Discrete Causal Representations from Heterogeneous Domains

A Bayesian Approach with Social Survey Applications

Ankur Garg[1] Michael Stettler[2] Aaron Schein[1] Julius von Kügelgen[3]

[1]Department of Statistics, University of Chicago, [2] University of Tübingen, [3]Seminar for Statistics, ETH Zürich

Contact: garga@uchicago.edu

arXiv:2606.06288

Causal representation learning[1] aims to extending causal modeling to settings in which the causal variables of interest are latent and only indirectly observed. Existing work in this area has primarily focused on identifiability under idealized assumptions, while practical methods for finite, real-world data remain less developed. **This work is a step towards bridging that gap and provides a successful application of a complete unsupervised CRL pipeline to complex real-world data.**

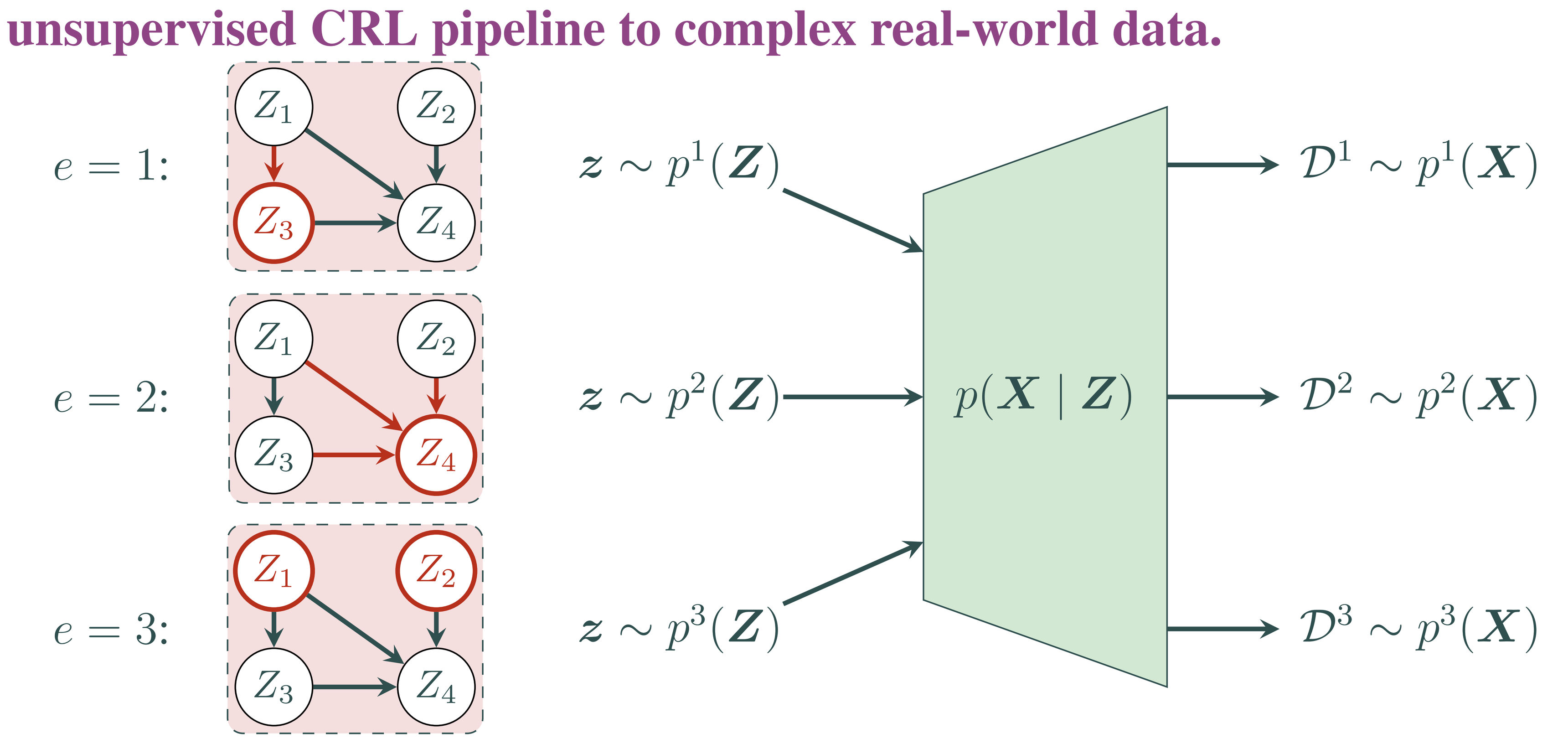

Figure 1: Multi-domain CRL with soft, multi-node interventions

Problem Setting

- We **observe samples** $\boldsymbol{x}^{e,j} \in \mathbb{R}^D$ from environments $e \in [E]$
- Each $\boldsymbol{x}^{e,j}$ is generated by some **discrete latent variables** $z^{e,j} \in [K]^L$
- Changes across environments are caused by **sparse interventions**.

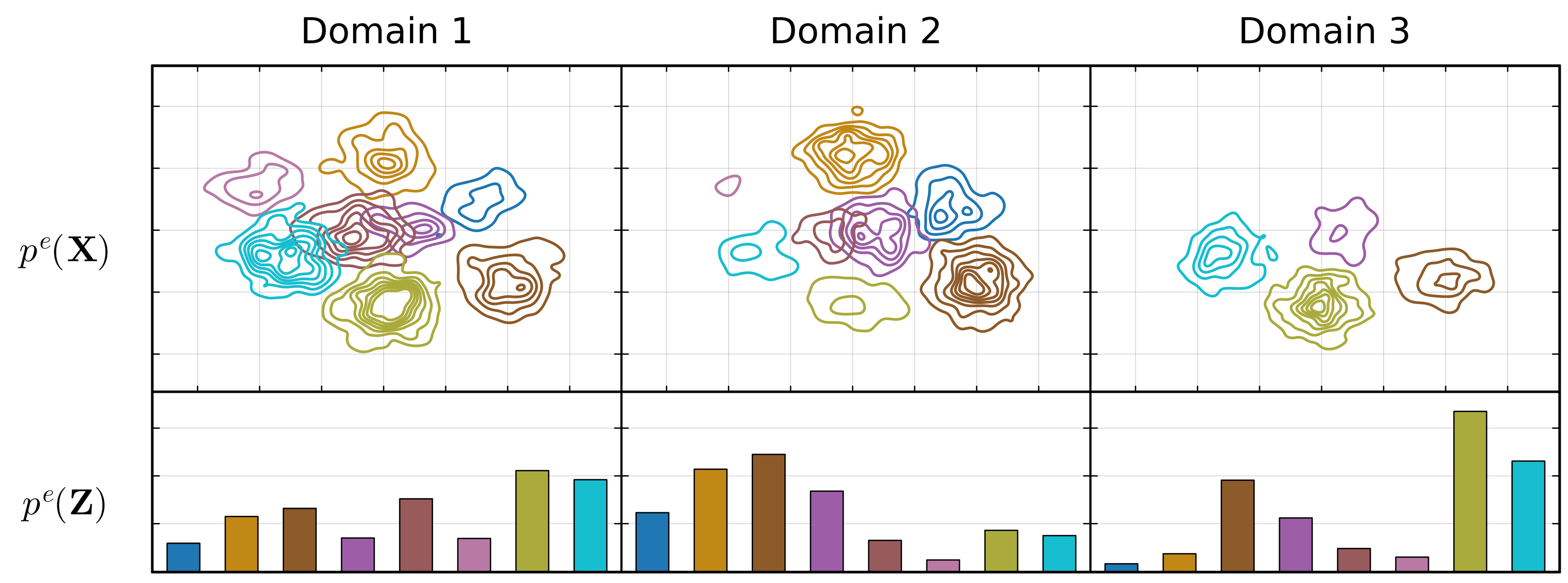

Figure 2: Observations and latents in multi-domain data

Measurement Model

$$\boldsymbol{x}^{e,j} \sim \mathcal{N}_D\left(\boldsymbol{m}_0 + \textstyle\sum_{\ell \in [L]} \boldsymbol{m}_\ell z_\ell^{e,j}, \operatorname{diag}(\boldsymbol{\sigma}^2)\right)$$
$$\boldsymbol{m}_\ell \sim \mathcal{N}_D(\boldsymbol{0}, \boldsymbol{I}), \quad \sigma_d^2 \sim \text{Inv}-\Gamma(\alpha_\sigma, \beta_\sigma)$$

Latent Mechanisms

$$\boldsymbol{\theta}_\ell(\mathbf{pa}_\ell) \sim \mathcal{N}_K(\boldsymbol{0}, \boldsymbol{\Sigma}_\Theta)$$
$$\mathcal{I}_\ell^e \sim \text{Beta-Bernoulli}(\alpha_c, \beta_c)$$
$$\boldsymbol{\delta}_\ell^e(\mathbf{pa}_\ell) \sim p(\boldsymbol{\delta} \mid \boldsymbol{\Sigma}_\Delta, \varepsilon) \propto \mathcal{N}_K(\boldsymbol{0}, \boldsymbol{\Sigma}_\Delta) \cdot \mathbb{1}\left\{\tfrac{1}{2}\boldsymbol{\delta}^\top \boldsymbol{H}\boldsymbol{\delta} \geq \varepsilon\right\}$$
$$z_\ell^{e,j} \sim \text{Categorical}\left(\text{softmax}\left(\boldsymbol{\theta}_\ell(\mathbf{pa}_\ell^{e,j}) + \mathcal{I}_\ell^e \boldsymbol{\delta}_\ell^e(\mathbf{pa}_\ell^{e,j})\right)\right)$$

Sparse Detectable Interventions

- **Sparsity:** The Beta-Bernoulli prior on $\mathcal{I}_\ell^e$ induces sparse interventions.
- **Detectability:** The truncation prior on $\boldsymbol{\delta}_\ell^e$ rules out arbitrarily small shifts by requiring every active intervention to induce an approximate KL change of at least ε: $\text{KL}(\text{softmax}(\boldsymbol{\theta}) \,\|\, \text{softmax}(\boldsymbol{\theta} + \boldsymbol{\delta})) \approx \frac{1}{2}\boldsymbol{\delta}^\top \boldsymbol{H}\boldsymbol{\delta} \geq \varepsilon$.

Posterior Landscape

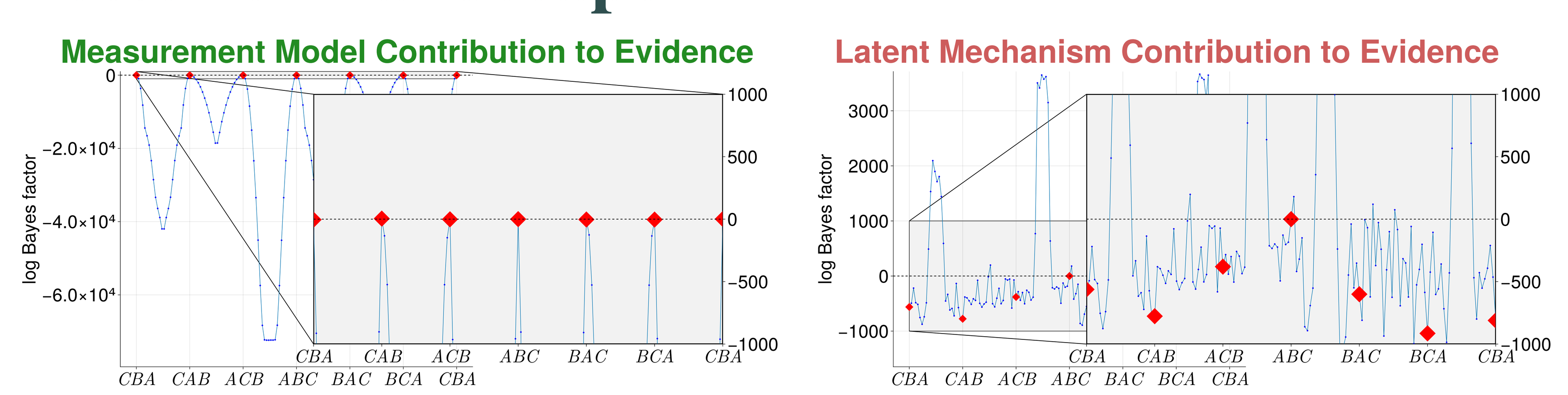

Figure 3: Multimodality across permutations of the causal ordering

- Competing posterior modes correspond to different causal orderings.
- The **Measurement Model** is permutation-invariant, creating the modes.
- The **Latent mechanisms** identifies the true causal order among them.

Case Study I: World Values Survey (WVS)

- We analyze wavs 7 of the WVS[2], a large-scale cross-national survey.
- We recover interpretable structure that matches global cultural variation.

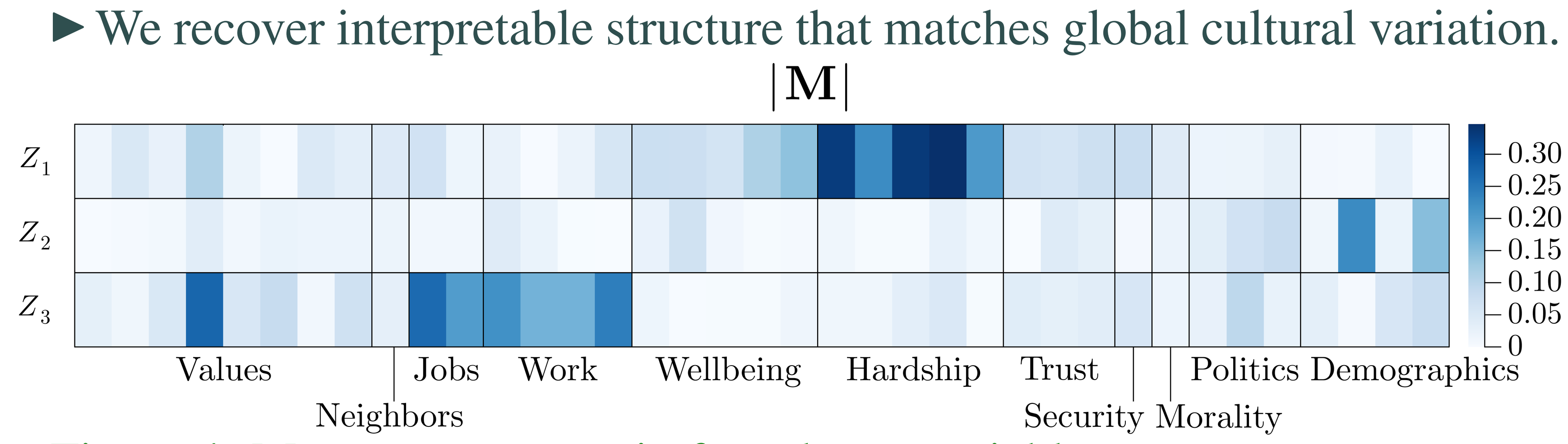

Figure 4: Measurement matrix from latent variables to survey responses.

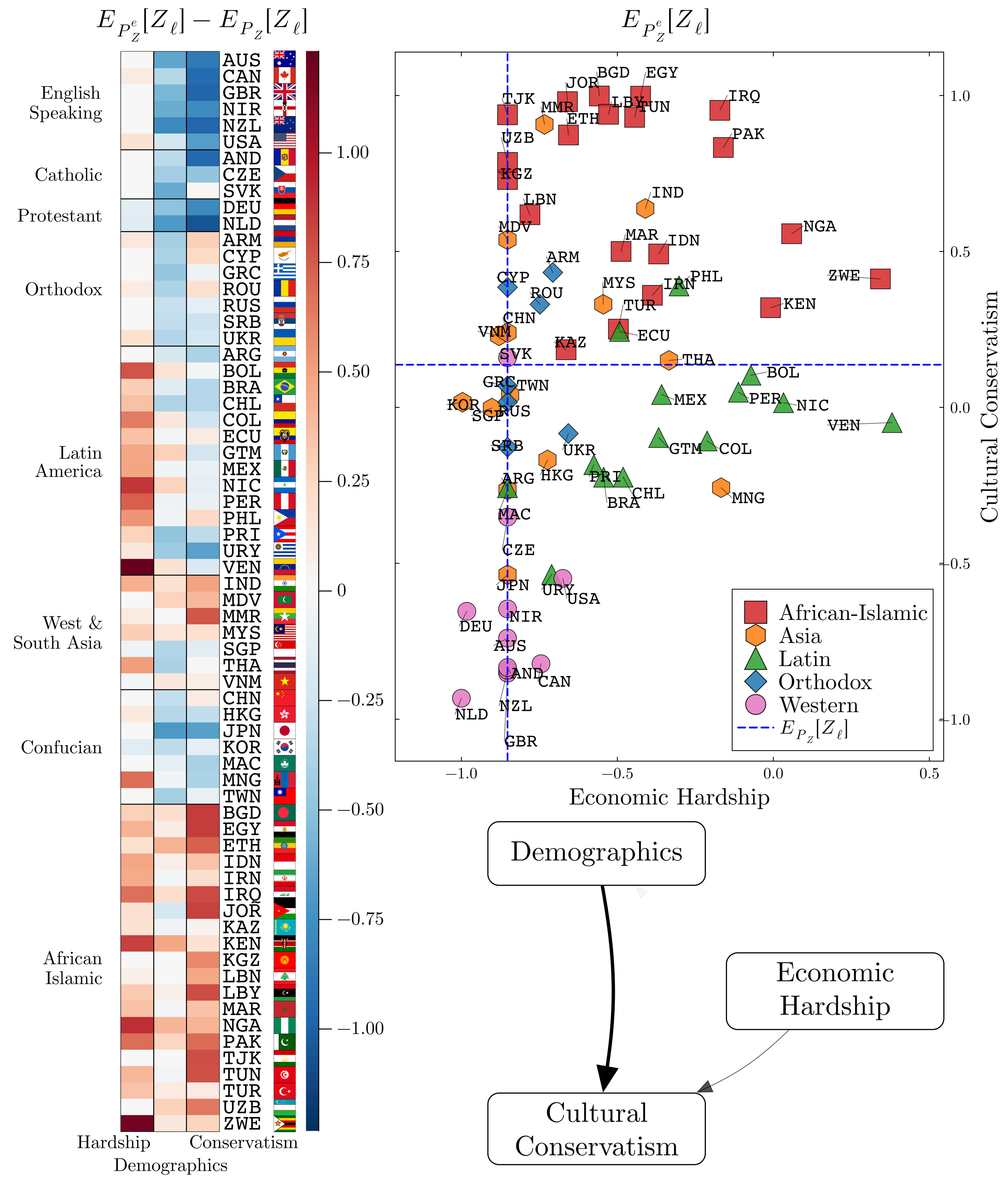

Figure 5: Marginal distributions and causal influence of latent concepts.

Sequential Monte Carlo Sampling (SMCS)

- The model admits efficient **closed-form conjugate updates**.
- Standard MCMC chains can get trapped in one causal ordering mode.
- SMCS[3] uses many weighted particles to explore modes in parallel.
- Likelihood tempering allows latent mechanisms to guide inference.

$$p_{T_k}(\underbrace{\boldsymbol{\Theta}, \mathcal{I}, \boldsymbol{\Delta}}_{\boldsymbol{\Phi}}, \underbrace{\boldsymbol{M}, \boldsymbol{\sigma}^2}_{\boldsymbol{\Psi}}, \boldsymbol{Z}^{\mathcal{E}} \mid \boldsymbol{X}^{\mathcal{E}}) \propto p(\boldsymbol{\Phi})p(\boldsymbol{\Psi})p\left(\boldsymbol{Z}^{\mathcal{E}}, \boldsymbol{X}^{\mathcal{E}} \mid \boldsymbol{\Phi}, \boldsymbol{\Psi}\right)^{\frac{1}{T_k}}.$$

Algorithm 1: Sequential Monte Carlo Sampler with Likelihood Tempering

Require: Temperatures $\{\infty = T_0 > \cdots > T_K = 1\}$, number of particles S, tempered posterior $p_{T_k}(\boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{Z} \mid \boldsymbol{X})$, and tempered p_{T_k}-invariant MCMC kernels (GibbsUpdate).

1: $(\boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{Z})_0^{(s)} \overset{\text{i.i.d.}}{\sim} p_{T_0}(\boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{Z})$ for all $s \in [S]$ ▷ Initialize particles
2: $w_0^{(s)} \leftarrow 1/S$ for all $s \in [S]$ ▷ Initialize weights
3: **for** $k = 1$ **to** K **do**
4: Set weights $w_k^{(s)} \propto w_{k-1}^{(s)} \frac{p_{T_k}(\boldsymbol{\Phi}^{(s)}, \boldsymbol{\Psi}^{(s)}, \boldsymbol{Z}^{(s)} \mid \boldsymbol{X})}{p_{T_{k-1}}(\boldsymbol{\Phi}^{(s)}, \boldsymbol{\Psi}^{(s)}, \boldsymbol{Z}^{(s)} \mid \boldsymbol{X})}$ for all $s \in [S]$ ▷ Reweight
5: $\{(\boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{Z})_k^{(s)}, w_k^{(s)}\}_{s=1}^S \leftarrow \text{Resample}\left(\left\{(\boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{Z})_k^{(s)}, w_k^{(s)}\right\}_{s=1}^S\right)$ ▷ Resample
6: $(\boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{Z})_k^{(s)} \leftarrow \text{GibbsUpdate}\left((\boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{Z})_k^{(s)}, T_k\right)$ for all $s \in [S]$ ▷ Mutate
7: **end for**
8: **return** $\left\{\left((\boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{Z})_K^{(s)}, w_K^{(s)}\right)\right\}_{s=1}^S$ as approximate posterior samples

[1] Bernhard Schölkopf, Francesco Locatello, Stefan Bauer, Nan Rosemary Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. Toward causal representation learning. *Proceedings of the IEEE*, 109(5):612–634, 2021.

[2] Christian Haerpfer, Ronald Inglehart, Alejandro Moreno, Christian Welzel, Kseniya Kizilova, Jaime Diez-Medrano, Marta Lagos, Pippa Norris, Eduard Ponarin, Bjorn Puranen, et al. World values survey: Round seven-country-pooled datafile version 5.0, 2022.

[3] Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. Sequential monte carlo samplers. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 68(3):411–436, 2006.