Learning Structural Causal Models through Deep Generative Models

A Survey on Methods, Guarantees, and Challenges

Audrey Poinsot[1,2], Alessandro Leite[1], Nicolas Chesneau[2], Michèle Sébag[1], Marc Schoenauer[1]

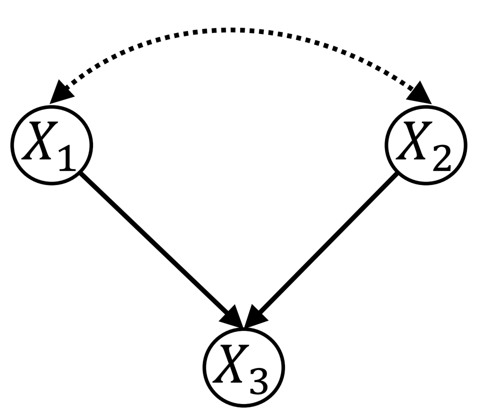
[1]INRIA Saclay, [2]Ekimetrics

Given a known causal structure and observational data, what are the capabilities of existing Deep Structural Causal Models in answering counterfactual questions?

Motivations

The true SCM M^*:

- $X_1 = f_1(U_1)$
- $X_2 = f_2(U_2)$
- $X_3 = f_3(X_1, X_2, U_3)$
- $P(U)$ s.t. $U_3 \perp\!\!\!\perp U_1$ and $U_3 \perp\!\!\!\perp U_2$



Counterfactual query of interest Q:

$Q(M^*) = P(X_3 | do(X_1), U)$

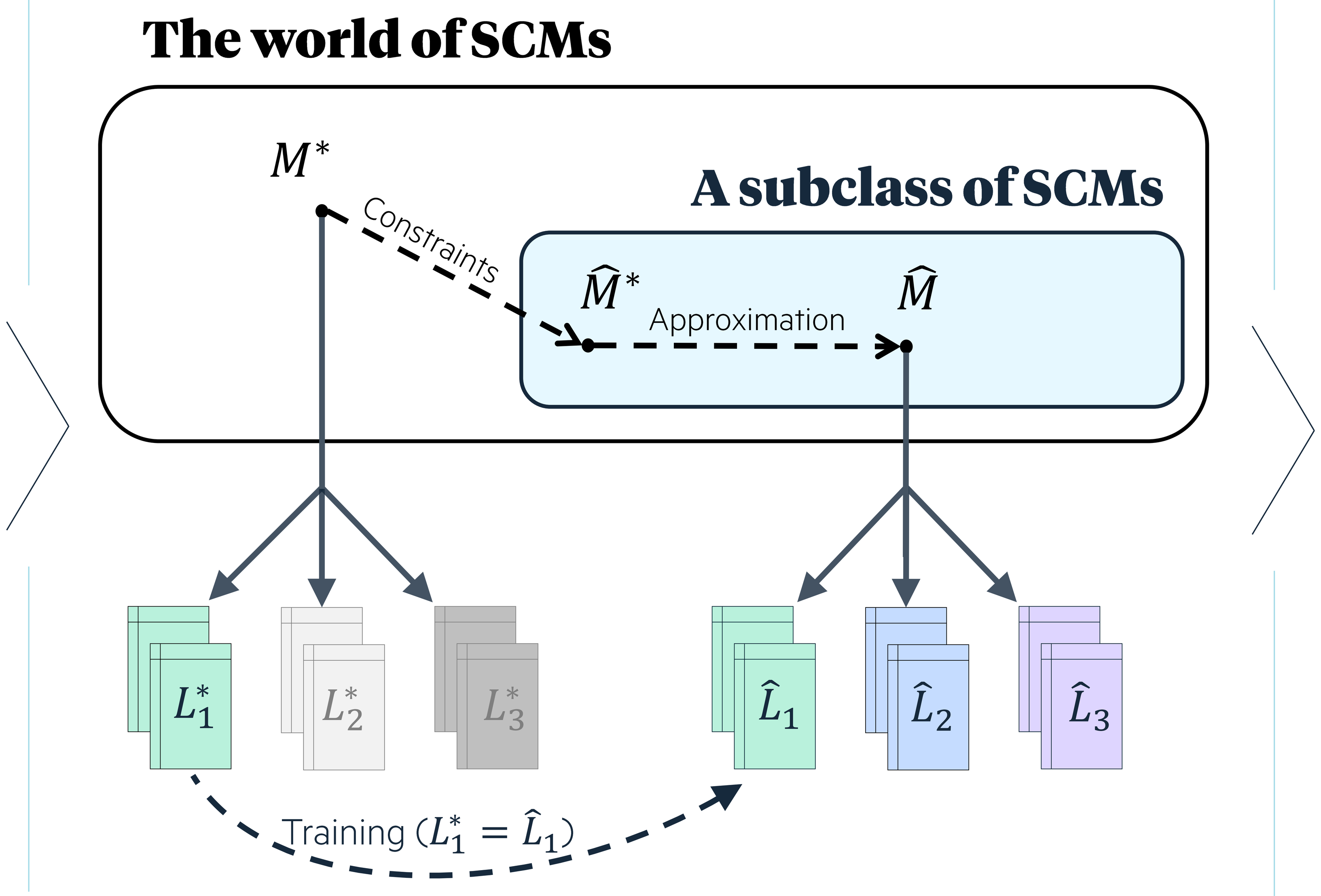
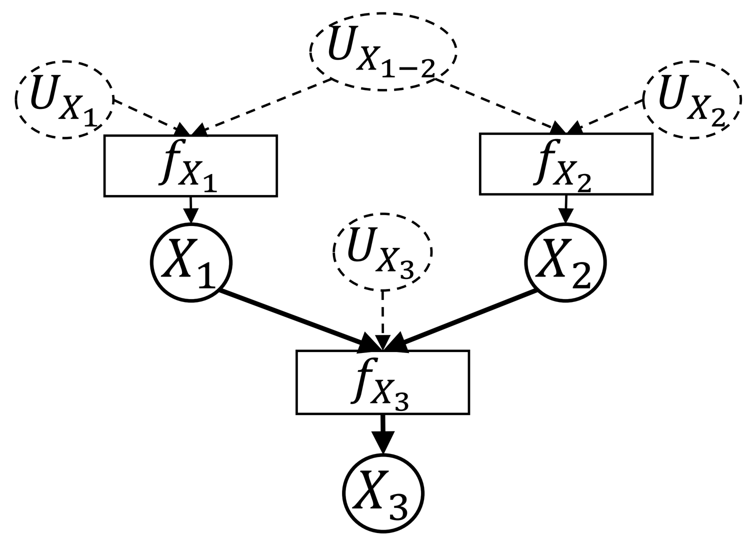


Figure inspired by [Xia *et al*, 2021]

The learned DSCM $\widehat{M}$:

- $X_1 = \hat{f}_1(U_1)$
- $X_2 = \hat{f}_2(U_2)$
- $X_3 = \hat{f}_3(X_1, X_2, U_3)$
- $P(\widehat{U})$ s.t. $\forall i, j,\ \widehat{U_i} \perp\!\!\!\perp \widehat{U_j}$

with f_i a Deep Generative Model $\forall i$



Estimated counterfactual query $\widehat{Q}$:

$\widehat{Q} = \hat{P}(X_3 | do(X_1), \widehat{U})$

Classification

Structural Causal Models

Numerous assumptions

Identification concerns

Deep Generative Models

Flexibility, few assumptions

Few guarantees

Learning Structural Causal Models through Deep Generative Models

DSCM

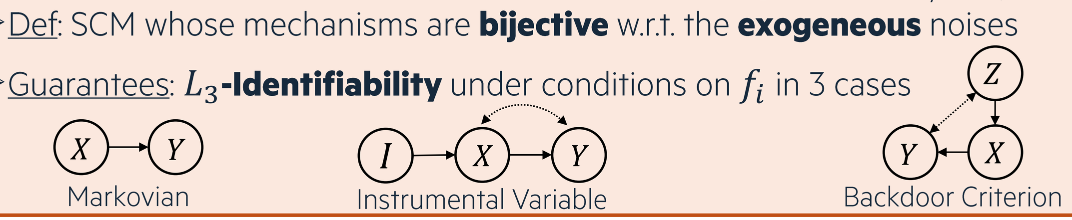
- Def: SCM whose mechanisms are **deep (conditional) generative models**
- No theoretical guarantees [Pawlowski *et al*, 2020]

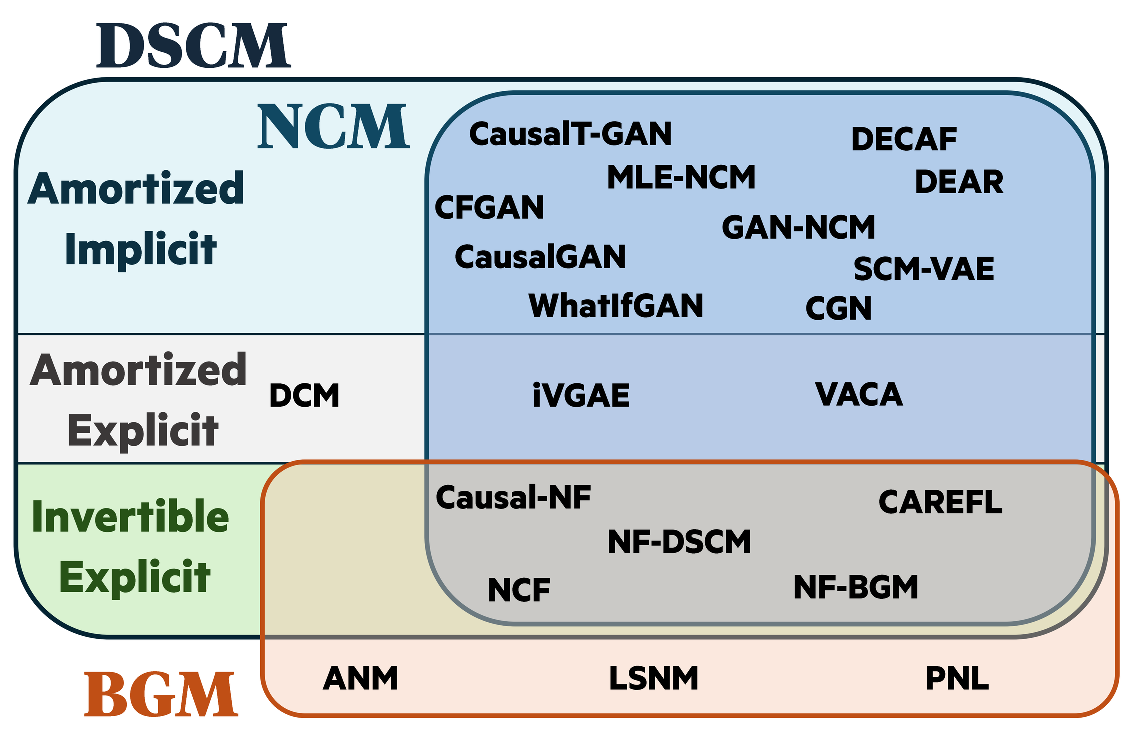
NCM [Xia *et al*, 2021 & 2023]

- Def: SCM whose mechanisms are **feedforward neural networks**
- Guarantees:
  - **Expressivity**: Given G there always exists an NCM L_3-consistent with the true SCM
  - **L_3-Identifiability** iif L_3-Identifiability holds for the true SCM

BGM [Nasr-Esfahany *et al*, 2023]

- Def: SCM whose mechanisms are **bijective** w.r.t. the **exogeneous** noises
- Guarantees: **L_3-Identifiability** under conditions on f_i in 3 cases





Amortized Implicit

- f_i is a conditional implicit-likelihood model learned with a loss implicitly considering U_i
- Adversarial learning, Loss to fit the distribution: **Causal-TGAN**, **CausalGAN**, **CFGAN**, **DECAF**, **WhatIfGAN**, **CGN**, **DEAR**, **GAN-NCM**, **MLE-NCM**, **SCM-VAE**

Amortized Explicit

- f_i is learned with an auto-encoder s.t. $g_i(PA(X_i), U_i) = f_i$ and $e_i(X_i, PA(X_i)) = U_i$
- (Variational) (Graph) Auto Encoders, Diffusion Models: **iVGAE**, **VACA**, **DCM**

Invertible Explicit

- f_i is supposed diffeomorphic w.r.t U_i s.t. $U_i = f_i^{-1}(X_i, PA(X_i))$
- Normalizing Flow: **Causal-NF**, **NF-DSCM**, **NCF**, **CAREFL**, **NF-BGM**

Takeaways

Theoretical comparison

Causal structure: Knowing the causal ordering is sufficient

Hidden confounding

- NF-BGM and WhatIfGAN consider dependent noises
- NCM and NCF deal with **semi-Markovian DAGs**

Abduction step: Only 7 methods implement the abduction step while the sample-rejection procedure is applicable to all the methods

Identifiability guarantees

- NCM **L_3-identifiability** result is applicable to all the methods except DCM
- DCM provides error bounds and L_3-identifiability under sufficiency and additional hypothèses
- **NeuralID** algorithm enables to automatically check for point identification of a query given a DAG and a dataset

Empirical comparison

Experimental evaluation

- High heterogeneity: datasets, causal task, metrics, …
- Lack of a unified benchmark
- Simulations lack sources of randomness (DAG, noise distribution, …)

Applications

- **Fairness**: counterfactual fairness, fair prediction in-processing and pre-processing
- **Explanability**: counterfactual explanations, scientific discovery
- **Machine Learning robustness**: Out-of-domain data augmentation, realistic dataset generation

Challenges & Opportunities

Lack of evaluation

- Lack of a proper benchmark: simulated data with different sources of randomness and different assumptions
- Lack of a complete evaluation strategy: data efficiency, computational time, robustness to unsatisfied assumptions, …

From identification to partial identification

- Strong and/or un-testable hypotheses are taken (e.g., known causal structure, no selection bias)
- Whenever partial identification is impossible or too hard to get, **sensitivity analysis** is a solution

Sensitive applications

- The assumptions, in particular the causal graph, much be validated by experts beforehand
- **NeuralID** enable to test for point identification
- Sensitivity analysis is crucial

audrey.poinsot@ekimetrics.com