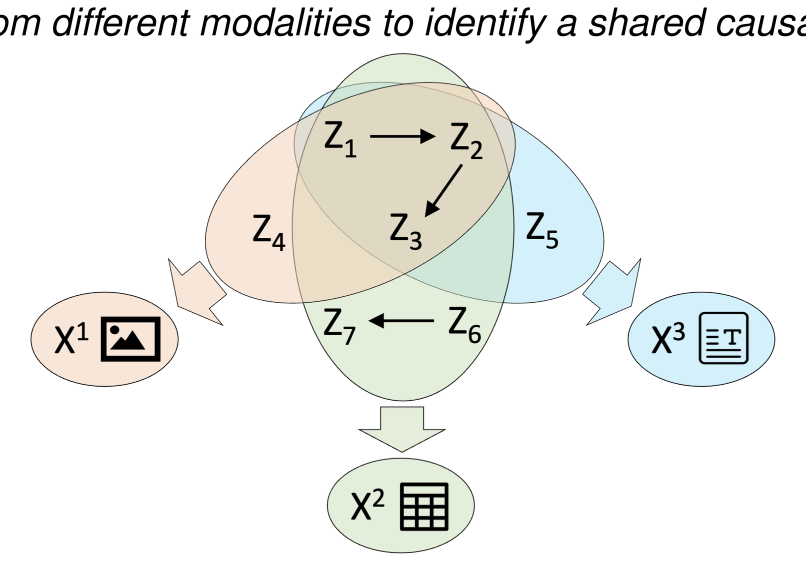
Research group Mathematical Statistics
TUM School of Computation, Information and Technology
Technical University of Munich

Unpaired Multi-Domain Causal Representation Learning

Nils Sturma[1], Chandler Squires[2], Mathias Drton[1], Caroline Uhler[2]

Motivation

"Integrate data from different modalities to identify a shared causal representation."

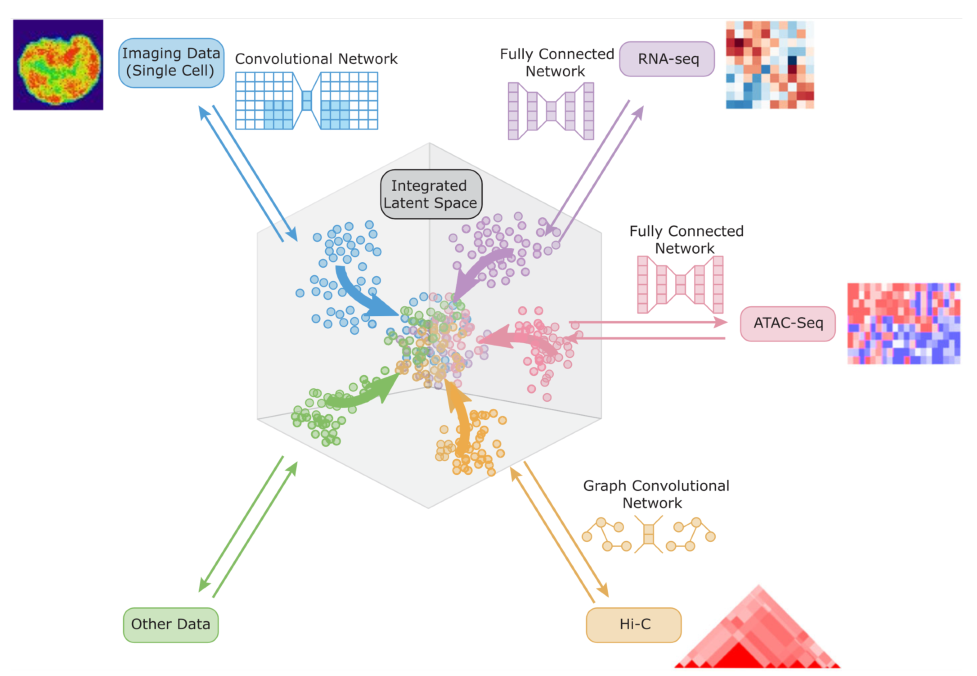


- **Causal representation**: Structural causal model among Z; *Shared* variables $Z_{\mathcal{L}}$ capture key causal relations.
- **Observed domains**: $X^e = g_e(Z_{S_e})$ such that $\mathcal{L} \subseteq S_e$.
- **Unpaired**: Joint distribution of X^e, X^f unknown.

Goal: Identifiability of shared causal representation in a **linear** setup.

Main Contributions:

- Sufficient and necessary conditions for identifiability of joint distribution.
- Sufficient conditions for identifiability of the shared causal structure.



Example:
Single-cell data in biology.

Setup and Graphical Perspective

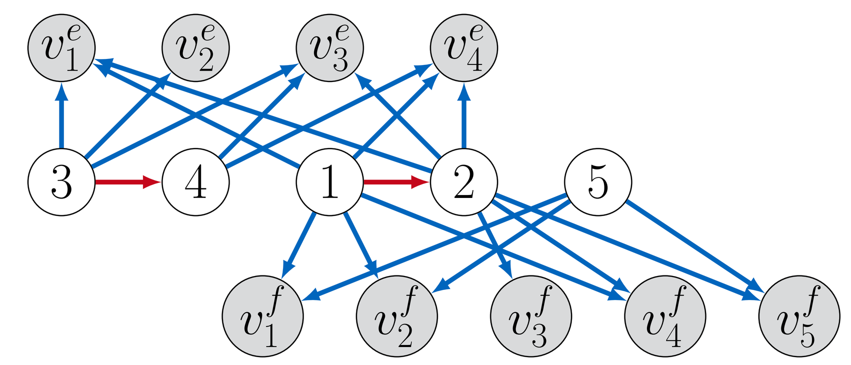
Latent: $(Z_h)_{h\in\mathcal{H}}$, where $Z = AZ + \varepsilon$.

Observed: $X^e \in \mathbb{R}^{d_e}$, where $X^e = G^e \cdot Z_{S_e}$.

$\mathcal{H} \supseteq S_e = \mathcal{L} \cup I_e =$ "shared" and "domain-specific".

m-domain graph:

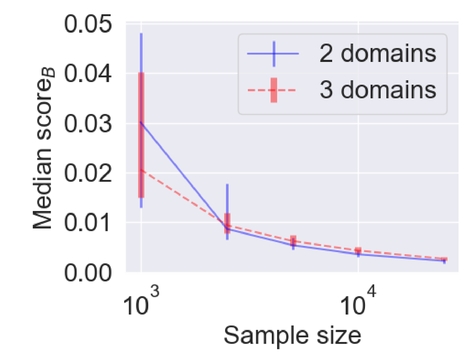
- Nodes $\mathcal{H} \cup V_1 \cup \cdots \cup V_m$, where $|V_e| = d_e$.
- Edges in $\mathcal{H}$ encode sparsity in A (acyclic).
- Edges from $\mathcal{H}$ to V_e encode sparsity in G^e.
- No edges from domain-specific to shared latents.

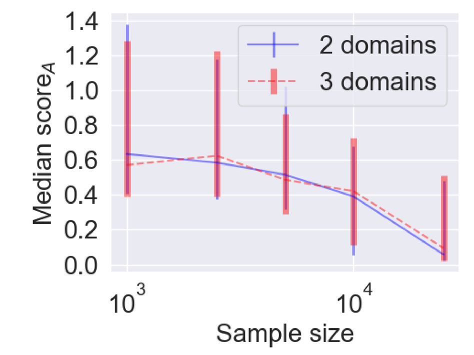


Example: $\mathcal{L} = \{1, 2\}$ and $I_e = \{3, 4\}$, $I_f = \{5\}$.

Important:
The graph, the set $\mathcal{L} \subseteq \mathcal{H}$ and the joint distribution (X^e, X^f) for $e \neq f$ are *unknown*.

Finite Samples





Joint Distribution

Let G be the "large" mixing matrix, that is, $G_{V_e,S_e} = G^e$.

$$\begin{pmatrix} X^1 \\ \vdots \\ X^m \end{pmatrix} = G \cdot Z = \underbrace{G \cdot (I - A)^{-1}}_{=:B} \cdot \varepsilon = \begin{pmatrix} B_{V_1,\mathcal{L}} & B_{V_1,I_1} & & \\ \vdots & & \ddots & \\ B_{V_m,\mathcal{L}} & & & B_{V_m,I_m} \end{pmatrix} \cdot \begin{pmatrix} \varepsilon_{\mathcal{L}} \\ \varepsilon_{I_1} \\ \vdots \\ \varepsilon_{I_m} \end{pmatrix}$$

Approach/ Algorithm:

1. Linear ICA in each domain.
2. Identify shared columns and shared ε_i by matching distributions.
3. Reconstruct B up to unknown (block)-permutation of the columns.

Assumptions:

1) Error distributions P_i of ε_i are **non-symmetric** and **pairwise different** ($P_i \neq P_j$ and $P_i \neq -P_j$ $\forall i, j \in \mathcal{H}$). Additionally: non-degenerate, mean zero, unit variance and independent.

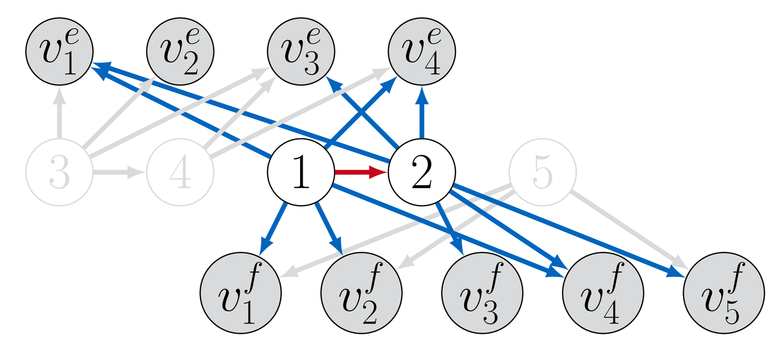
2) The matrix G_{V_e,S_e} is of full column rank for each $e = 1, \ldots, m$.

Theorem: "B and $P = (P_h)_{h\in\mathcal{H}}$ are identifiable up to signed block-permutation."

Shared Latent Graph

We identify the shared causal graph $\mathcal{G}_{\mathcal{L}}$ from the matrix $\widehat{B}_{V,\mathcal{L}} = B_{V,\mathcal{L}}\Psi$, where Ψ is a signed permutation matrix.

Definition: $v \in V$ is a partial pure child of $h \in \mathcal{L}$ if $\text{pa}(v) \cap \mathcal{L} = \{h\}$.



Example:
v_1^f is a partial pure child of 1.

Crucial Observation: rank$(B_{\{v,w\},\mathcal{L}}) = 1 \iff$ there is a node $h \in \mathcal{L}$ such that both v and w are partial pure children of h.

Assumptions:

3) Each shared latent node $h \in \mathcal{L}$ has **two partial pure children.**

4) Rank faithfulness of $B_{V,\mathcal{L}}$

Theorem: "The graph $\mathcal{G}_{\mathcal{L}}$ and A are identifiable up to a (signed) permutation that is consistent with the true graph $\mathcal{G}_{\mathcal{L}}$."

[1]Technical University of Munich; Munich Center for Machine Learning
[2] LIDS, Massachusetts Institute of Technology; Broad Institute of MIT and Harvard

NeurIPS 2023
December 10-16, 2023, New Orleans