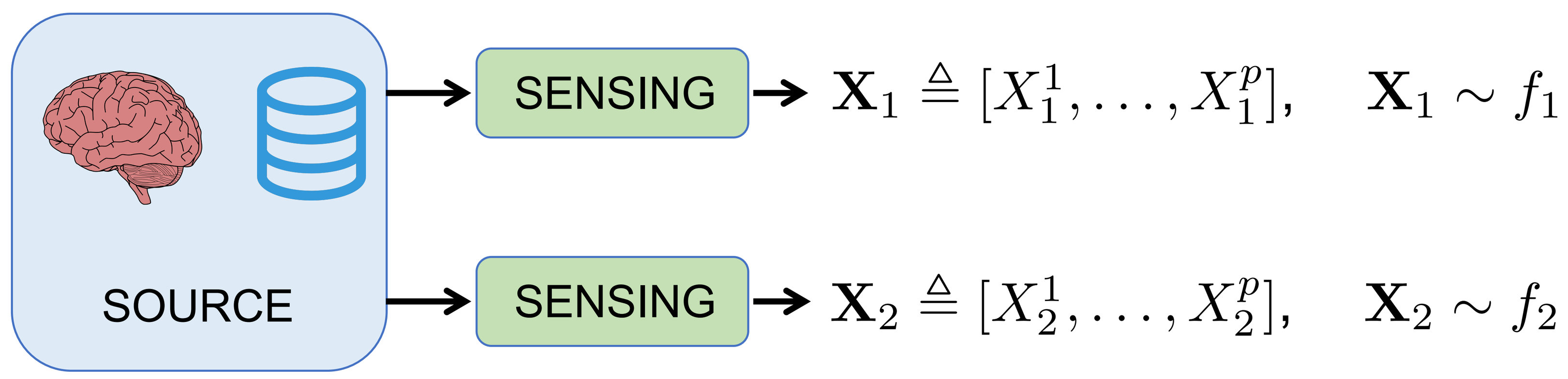
Learning Shared Subgraphs in Ising Model Pairs

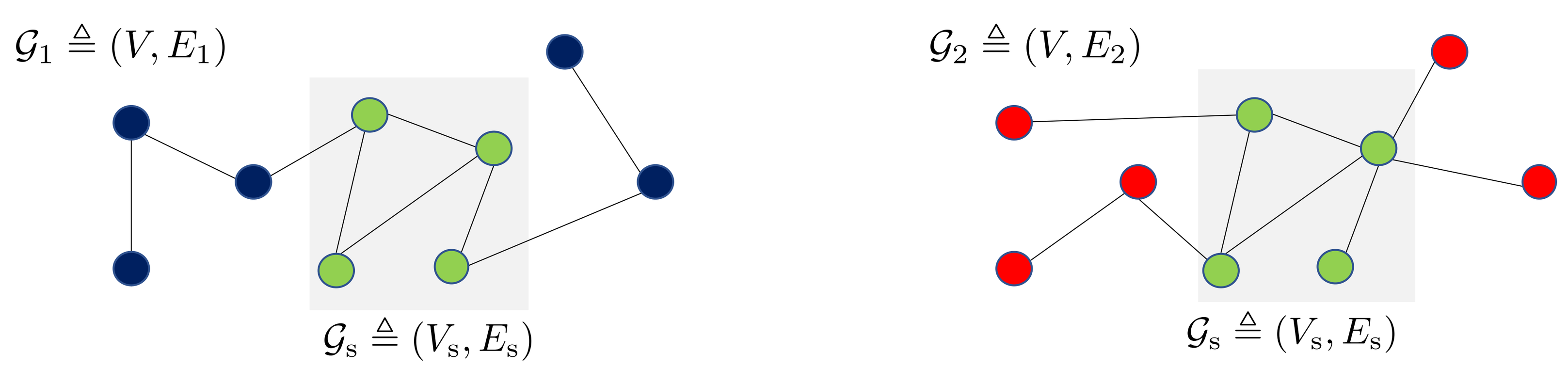
Burak Varıcı, Saurabh Sihag, and Ali Tajer

Rensselaer Polytechnic Institute

Multiple Information Networks



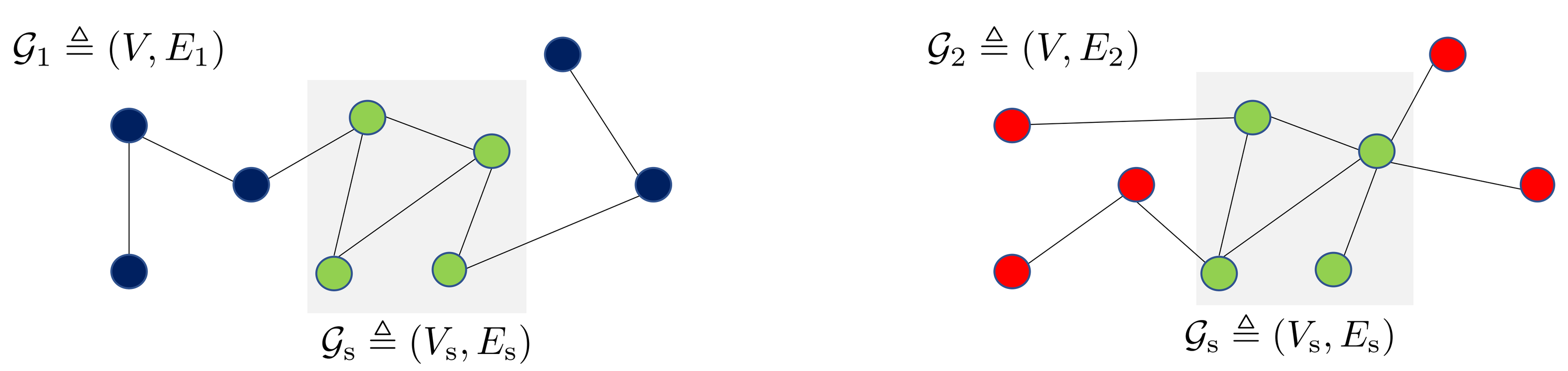
- Graphical model: describe complex dependency structures.
- Information layer → distinct data distribution → distinct representation.



Green vertices and edges between : $\mathcal{G}_s \triangleq (V_s, E_s)$

- **Shared subgraphs – joint information**
- Multiple brain imaging techniques → how to utilize together?
- Finding similar molecular structures for drug discovery.

Structure Learning



- Graph samples: $\mathbf{X}_1 \triangleq [X_1^1, \ldots, X_1^p], \quad \mathbf{X}_1 \sim f_1$
  $\mathbf{X}_2 \triangleq [X_2^1, \ldots, X_2^p], \quad \mathbf{X}_2 \sim f_2$
- Objective: Observe $\mathbf{X}_1, \mathbf{X}_2 \rightarrow$ estimate $\mathcal{G}_s = (V_s, E_s)$

  Estimate $E_1, E_2 \rightarrow E_s = E_1 \cap E_2$: vastly inefficient
- Introduce joint learning of only $\mathcal{G}_s$
- Ising Model: $f(\mathbf{X}) = \frac{1}{Z} \exp\left(\sum_{(u,v)\in E} \lambda X^u X^v \right)$

Problem Formulation

n graph samples $\mathbf{X}_1^n, \mathbf{X}_2^n$ → graph decoder → shared subgraph estimates

$$\psi_s = \{-1,+1\}^{n\times p} \times \{-1,+1\}^{n\times p} \rightarrow (V_s, E_s)$$

- Exact recovery: Perfectly learn $\mathcal{G}_s$

$$\mathrm{P_L}(\mathcal{I}_p^s) \triangleq \max_{\mathcal{G}_1,\mathcal{G}_2 \in \mathcal{I}_p} \mathbb{P}(|E_s \Delta \hat{E}_s| \neq 0)$$

- Vertex sample complexity: $\mathrm{N}(n_{\mathrm{T}}) = \sum_{k=1}^{n_{\mathrm{T}}} |\hat{V}_s(k)|$

  $\downarrow$ number of samples $\quad \downarrow$ adaptive V_s estimation

Pruning

- Form coarse estimates $\hat{V}_s(k), \hat{E}_s(k)$ at each iteration k with the rule:

$$\min_{i\in\{1,2\}} \bar{\mathbb{E}}_k[X_i^u X_i^v] > \tanh\lambda - \sqrt{\alpha \log p / 2k}$$

- Importance: Narrow down sampling to **only** V_s adaptively, results in significant savings in sampling.
- At any iteration k, $\quad \mathbb{P}(V_s \subseteq \hat{V}_s(k)) \geq 1 - 2p^{2-\alpha}$
- Sample complexity: $k = O\left(\frac{\alpha \log p}{\lambda^2}\right)$ in correlation decay regime ensures

$$\mathbb{P}(\hat{V}_s(k) = V_s) \geq 1 - 2p^{2-\alpha}$$

Joint Learning

- **Joint** multiplicative updates at every iteration

$$w_i^{uv}(k+1) = w_i^{uv}(k) \cdot \exp\left(\frac{\beta}{2}(\ell_1^{uv}(k) + \ell_2^{uv}(k))\right), \qquad u, v \in \hat{V}_s(k)$$

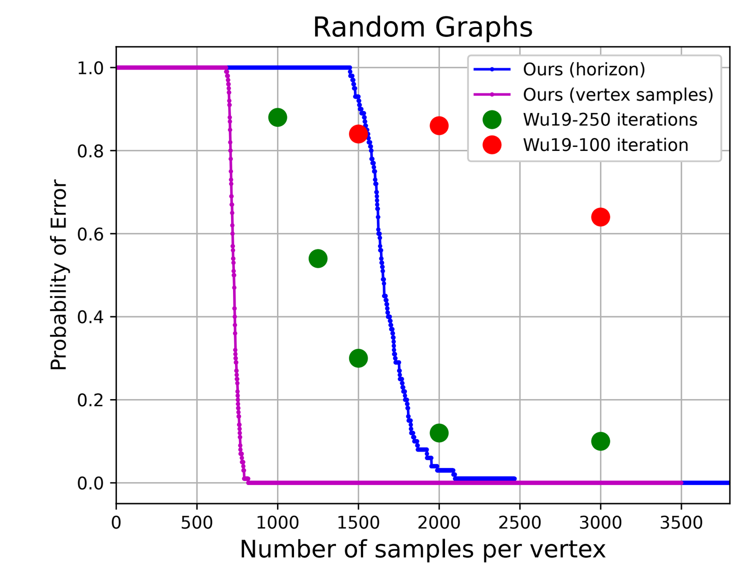
- Importance: Joint updates improve learning of E_s, and samples from only $\hat{V}_s(k)$ suffice.
- Sample complexity: When $\mathcal{G}_s$ is isolated, and pruning localizes V_s, for ensuring $\mathrm{P_L}(\mathcal{I}_p^s) \leq (1 - \frac{2}{\rho})$,

Joint (ours): $O\left(\frac{1}{\lambda^2}\exp(\lambda d) \log \frac{\rho q}{\lambda}\right)$ where $q = |V_s| < p$

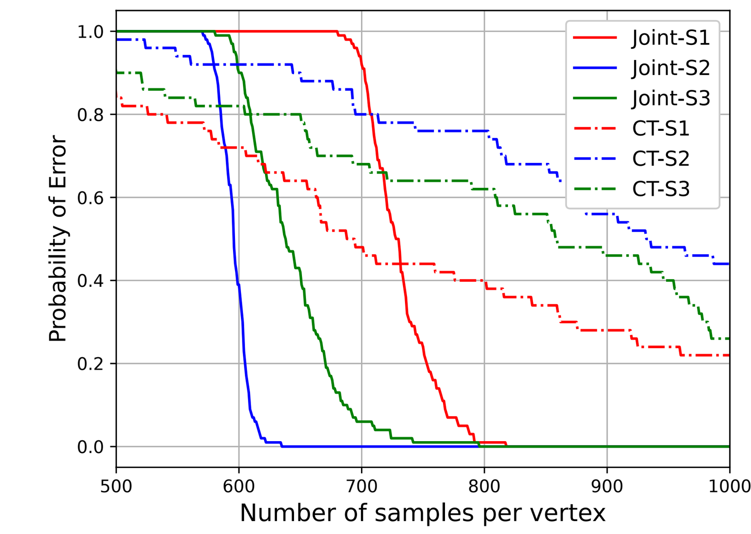
Independent: $O\left(\frac{1}{\lambda^2}\exp(\lambda d) \log \frac{\rho p}{\lambda}\right)$ (Klivans and Meka, FOCS 2017)

Simulations

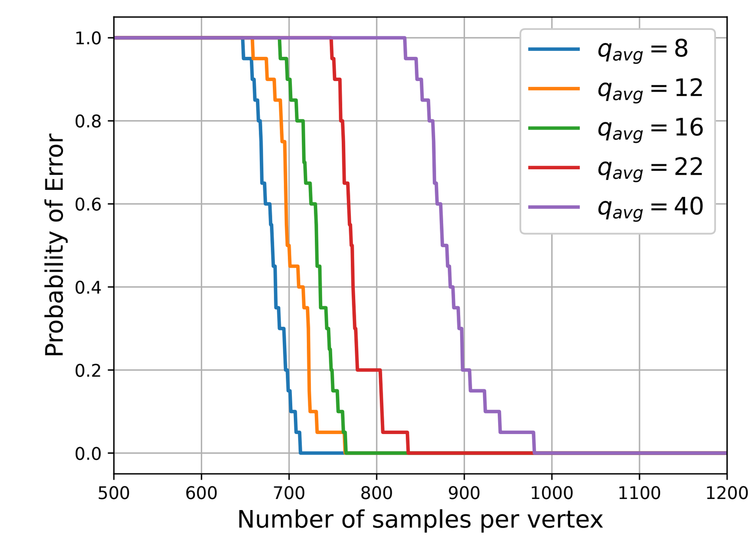
- Generate Erdös-Renyi random graphs with $p = 200$ vertices.
- Baseline: Learn E_1, E_2 separately and form $E_s = E_1 \cap E_2$.
- Comparison with sparse logistic regression (Wu et al. NeurIPS'19).



- Comparison with correlation thresholding (Anandkumar et al. 2010).



- Algorithm can handle various subgraph sizes q.



Conclusions

- Novel problem of learning the shared structure of two graphs.
- An algorithmic framework and its evaluation in different regimes.
- Sample complexity analysis for specific settings.