UAI 2026 42nd Conference on Uncertainty in Artificial Intelligence · Amsterdam

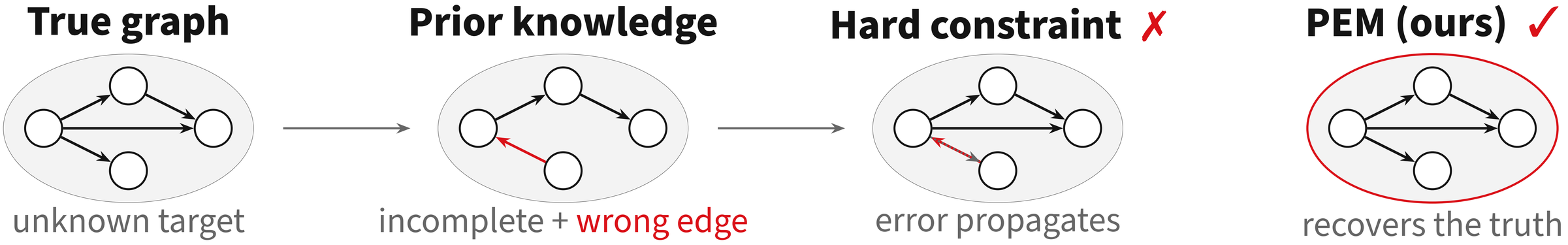
Probabilistic Edge Modulation for High-Dimensional Causal Discovery

Seyong Hwang[1] Kyoungjae Lee[2] Sunmin Oh[3] Gunwoong Park[3]

[1]University of Michigan [2]Sungkyunkwan University [3]Seoul National University

Paper on OpenReview

KEY IDEA **Prior structure as graded beliefs, not hard constraints: one spike-and-slab dial per edge, with high-dimensional consistency guarantees.**



1 Encode beliefs as $\eta_{j,k}$

LBN model:

$$X = BX + \epsilon$$

where $B_{j,k} \neq 0$ denotes $k \to j$.

Prior scores $w_{k\to j} \in [0,1]$ map to edge-specific inclusion strengths

$$\eta_{j,k} = \max\{w_{j\to k},\ w_{k\to j}\}$$

- $\eta_{j,k}$ **continuously** interpolates between exclusion and inclusion of edge (j,k).
- Hard constraints = boundary cases $\eta_{j,k} \in \{0,1\}$: PEM **generalizes** them.

2 Modulate precision

Edge-specific spike-and-slab prior on $\Omega_{j,k}$ $(j<k)$, with $v_1 > v_0 > 0$:

$$\Omega_{j,k} \mid r_{j,k}=0 \sim \text{Laplace}(v_0) \text{ spike}$$
$$\Omega_{j,k} \mid r_{j,k}=1 \sim \text{Laplace}(v_1) \text{ slab}$$
$$r_{j,k} \sim \text{Bern}(\eta_{j,k})$$

One unified **PEM-MAP estimator** per elimination step r:

$$\widehat{\Omega}^{(r)} = \arg\min_{\Theta \succ 0} \left\{ \frac{n}{2} D_{\text{LD}}\left(\widehat{\Sigma}^{(r)} \,\middle\|\, \Theta^{-1}\right) + \mathcal{P}_\eta(\Theta) \right\}$$

- η shrinks **off-diagonals** and shifts the ordering-deciding **diagonals** — priors guide **both ordering and parent selection**.

3 Learn the DAG

Backward elimination over r on active set $S(r)$:

Estimate PEM-MAP $\widehat{\Omega}^{(r)}$

Order $\widehat{\pi}_{p+1-r} \leftarrow \arg\min_{j \in S(r)} \left[\widehat{\Omega}^{(r)}\right]_{j,j}$

Select parents with posterior inclusion probability $p^{(r)}_{j,k} \geq T$ (default T=0.5):

$$p^{(r)}_{j,k} = \frac{\frac{\eta_{j,k}}{2v_1} e^{-|\widehat{\Omega}^{(r)}_{j,k}|/v_1}}{\frac{\eta_{j,k}}{2v_1} e^{-|\widehat{\Omega}^{(r)}_{j,k}|/v_1} + \frac{1-\eta_{j,k}}{2v_0} e^{-|\widehat{\Omega}^{(r)}_{j,k}|/v_0}}$$

- $O(p^4)$ total — like **LISTEN / BHLSM**; closed form, no search, no MCMC.

Theory

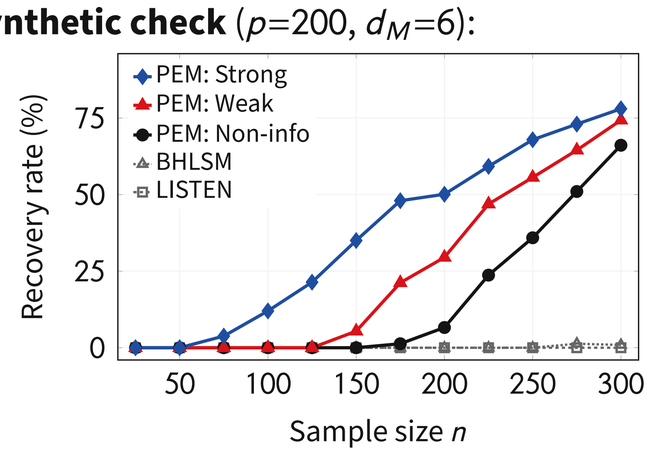
Theorem (High-dim. consistency). Under structural separation & estimation stability, $\Pr(\widehat{G} = G) \to 1$ as soon as

$$n = \Omega\left(d_M^2 \log p\right) \text{ (sub-Gaussian)}$$
$$n = \Omega\left(d_M^2\, p^{2/m}\right) \text{ (4}m\text{-th moment)}$$

- Same rates as prior-free methods, but with **no max-indegree bound** and no incoherence conditions.
- Informative η widens the ordering gaps: priors **provably relax** the finite-sample burden.

Proposition (Oracle prior). $\eta_{j,k} = 1\{(j,k) \in E\}$ $\Rightarrow$ exact recovery, deterministically.

Synthetic check (p=200, d_M=6):

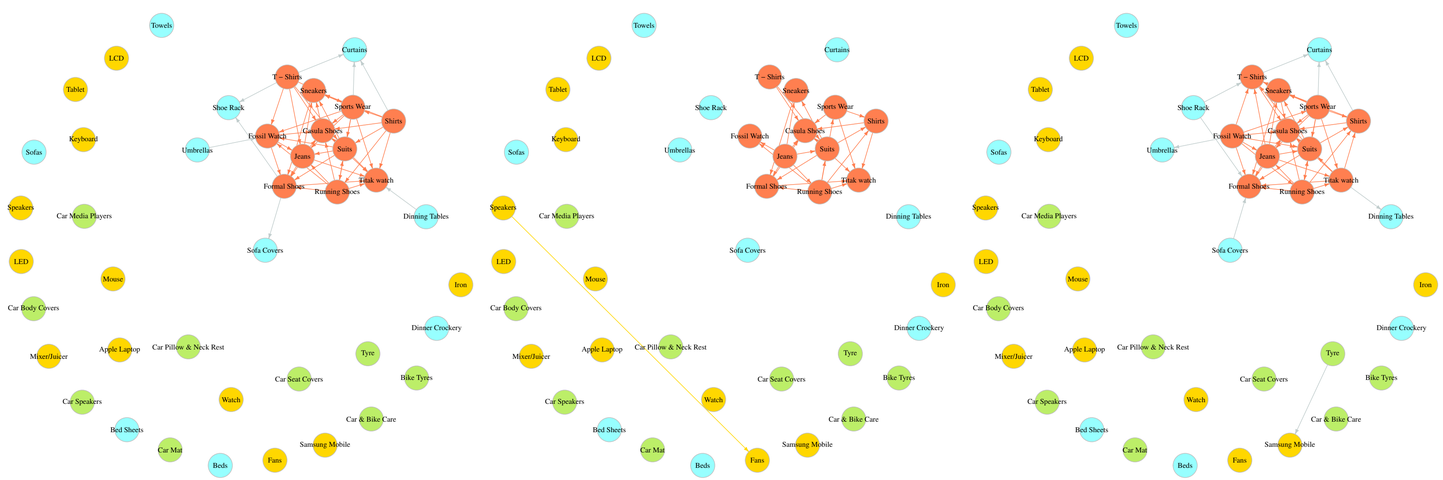


- Large gains in **small-n, dense regimes**.
- Gains survive flipping 20% of η; heavier noise decays to the no-prior baseline.

Real Data: Prior Transfer on Sparse E-commerce

Kaggle e-commerce: 38,997 customers × 42 products, split by payment method.

- Sparse, zero-inflated counts ⇒ data-only fits of the small groups collapse to **near-empty graphs**.
- PEM: transfer the credit-card reference DAG as **soft beliefs** η.



credit card, n=31,055 (reference) money order, n=9,089 e-wallet, n=2,746

- Without the prior, each group collapses to **a single edge**; with it, 28 edges (money order) and 51 (e-wallet).
- **Not a copy** of the reference: only ~1/3 of its edges are shared.
- The reference's 8 cross-category links are **reflected differently**: none in money order, 9 in e-wallet.
- Directions are interpretable: high-volume products become **out-degree hubs** (Spearman 0.80–0.88).