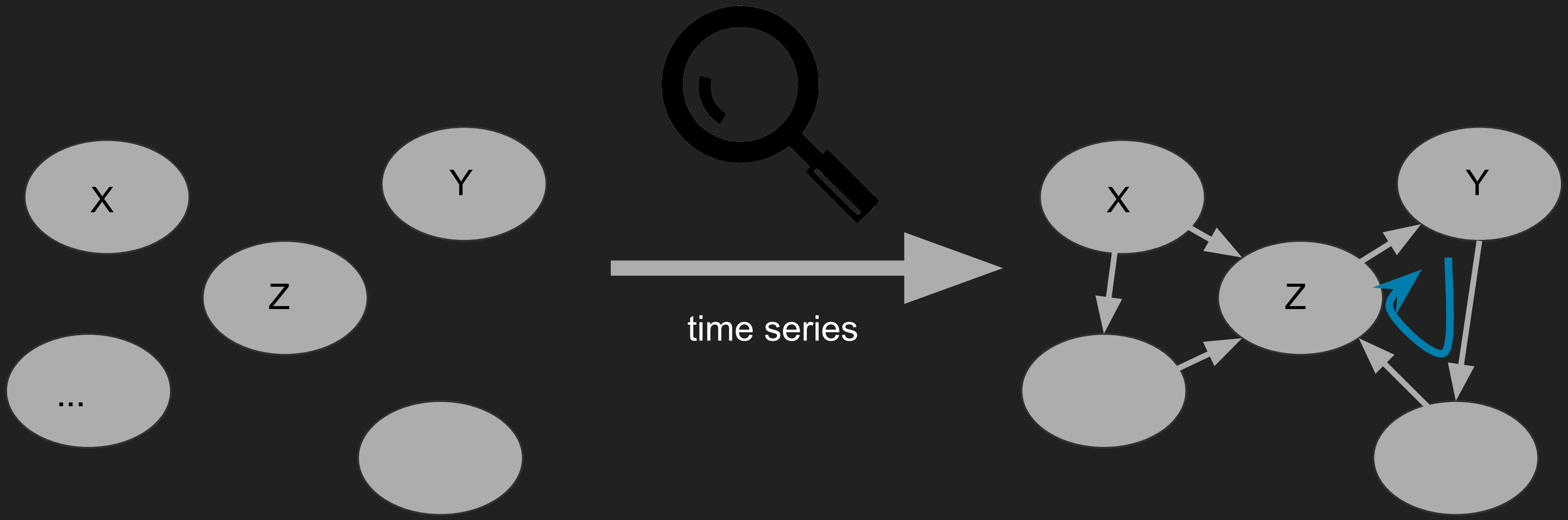
Active Learning of Continuous-time Bayesian Networks through Interventions

Dominik Linzner and Heinz Koeppl

The Why Company GmbH | Bio-inspired Communication Systems Lab
@ Technische Universität Darmstadt

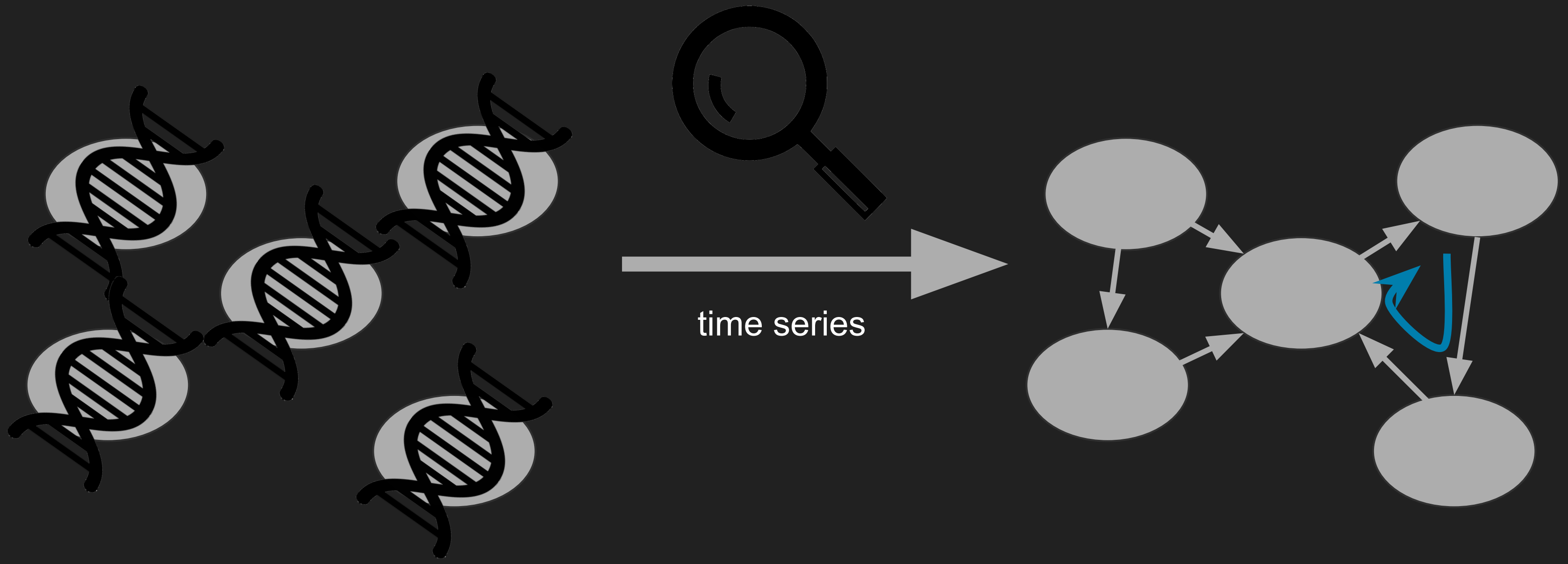
Goal: Learning directed (cyclic) networks



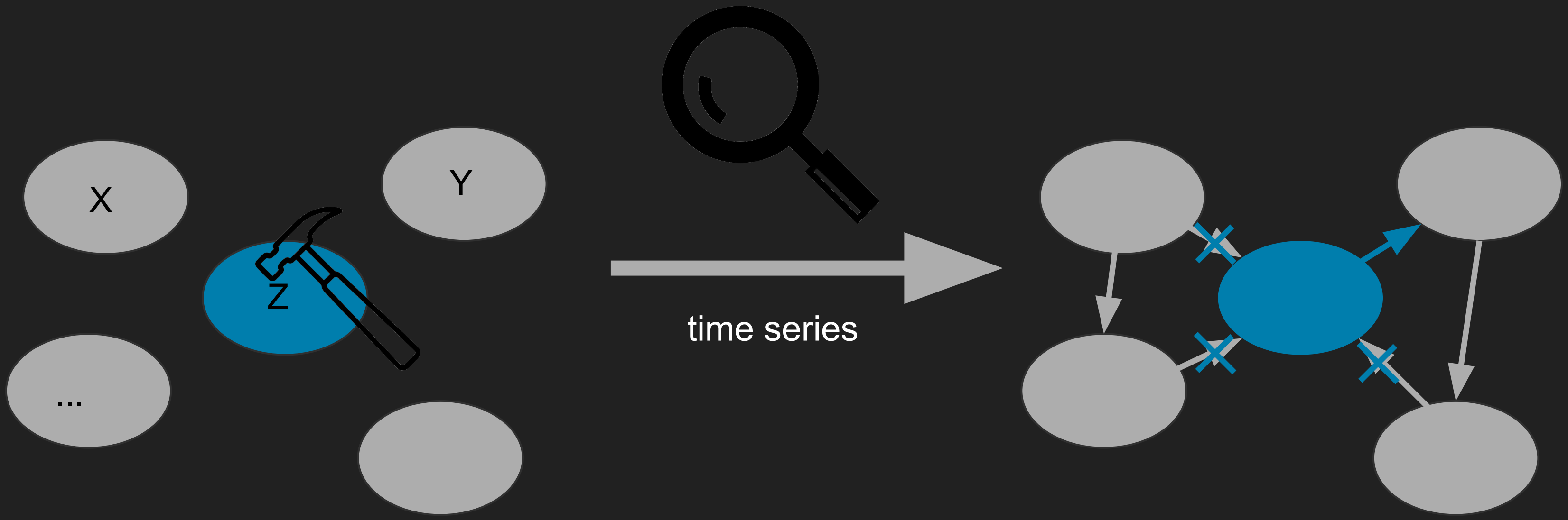
Causal models with cycles:

Nodelman, U., Shelton, C. R., & Koller, D. (2002). Continuous Time Bayesian Networks. UAI

Problem: Generating data is expensive



Solution: Speed up using active learning through interventions



e.g. gene knock-out experiments in biology

Our contributions

1. First active learning scheme for continuous-time Bayesian networks (CTBNs)
2. Studying the effect of interventions on CTBNs
3. Introduce a variational criterion for active learning suitable for high-dimensional problems

Our contributions

1. First active learning scheme for continuous-time Bayesian networks (CTBNs)
2. Studying the effect of interventions on CTBNs
3. Introduce a variational criterion for active learning suitable for high-dimensional problems

Active learning in high-dimensions

Bayesian Optimal Experimental Design (BOED)

Integrate over all possible experimental outcomes for all possible models (Expected Information Gain)

$$\mathrm{EIG} = \mathsf{E}_{p(D,\Theta)}\left[\ln \frac{p(\Theta \mid D)}{p(\Theta)}\right]$$

Box, G. E. P., & Hill, W. J. (1967). Discrimination Among Mechanistic Models. Technometrics, 9(1), 57–71.

Active learning in high-dimensions

Bayesian Optimal Experimental Design (BOED)

Integrate over all possible experimental outcomes for all possible models (Expected Information Gain)

$$\mathrm{EIG} = \mathsf{E}_{p(D,\Theta)}\left[\ln \frac{p(\Theta \mid D)}{p(\Theta)}\right] \leq \mathrm{VBHC}(\kappa)$$

Variational Box-Hill criterion (VBHC)

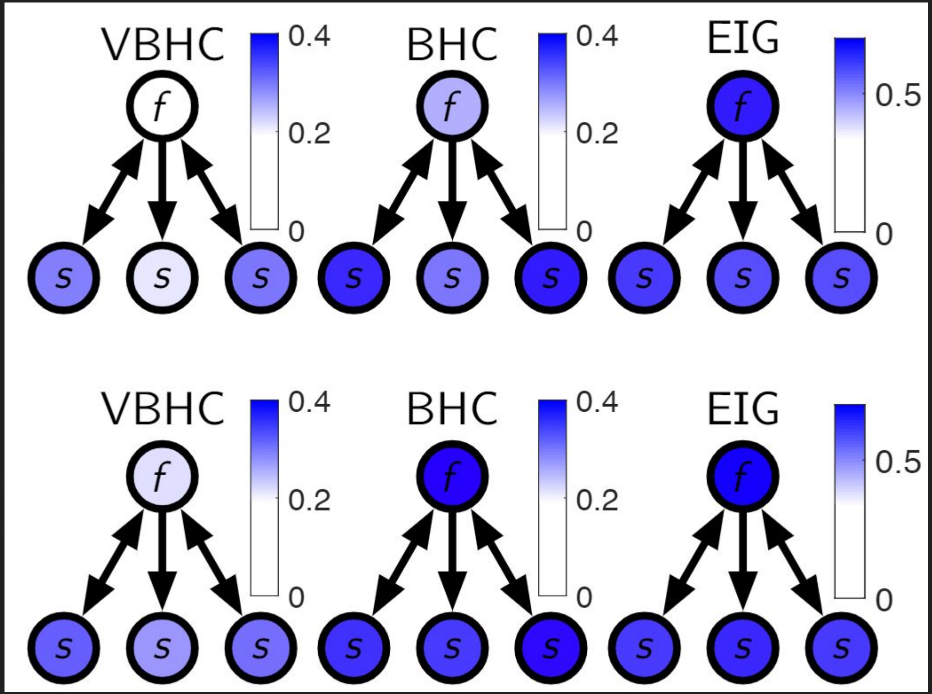
Variational extension of Box-Hill criterion for model discrimination

$$\mathrm{VBHC}(\kappa) = \mathsf{E}_{p(\Theta)q_\kappa(\tilde{\Theta})}\left[\mathrm{KL}(p(D \mid \Theta) \,||\, p(D \mid \tilde{\Theta})\right] + \mathrm{KL}(p(\Theta) \,||\, q_\kappa(\Theta))$$

Box, G. E. P., & Hill, W. J. (1967). Discrimination Among Mechanistic Models. Technometrics, 9(1), 57–71.

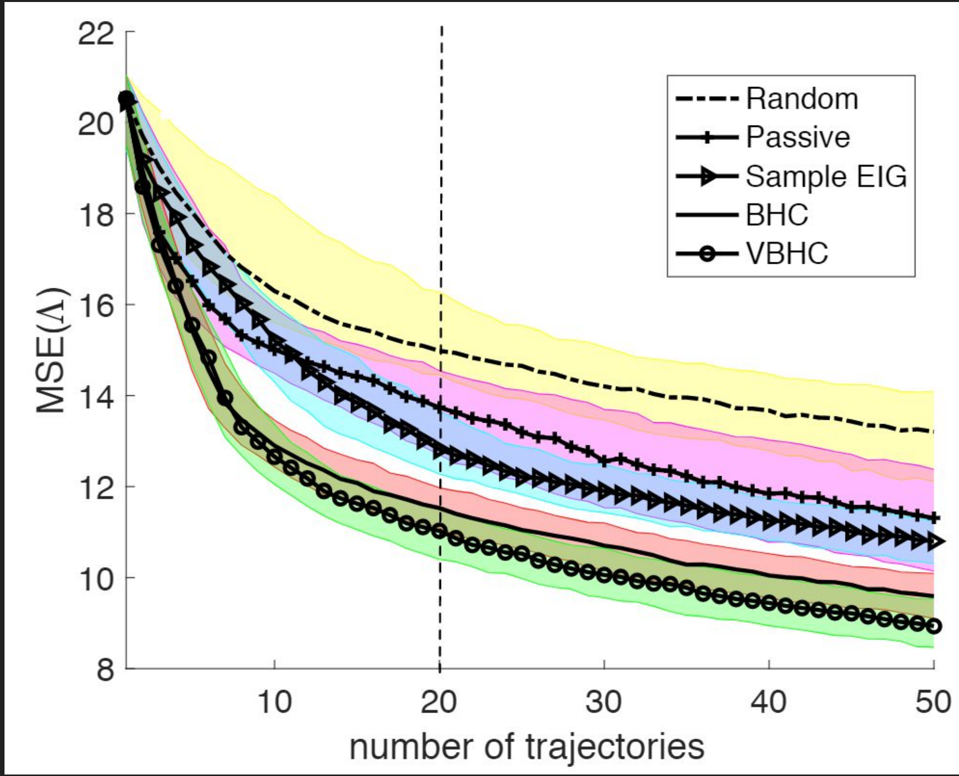
Variational Box-Hill criterion (VBHC)

- Synthetic scenario (4 variables)
- time-scale separation (fast and slow)
- Color code is probability of intervention
- VBHC (and BHC) exploit time-scale separation (sample estimate of EIG does not)



Variational Box-Hill criterion (VBHC)

- Synthetic scenario (4 variables)
- time-scale separation (fast and slow)
- Mean-Squared-Error for parameter estimate
- VBHC can improve on BHC



For more details (also on CTBNs and interventions) =>Check out our poster & paper

Active Learning of Continuous-time Bayesian Networks through Interventions

Dominik Linzner [1 2] **Heinz Koeppl** [1 3]

Abstract

We consider the problem of learning structures and parameters of Continuous-time Bayesian Networks (CTBNs) from time-course data under minimal experimental resources. In practice, the cost of generating experimental data poses a bottleneck, especially in the natural and social sciences. A popular approach to overcome this is Bayesian optimal experimental design (BOED). However, BOED becomes infeasible in high-dimensional settings, as it involves integration over all possible experimental outcomes. We propose a novel criterion for experimental design based on a variational approximation of the expected information gain. We show that for CTBNs, a semi-analytical expression for this criterion can be calculated for structure and parameter learning. By doing so, we

to learn these dependencies. This is a problem when data is acquired under limited resources, which is the case in dedicated experiments, e.g. in molecular biology or psychology (Steinke et al., 2007; Zechner et al., 2012; Liepe et al., 2013; Myung & Pitt, 2015; Dehghannasiri et al., 2015; Prangemeier et al., 2018). Active learning schemes pave a principled way to design sequential experiments such that the required resources are minimized.

The framework of Bayesian optimal experimental design (BOED) (Chaloner, 1987; Ryan et al., 2016) allows for the design of active learning schemes, which are provably (Lindley, 1956; Sebastiani & Wynn, 2000) one-step optimal. However, BOED becomes infeasible in high-dimensional settings, as it involves integration over possible experimental outcomes.

While active learning schemes have been previously applied in order to learn dependency structures (Tong & Koller