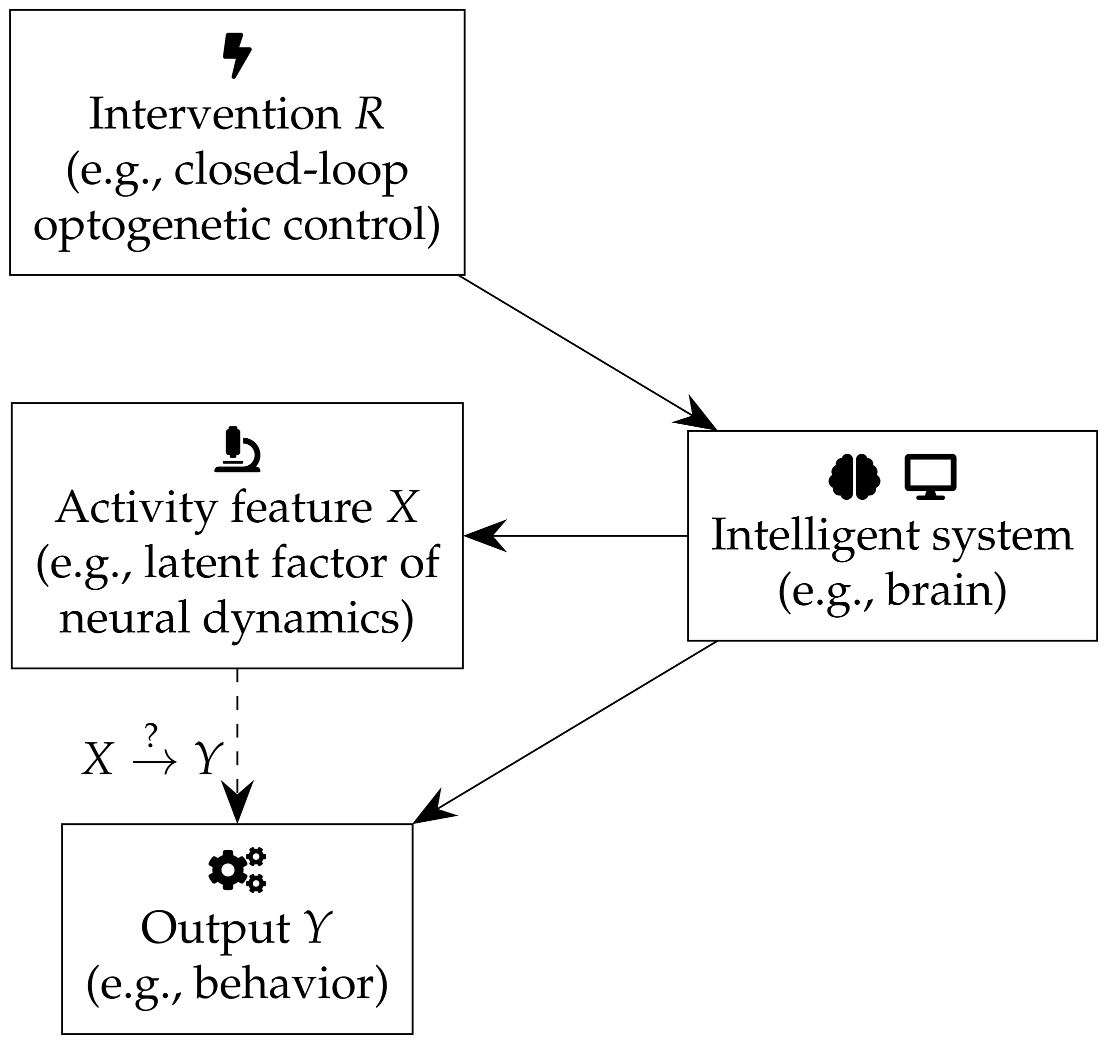
A foundation for causal characterization of latent neural dynamics with limited observational and interventional capacity

Kyle A. Johnsen[1,2] (kjohnsen@gatech.edu), Christopher J. Rozell[1]
1. Georgia Institute of Technology, Atlanta, GA, USA 2. Emory University, Atlanta, GA, USA

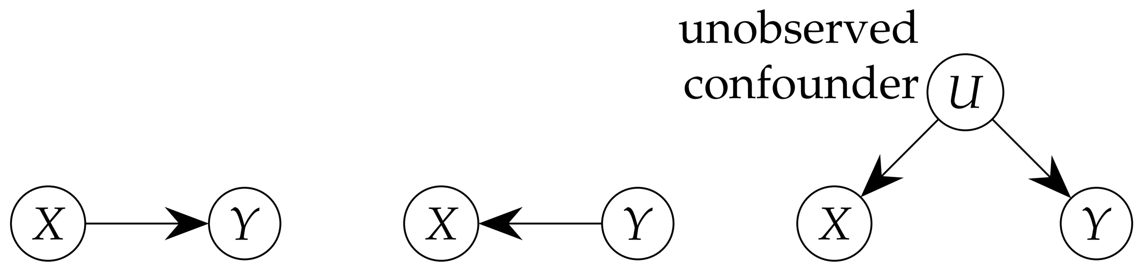
Closed-loop control can help causally probe intelligent systems.

Introduction

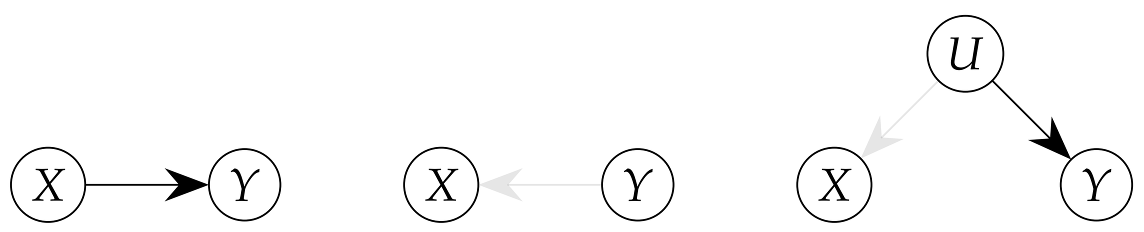
When studying intelligent systems in a gray-box fashion, we often want to understand how some measure of intermediate activity relates to function:



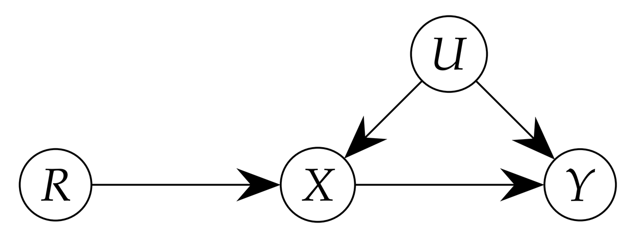
Correlation between activity X and output Y is not sufficient to infer causation:



Controlling X removes ambiguity when characterizing $X \to Y$:



When we can't perfectly control X, we can use an "instrumental variable" R to uncover the causal effect:



In our case, R is the reference value of an optimal feedback controller, representing what we'd like X to be. Framing the problem this way lets us leverage a rich set of instrumental variable (IV) estimation methods.

Methods

As a toy system, we implement the reference causal graph with linear-Gaussian relationships:

$$X = m_{UX}U + m_{RX}R + \epsilon_X$$
$$Y = m_{XY}X + m_{UY}U + \epsilon_Y$$

$$R \sim \mathcal{N}(0, \sigma_R) \qquad U \sim \mathcal{N}(0, \sigma_U)$$
$$\epsilon_X \sim \mathcal{N}(0, \sigma_X) \qquad \epsilon_Y \sim \mathcal{N}(0, \sigma_Y)$$

We can decompose the variance of Y into effects from X and U:

$$\mathrm{Var}(Y) = m_{XY}^2(m_{UX}^2\sigma_U^2 + m_{RX}^2\sigma_R^2 + \sigma_X^2) + m_{UY}^2\sigma_U^2 + \sigma_Y^2$$

And compute the correlation for a purely observational experiment (no control, $\sigma_R = 0$):

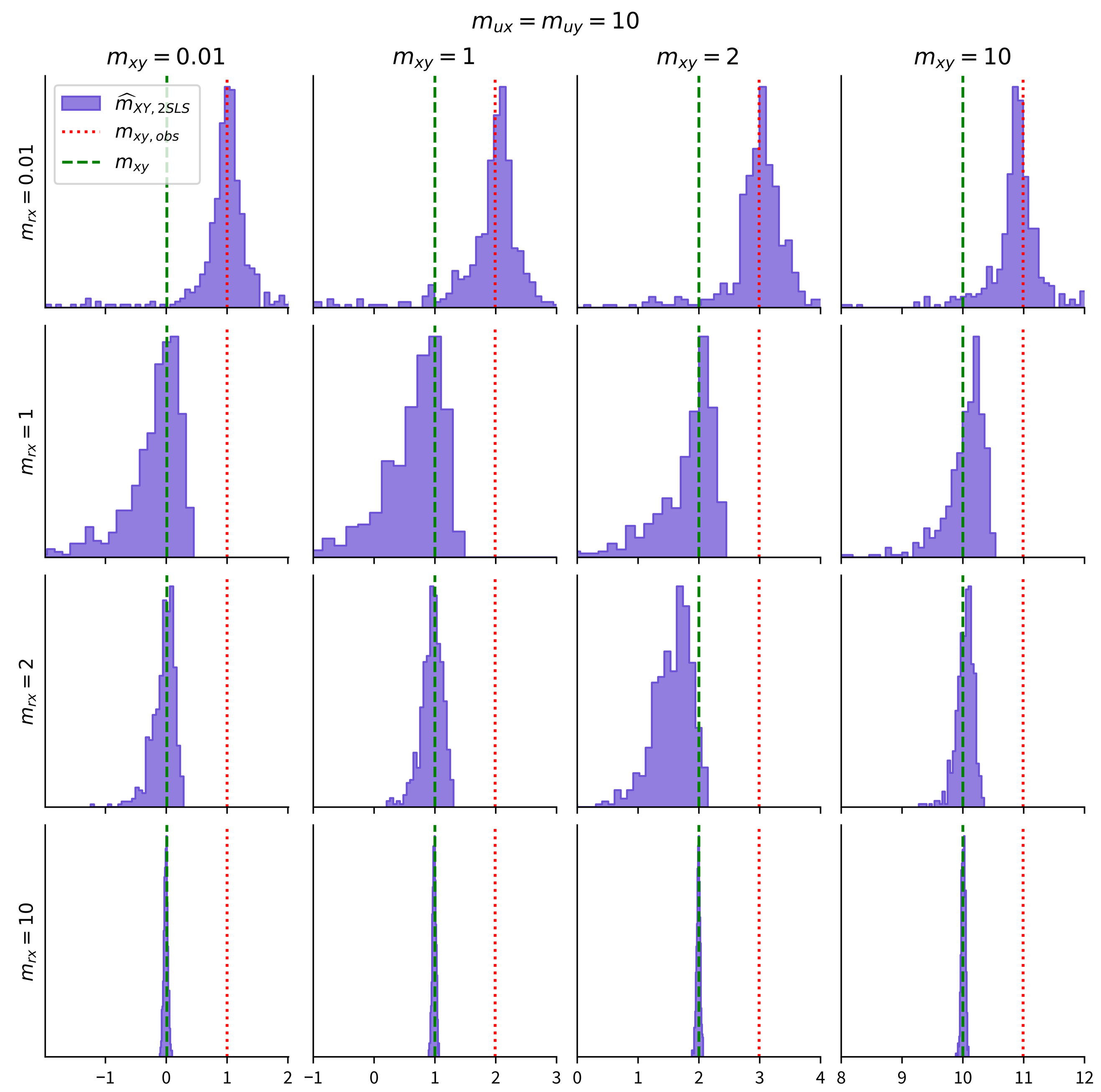
$$\mathrm{Cov}_{\mathrm{obs}}(X,Y) = m_{XY}^2(m_{UX}^2\sigma_U^2 + \sigma_X^2) + m_{UY}^2\sigma_U^2 + \sigma_Y^2$$
$$r_{\mathrm{obs}}(X,Y) = \frac{\mathrm{Cov}_{\mathrm{obs}}(X,Y)}{\sqrt{\mathrm{Var}_{\mathrm{obs}}(X)\mathrm{Var}_{\mathrm{obs}}(Y)}}$$

Two-stage least squares (2SLS) is a standard IV method for estimating m_{XY}:

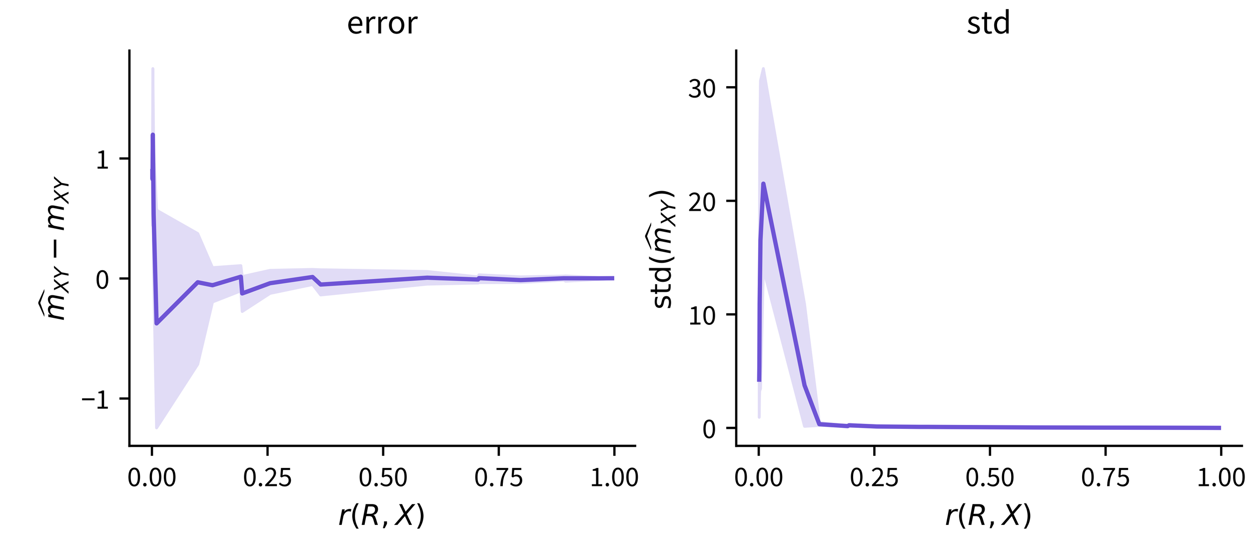
$$\hat{X} = \hat{m}_{RX}R \qquad \hat{Y} = \hat{m}_{XY}\hat{X}$$

Demonstration

Bias and variance of causal effect estimates decrease as the instrument strength (our control performance) increases:



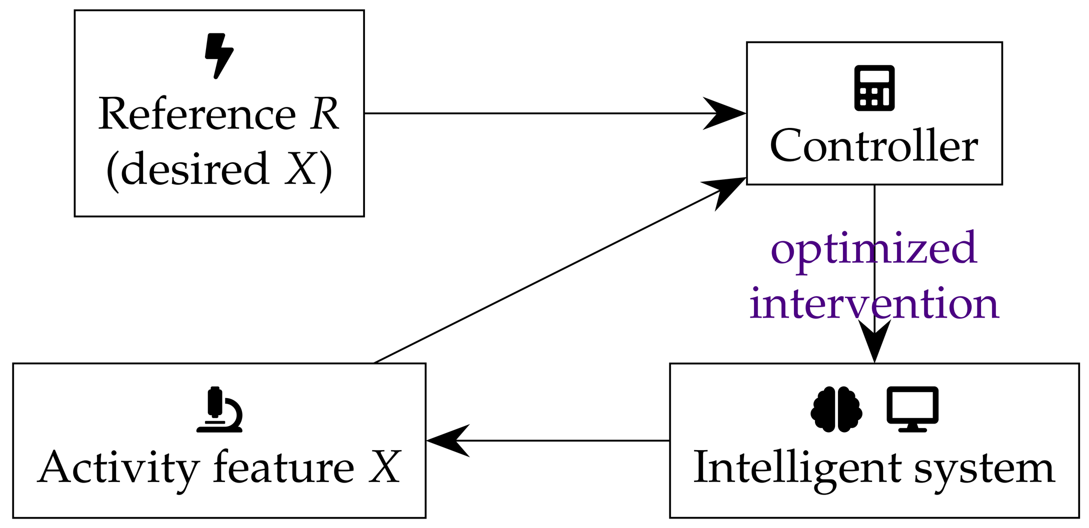
Results across $m_{RX}, m_{XY}, m_{UX}, m_{UY}$ conditions:

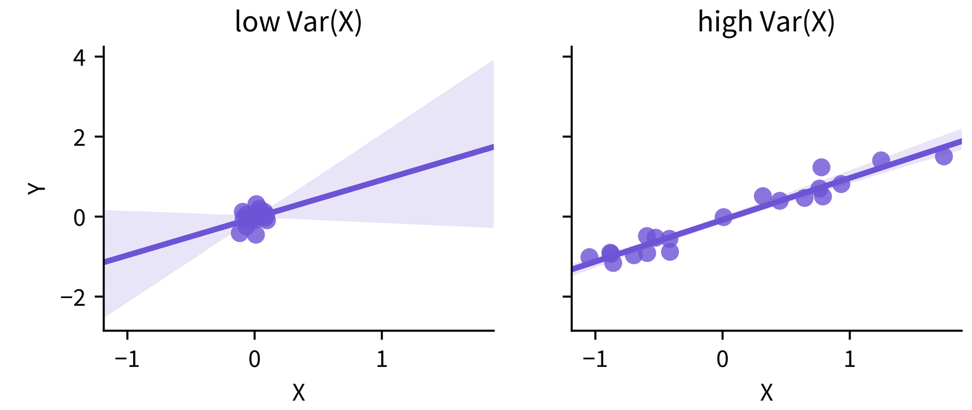


Future directions

When is closed-loop control helpful or necessary compared to random intervention?

Closed-loop optimization could probe X more efficiently, facilitating causal inference especially for high-dimensional input (e.g., multichannel optogenetics) unlikely to affect X by chance:





Modern machine-learning methods may be needed for nonlinear $R \to X$ and $X \to Y$ relationships:

For example, DeepIV [1], DeepGMM [2], KernelIV [3], MMR-IV [4]. Figure from Bakhitov and Singh [5]:

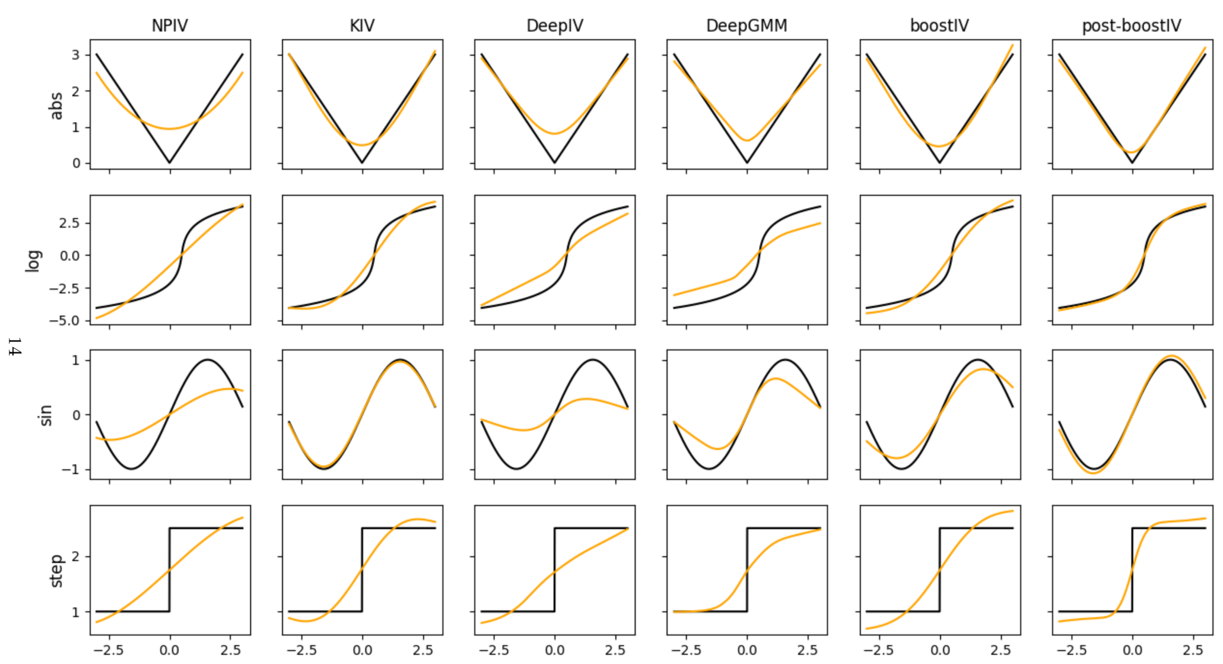
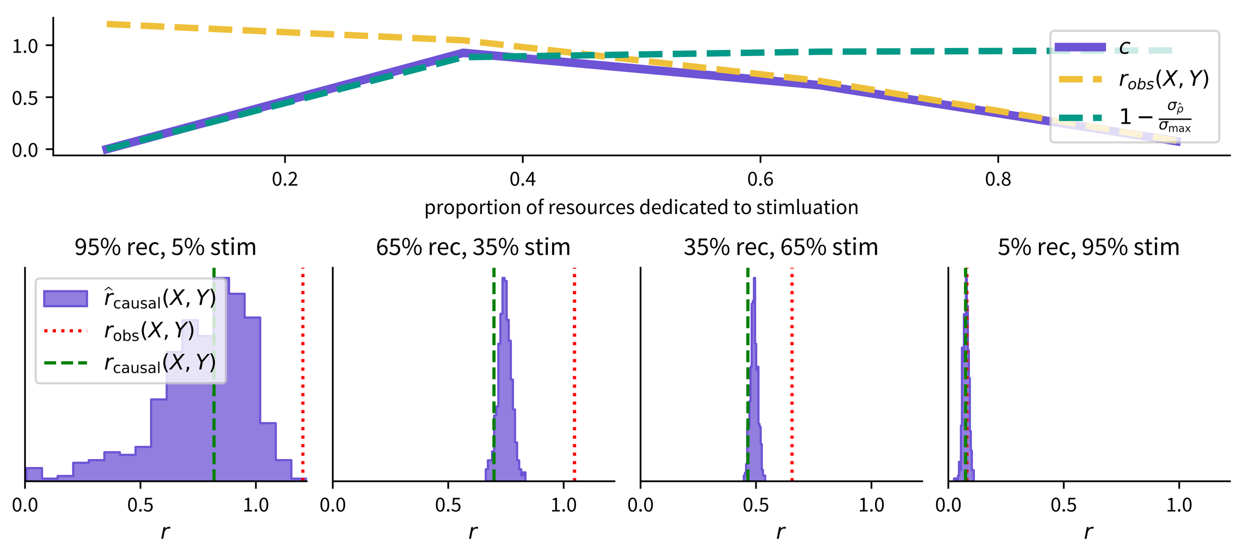


Figure 1. Out-of-sample average fit across simulations. The black line is the true function, the orange line is the fit.

When designing causal inference experiments with limited instrumentation capacity, we must trade off measurement and manipulation:

Proposed causal discovery metric $c = \underbrace{\rho_{\mathrm{obs}}(X,Y)}_{\text{feature quality}} \underbrace{\left(1 - \frac{\sigma_{\hat{\rho}}}{\sigma_{\hat{\rho},\max}}\right)}_{\text{control quality}}$



Acknowledgments

Supported by NIH CRCNS grant R01NS115327.
Kyle Johnsen was supported by the Georgia Tech/Emory NIH/NIBIB Training Program in Computational Neural Engineering (T32EB025816).

References

[1] Jason Hartford, Greg Lewis, Kevin Leyton-Brown, and Matt Taddy. Deep IV: A Flexible Approach for Counterfactual Prediction. In *Proceedings of the 34th International Conference on Machine Learning*, pages 1414–1423. PMLR, July 2017. URL https://proceedings.mlr.press/v70/hartford17a.html.

[2] Andrew Bennett, Nathan Kallus, and Tobias Schnabel. Deep Generalized Method of Moments for Instrumental Variable Analysis. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://papers.nips.cc/paper_files/paper/2019/hash/15d185eaa7c954e77f5343d941e25fbd-Abstract.html.

[3] Rahul Singh, Maneesh Sahani, and Arthur Gretton. Kernel Instrumental Variable Regression. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper/2019/hash/17b3c7061788dbe82de5abe9f6fe22b3-Abstract.html.

[4] Rui Zhang, Masaaki Imaizumi, Bernhard Schölkopf, and Krikamol Muandet. Instrumental Variable Regression via Kernel Maximum Moment Loss, February 2023. URL http://arxiv.org/abs/2010.07684.

[5] Edvard Bakhitov and Amandeep Singh. Causal Gradient Boosting: Boosted Instrumental Variable Regression, January 2021. URL http://arxiv.org/abs/2101.06078.