LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
LMU MUNICH SCHOOL OF MANAGEMENT
INSTITUTE OF ARTIFICIAL INTELLIGENCE (AI) IN MANAGEMENT
Munich Center for Machine Learning

Partial Counterfactual Identification of Continuous Outcomes with a Curvature Sensitivity Model

Valentyn Melnychuk, Dennis Frauen, Stefan Feuerriegel

LMU Munich & Munich Center for Machine Learning (MCML), Munich, Germany

NeurIPS 2023, Spotlight Presentation

Introduction: Counterfactual identification in Markovian SCMs

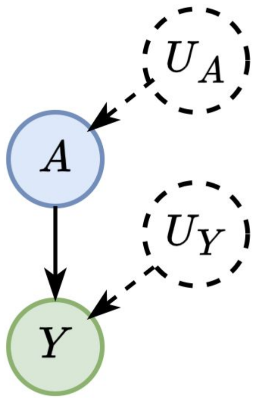
Why this is important?

- Counterfactual inference is widely used in data-driven decision-making: it aims to answer **retrospective** “what if” questions
- Counterfactual identifiability is only possible with unnatural or unrealistic assumptions (e.g. monotonicity of the functions in the Markovian structural causal models (SCMs))

Problem formulation

Given observational dataset from $\mathbb{P}^{\mathcal{M}}(Y, A)$, induced by some bivariate SCM $\mathcal{M}$ with

A treatments
Y (factual) outcomes



$$\mathcal{M} = \langle \mathbf{U}, \mathbf{V}, \mathbb{P}(\mathbf{U}), \mathcal{F} \rangle$$
$$\mathbf{U} = \{U_A \in \{0,1\}, U_Y \in [0,1]^d\}$$
$$\mathbf{V} = \{A \in \{0,1\}, Y \in \mathbb{R}\}$$
$$\mathbb{P}(\mathbf{U})\colon U_A \sim \text{Bern}(p_A),\ 0 < p_A < 1,$$
$$U_Y \sim \text{Unif}(0,1)^d$$
$$\mathcal{F} = \{f_A(U_A) = U_A, f_Y(A, U_Y)\}$$

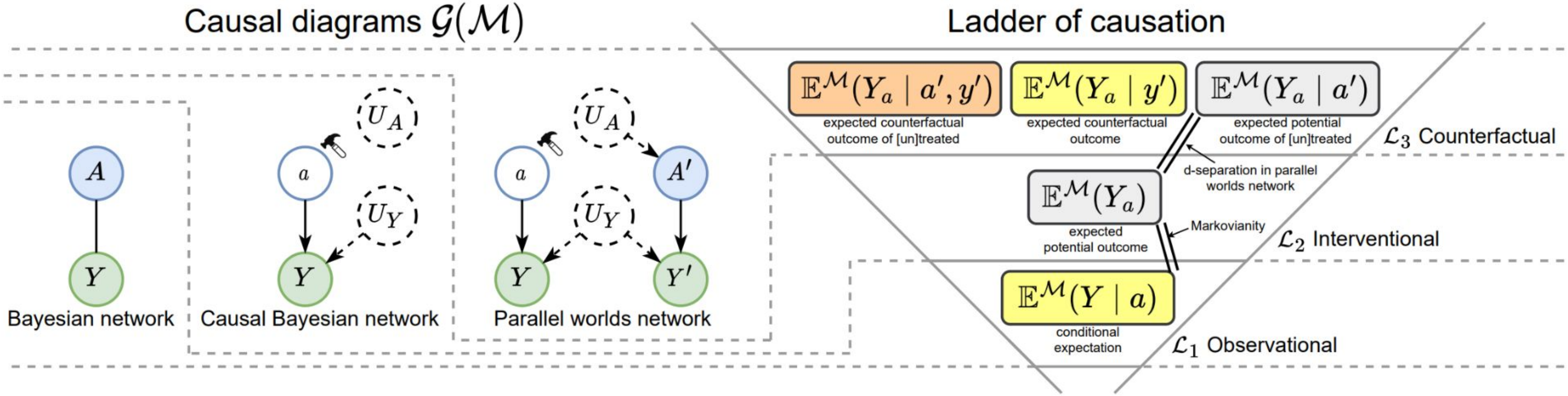
we want to perform a partial identification of an **expected counterfactual outcome of [un]treated** ECOU [ECOT]

$$Q^{\mathcal{M}}_{a' \to a}(y') = \mathbb{E}^{\mathcal{M}}(Y_a \mid a', y')$$

Introduction: Task complexity – Related work

- Counterfactual queries in general are not identifiable from both L1 and L2 data even for Markovian SCMs
- Partial identification of L3 discrete outcomes / L2 continuous outcomes does not generalize -> we need brand new mathematical tools for L3 partial identification with continuous outcomes

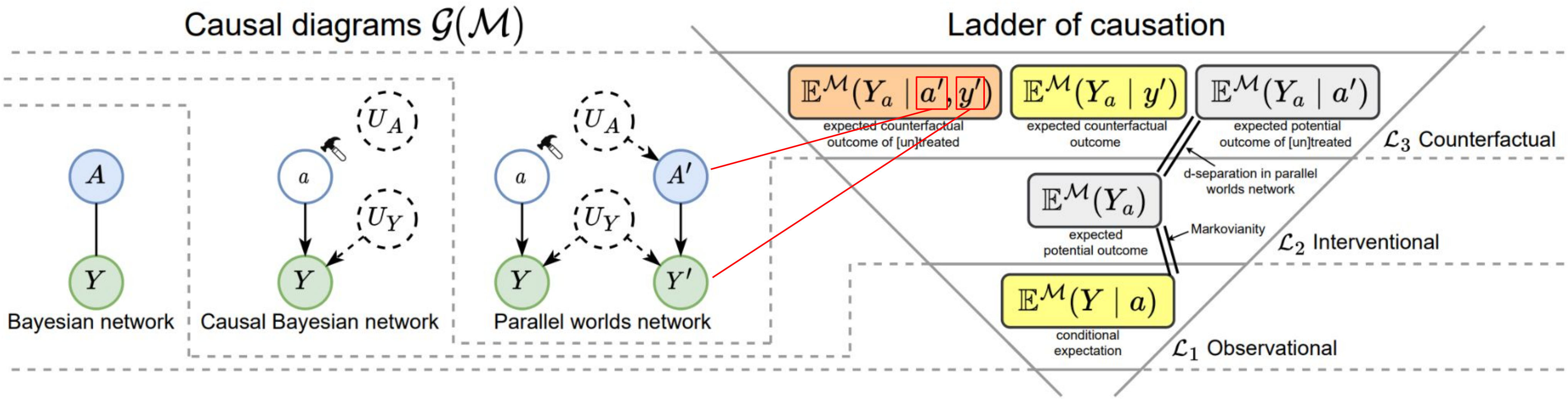
Why this is hard?



Introduction: Task complexity – Related work

- Counterfactual queries in general are not identifiable from both L1 and L2 data even for Markovian SCMs
- Partial identification of L3 discrete outcomes / L2 continuous outcomes does not generalize -> we need brand new mathematical tools for L3 partial identification with continuous outcomes

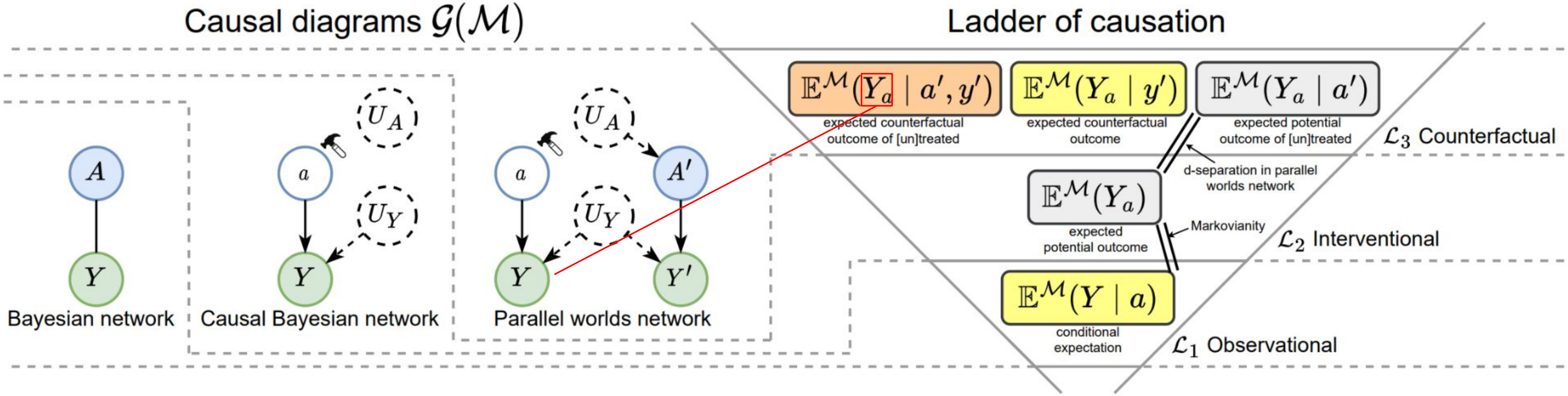
Why this is hard?



Introduction: Task complexity – Related work

- Counterfactual queries in general are not identifiable from both L1 and L2 data even for Markovian SCMs
- Partial identification of L3 discrete outcomes / L2 continuous outcomes does not generalize -> we need brand new mathematical tools for L3 partial identification with continuous outcomes

Why this is hard?



Introduction: Task complexity – Related work

Related work

Layer	M/SM	Symbolic identifiability	Point identification methods	Partial identification methods: Discrete outcomes	Partial identification methods: Continuous outcomes
$\mathcal{L}_2$ Interventional	M	Always via back-door criterion [6]	Deep generative models [70, 138]	—	—
	SM	Do-calculus & rules of probability [52, 75, 114]	Potential outcomes framework [12, 26, 48, 85, 105, 124] ; binary IV [36, 54, 125] ; proxy variables [81, 87]	Partially observed back-/front-door variables [77]; canonical SCM [130]	No-assumptions bound [83]; MSM [13, 33, 37, 56, 57, 58, 84, 92, 120]; outcome sensitivity models [13, 100]; confounding functions [9, 15, 103]; noisy proxy variables [45]; IV [44, 51, 64, 139]; ATD [3]; clustered DAGs [93]
$\mathcal{L}_3$ Counterfactual	M	Parallel worlds networks [2, 115], counterfactual unnesting theorem [25]	Deep generative models [19, 27, 66, 94, 107, 108, 111, 112] ; Markovian BGMs [55, 62, 88, 89, 117, 142] ; transport-based counterfactuals [30]	PN, PS, PNS [4, 76, 78, 97, 121]; response functions framework / canonical partitioning [4, 91, 106, 131, 135, 136, 137, 141]; causal marginal problem [42, 109]; deep twin networks [123]	**CSM (this paper)**
	SM		ETT [113] ; path-specific effects [116, 140] ; deep generative models [29, 80, 128, 129, 134] ; semi-Markovian BGMs [88]		Future work (see discussion in Appendix E); ANMs with hidden confounding [65]

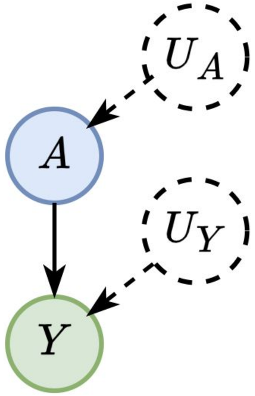
Legend:

- M/SM: Markovian SCM (M), semi-Markovian SCM (SM)

Introduction: Assumptions - Motivating example

Assumptions

- **Bivariate Markovian SCMs** with continuously-differentiable functions and high-dimensional latent noise: $\mathfrak{B}(C^k, d)$, k >= 0, d > 0

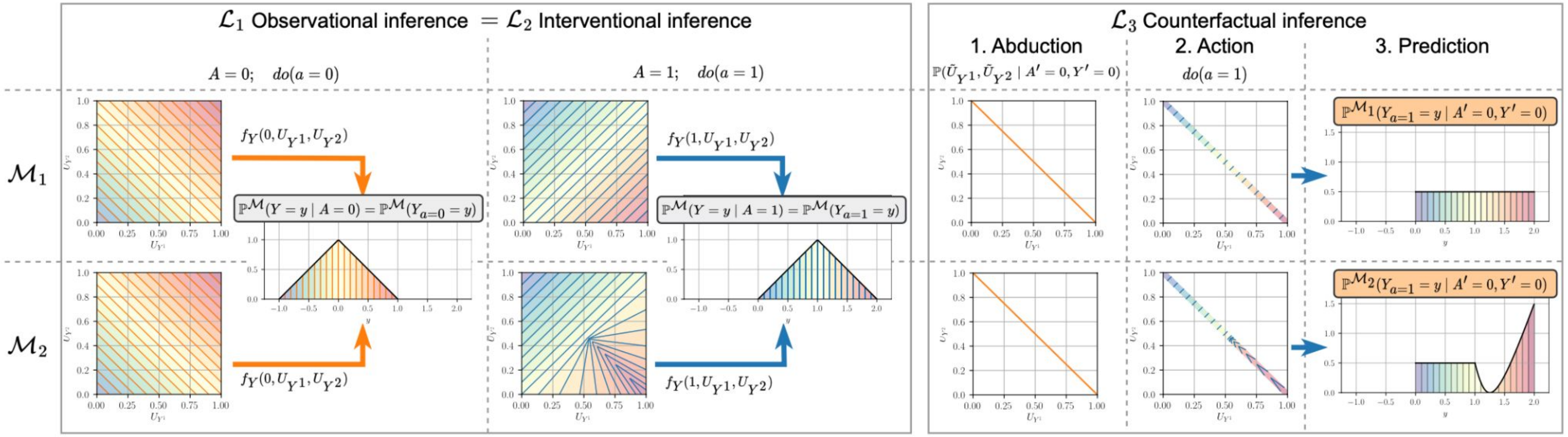


$$\mathcal{M} = \langle \mathbf{U}, \mathbf{V}, \mathbb{P}(\mathbf{U}), \mathcal{F} \rangle$$
$$\mathbf{U} = \{U_A \in \{0,1\}, U_Y \in [0,1]^d\}$$
$$\mathbf{V} = \{A \in \{0,1\}, Y \in \mathbb{R}\}$$
$$\mathbb{P}(\mathbf{U})\text{: } U_A \sim \mathrm{Bern}(p_A),\ 0 < p_A < 1,\quad U_Y \sim \mathrm{Unif}(0,1)^d$$
$$\mathcal{F} = \{f_A(U_A) = U_A, f_Y(A, U_Y)\}$$

+ $f_Y(a, \cdot) \in C^k$

Motivating example

- ECOU [ECOT] is non-identifiable in $\mathfrak{B}(C^k, d)$



$$Q^{\mathcal{M}_1}_{0\to1}(0) = 1$$

$$Q^{\mathcal{M}_2}_{0\to1}(0) \approx 1.114$$

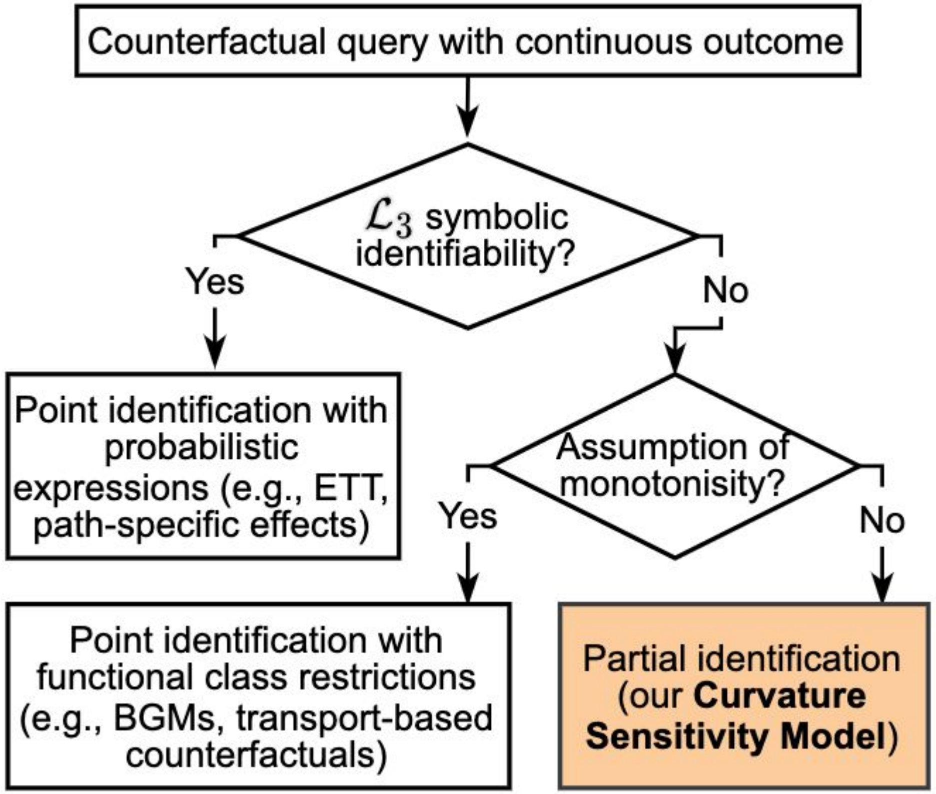
Introduction: Research gap – Our contributions

Research gap

- We are the first to propose a sensitivity model for partial counterfactual identification of continuous outcomes in Markovian SCMs

Our contributions

- We prove that the expected counterfactual outcome of [un]treated has **non-informative bounds** in $\mathfrak{B}(C^k, d)$
- We propose the first sensitivity model, namely, **Curvature Sensitivity Model (CSM),** to obtain informative bounds.
- We introduce a novel deep generative model called **Augmented Pseudo-Invertible Decoder (APID)** to perform partial counterfactual inference under our CSM



Partial Counterfactual Identification: Formulation

- Given the observational distributions, $\mathbb{P}(Y \mid a)$, we want to solve a constrained variational problem, which involves partial derivatives and Hausdorff integrals:

$$\underline{Q_{a' \to a}(y')} = \inf_{\mathcal{M} \in \mathfrak{B}(C^k, d)} Q^{\mathcal{M}}_{a' \to a}(y') \quad s.t. \; \forall a \in \{0, 1\} : \mathbb{P}(Y \mid a) = \mathbb{P}^{\mathcal{M}}(Y \mid a)$$

$$\overline{Q_{a' \to a}(y')} = \sup_{\mathcal{M} \in \mathfrak{B}(C^k, d)} Q^{\mathcal{M}}_{a' \to a}(y') \quad s.t. \; \forall a \in \{0, 1\} : \mathbb{P}(Y \mid a) = \mathbb{P}^{\mathcal{M}}(Y \mid a)$$

Partial counterfactual identification of ECOU [ECOT]

Partial Counterfactual Identification: Formulation

Partial counterfactual identification of ECOU [ECOT]

- Given the observational distributions, $\mathbb{P}(Y \mid a)$, we want to solve a constrained variational problem, which involves partial derivatives and Hausdorff integrals:

$$\underline{Q_{a' \to a}(y')} = \inf_{\mathcal{M} \in \mathfrak{B}(C^k, d)} Q^{\mathcal{M}}_{a' \to a}(y') \quad s.t.\ \forall a \in \{0, 1\} : \mathbb{P}(Y \mid a) = \mathbb{P}^{\mathcal{M}}(Y \mid a)$$

$$\overline{Q_{a' \to a}(y')} = \sup_{\mathcal{M} \in \mathfrak{B}(C^k, d)} Q^{\mathcal{M}}_{a' \to a}(y') \quad s.t.\ \forall a \in \{0, 1\} : \mathbb{P}(Y \mid a) = \mathbb{P}^{\mathcal{M}}(Y \mid a)$$

- Observational distribution is a pushforward distribution:

$$\mathbb{P}^{\mathcal{M}}(Y = y \mid a) = \int_{E(y,a)} \frac{1}{\|\nabla_{u_Y} f_Y(a, u_Y)\|_2} \, \mathrm{d}\mathcal{H}^{d-1}(u_Y).$$

Partial Counterfactual Identification: Formulation

Partial counterfactual identification of ECOU [ECOT]

- Given the observational distributions, $\mathbb{P}(Y \mid a)$, we want to solve a constrained variational problem, which involves partial derivatives and Hausdorff integrals:

$$\underline{Q_{a' \to a}(y')} = \inf_{\mathcal{M} \in \mathfrak{B}(C^k, d)} Q^{\mathcal{M}}_{a' \to a}(y') \quad s.t.\ \forall a \in \{0,1\} : \mathbb{P}(Y \mid a) = \mathbb{P}^{\mathcal{M}}(Y \mid a)$$

$$\overline{Q_{a' \to a}(y')} = \sup_{\mathcal{M} \in \mathfrak{B}(C^k, d)} Q^{\mathcal{M}}_{a' \to a}(y') \quad s.t.\ \forall a \in \{0,1\} : \mathbb{P}(Y \mid a) = \mathbb{P}^{\mathcal{M}}(Y \mid a)$$

- Observational distribution is a pushforward distribution:

$$\mathbb{P}^{\mathcal{M}}(Y = y \mid a) = \int_{E(y,a)} \frac{1}{\|\nabla_{u_Y} f_Y(a, u_Y)\|_2} \, \mathrm{d}\mathcal{H}^{d-1}(u_Y).$$

- Counterfactual queries are expectations of pushforward distributions:

$$\mathbb{P}^{\mathcal{M}}(Y_a = y \mid a', y') = \frac{1}{\mathbb{P}^{\mathcal{M}}(Y = y' \mid a')} \int_{E(y',a')} \frac{\delta(f_Y(a, u_Y) - y)}{\|\nabla_{u_Y} f_Y(a', u_Y)\|_2} \, \mathrm{d}\mathcal{H}^{d-1}(u_Y)$$

$$Q^{\mathcal{M}}_{a' \to a}(y') = \mathbb{E}^{\mathcal{M}}(Y_a \mid a', y') = \frac{1}{\mathbb{P}^{\mathcal{M}}(Y = y' \mid a')} \int_{E(y',a')} \frac{f_Y(a, u_Y)}{\|\nabla_{u_Y} f_Y(a', u_Y)\|_2} \, \mathrm{d}\mathcal{H}^{d-1}(u_Y)$$

Partial Counterfactual Identification: Non-Informative Bounds

Solution for d = 1

- Partial counterfactual identification of ECOU [ECOT] has two solutions in class $\mathfrak{B}(C^1, 1)$ (d = 1, k = 1), when $f_Y(a, \cdot)$ is strictly monotonous:

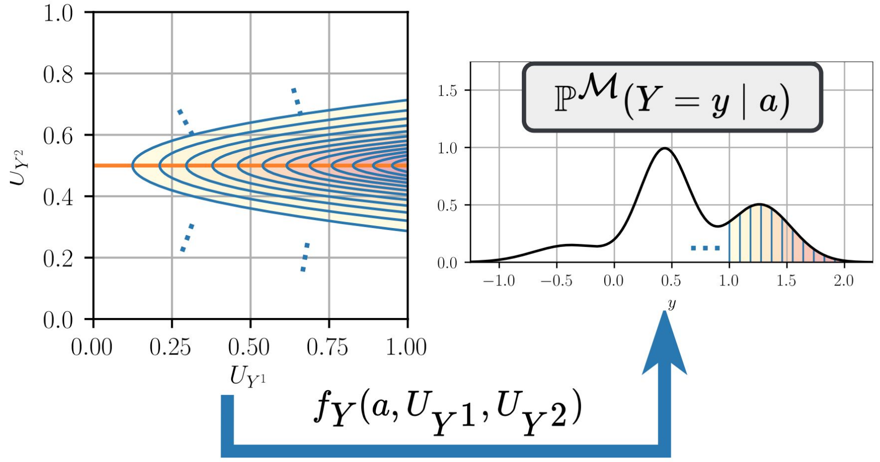
$$Q^{\mathcal{M}}_{a' \to a}(y') = \mathbb{F}_a^{-1}(\pm \mathbb{F}_{a'}(y') \mp 0.5 + 0.5)$$

where $\mathbb{F}_a^{-1}$ *is an inverse CDF of the observational distribution* $\mathbb{P}(Y \mid a)$

- This class is known as **bijective generative mechanisms**[1] (BGMs)

Non-informative bounds

- **Theorem 1 (informal).** The ignorance interval for the partial identification of the ECOU [ECOT] has **non-informative** bounds for SCMs in $\mathfrak{B}(C^k, d)$ for every k > 1



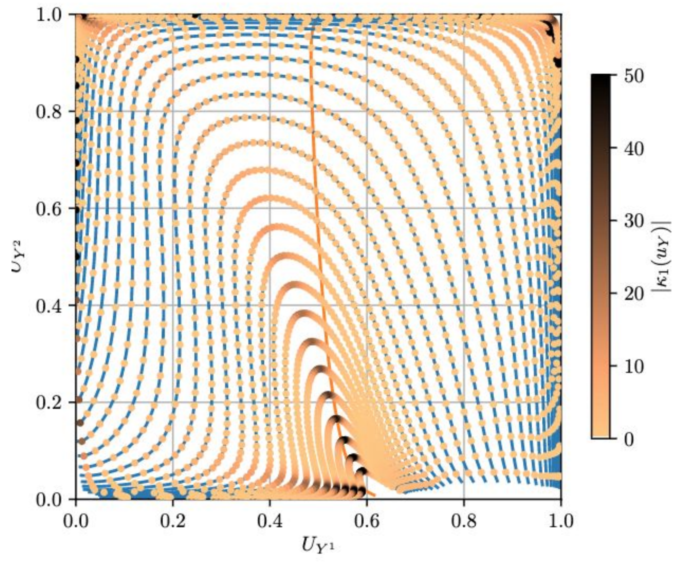
[1] Arash Nasr-Esfahany, Mohammad Alizadeh, and Devavrat Shah. "Counterfactual identifiability of bijective causal models". In: International Conference on Machine Learning. 2023.

CSM: Assumption Kappa - Informative bounds

Curvature sensitivity model (CSM) = Assumption κ

- (Informal) we assume that κ ≥ 0 is the upper bound of the absolute **curvature** for the level sets:

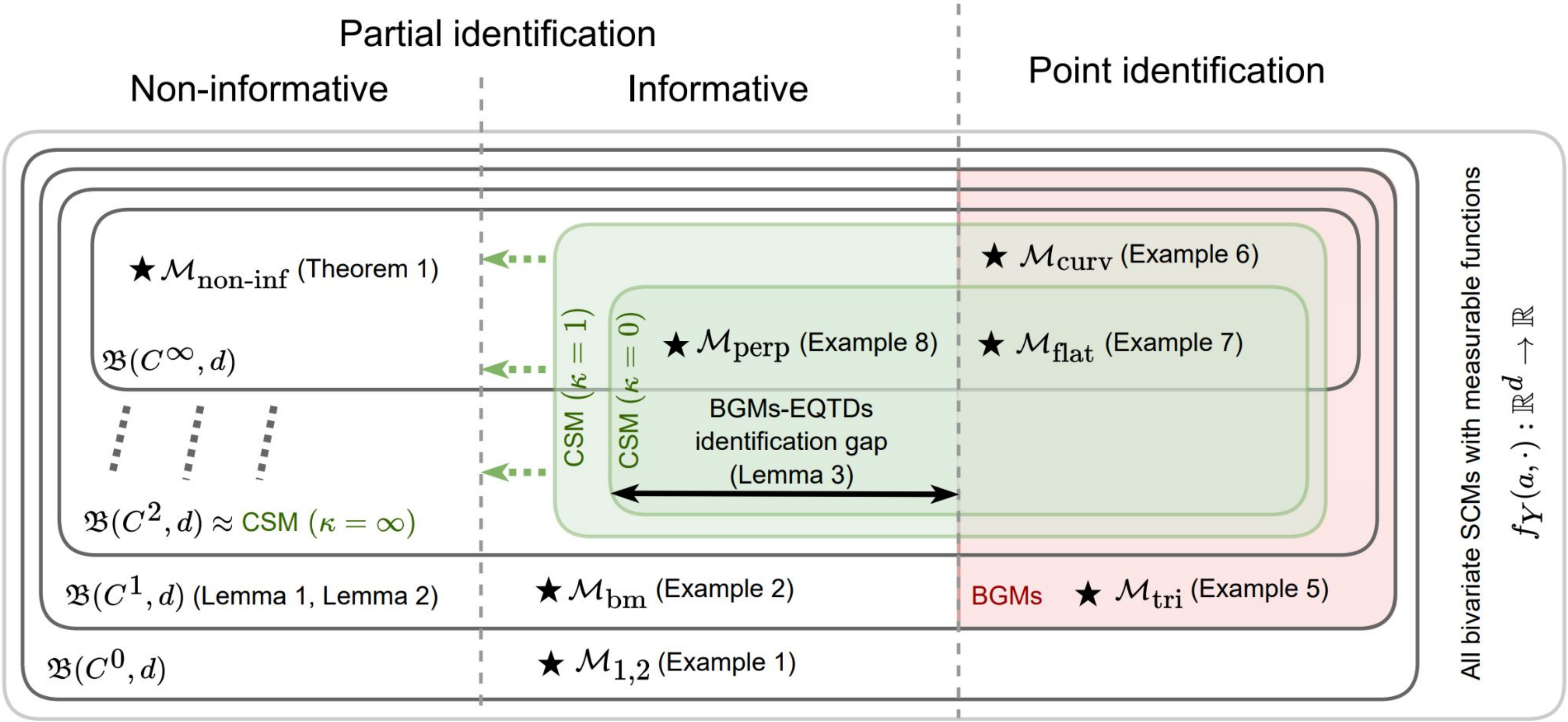
$$\kappa_1(u_Y) = -\frac{1}{2}\nabla_{u_Y}\left(\frac{\nabla_{u_Y} f_Y(a, u_Y)}{\|\nabla_{u_Y} f_Y(a, u_Y)\|_2}\right)$$



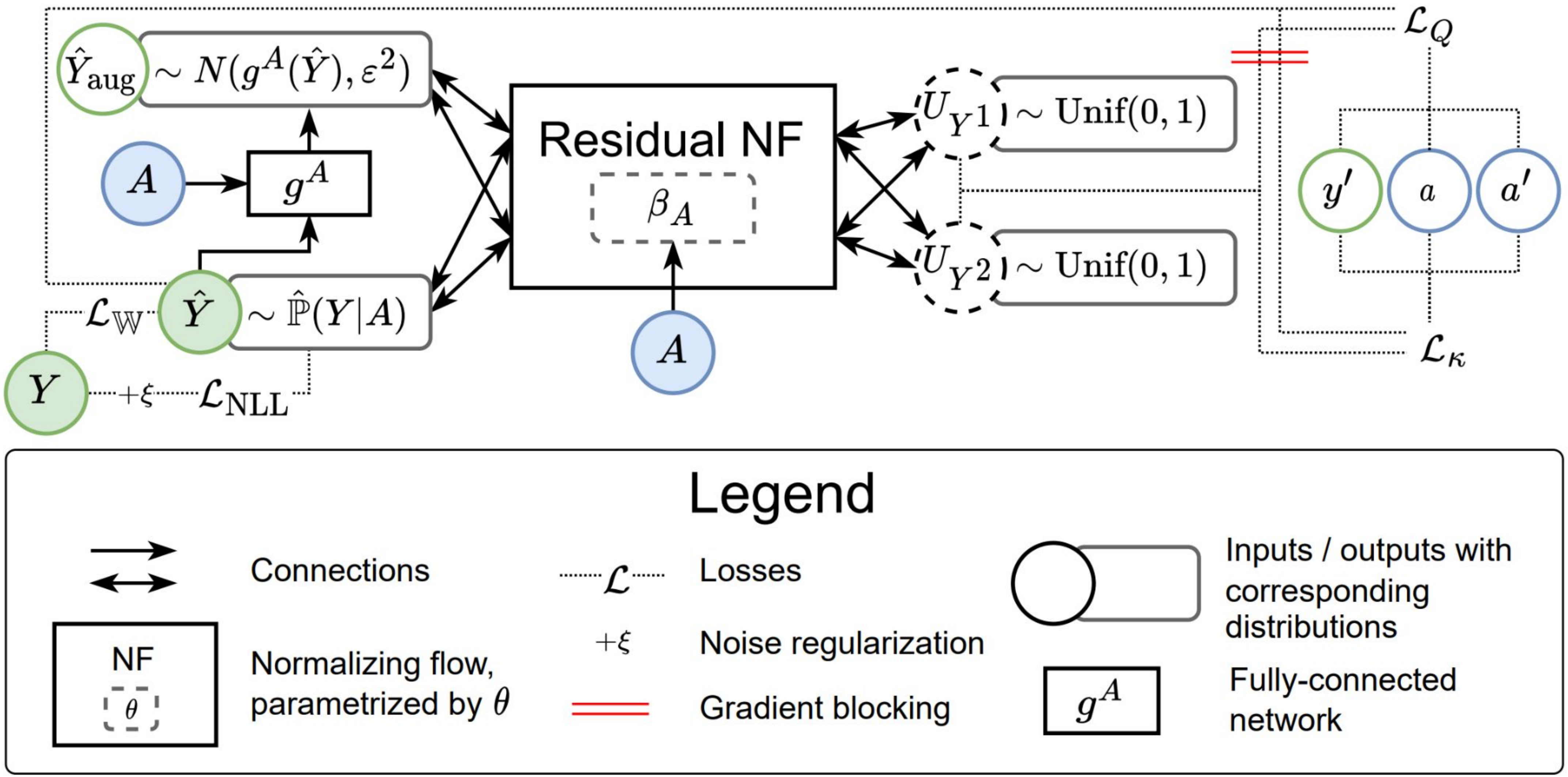
Partial identification with informative bounds

- **Theorem 2 (informal).** Under Assumption κ, the ignorance interval for the partial identification of the ECOU [ECOT] has **informative bounds** for SCMs in $\mathfrak{B}(C^k, d)$ for k = 2 and d > 1
- When κ = 0, we do not obtain a point identification, but a **BGMs-EQTDs identification gap**

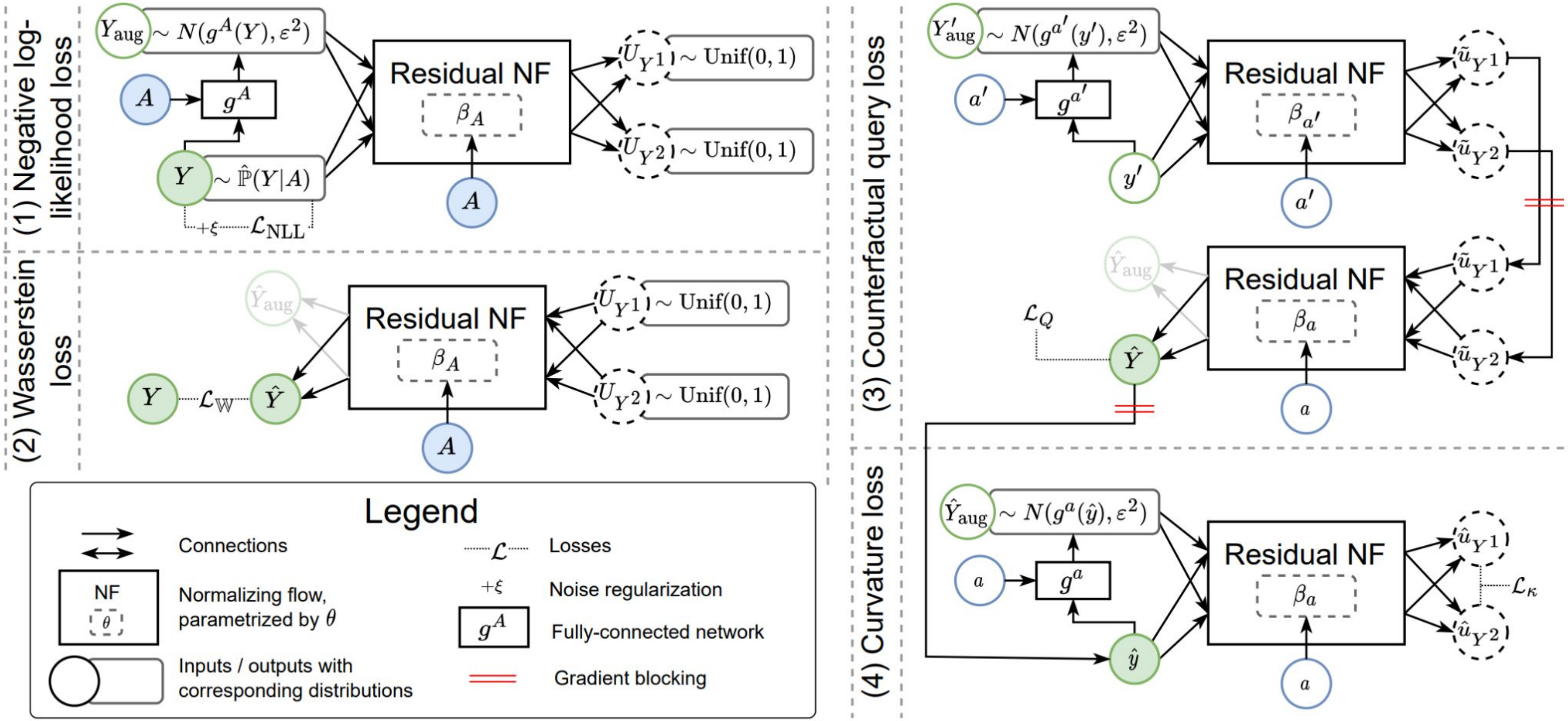
CSM: Identification spectrum



Augmented Pseudo Invertible Decoder: Novel deep generative model



Augmented Pseudo Invertible Decoder: Training



Experiments: Datasets – Results

Datasets

- We evaluate APID based on 2 synthetic dataset and 1 real-world COVID-19 pandemic data

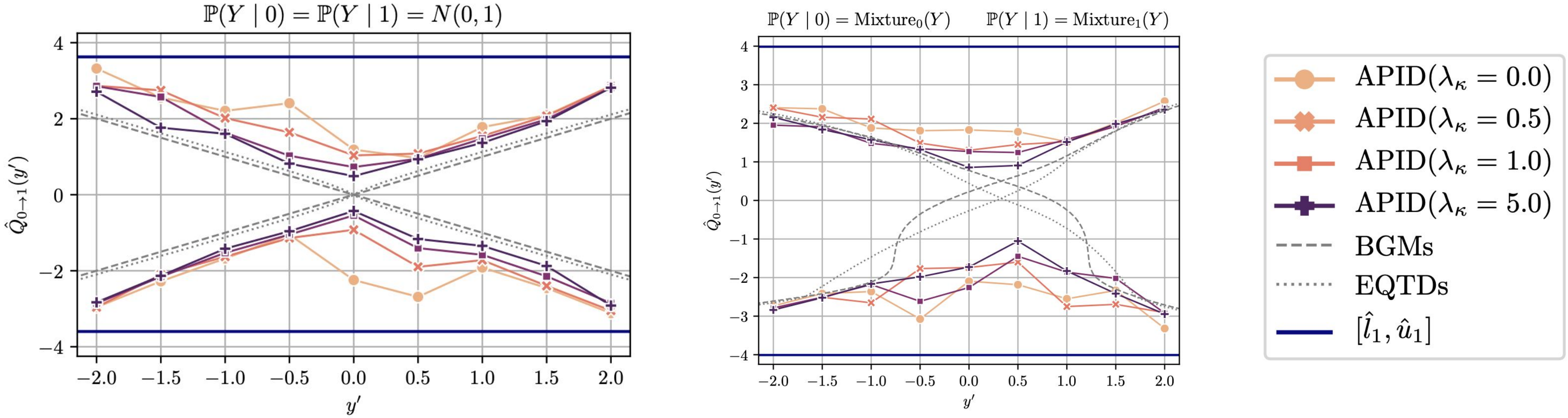
$$\begin{cases} Y \mid 0 \sim \mathbb{P}(Y \mid 0) = N(0,1) \\ Y \mid 1 \sim \mathbb{P}(Y \mid 1) = N(0,1) \end{cases}$$

$$\begin{cases} Y \mid 0 \sim \mathbb{P}(Y \mid 0) = \text{Mixture}(0.7\, N(-0.5, 1.5^2) + 0.3\, N(1.5, 0.5^2)), \\ Y \mid 1 \sim \mathbb{P}(Y \mid 1) = \text{Mixture}(0.3\, N(-2.5, 0.35^2) + 0.4\, N(0.5, 0.75^2) + 0.3\, N(2.0, 0.5^2)) \end{cases}$$

- Even for synthetic data, we the GT counterfactual queries are intractable

Results

- APID is consistent with the BGMs-EQTDs identification gap



LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

LMU MUNICH SCHOOL OF MANAGEMENT

INSTITUTE OF ARTIFICIAL INTELLIGENCE (AI) IN MANAGEMENT

Conclusion

Our work is the first to present a **sensitivity model for partial counterfactual identification** of **continuous outcomes** in Markovian SCMs

Our work rests on the assumption of the **bounded curvature of the level sets**, yet which should be sufficiently broad and realistic to cover many models from physics and medicine

Source Code:
github.com/Valentyn1997/CSM-APID

ArXiv Paper:
arxiv.org/abs/2306.01424