MetaCaDI: A Meta-Learning Framework for Causal Discovery from Multiple Environments with Unknown Interventions

Hans Jarett Ong[1], **Yoichi Chikahara**[2], **Tomoharu Iwata**[2]

[1]Nara Institute of Science and Technology [2]NTT Communication Science Laboratories

The Challenge: Few-Shot Causal Discovery with Unknown Interventions

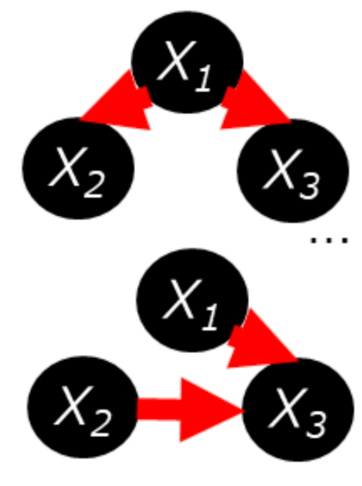
Input: Small Dataset $\mathbf{D}$ (e.g., 3 samples)
Output: Unknown **Causal Graph** and Unknown **Intervention Targets**

This task is very difficult!

Proposed Solution: Use Related Data
Assume each $\mathbf{D}$ is from a **Shared Causal Graph** with **Different Intervention Targets**

Posterior Distribution $\mathbf{P}(\mathbf{A}, \mathbf{m} \mid \mathbf{D})$

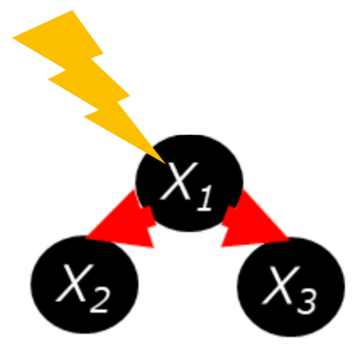
Causal Graph



0.084

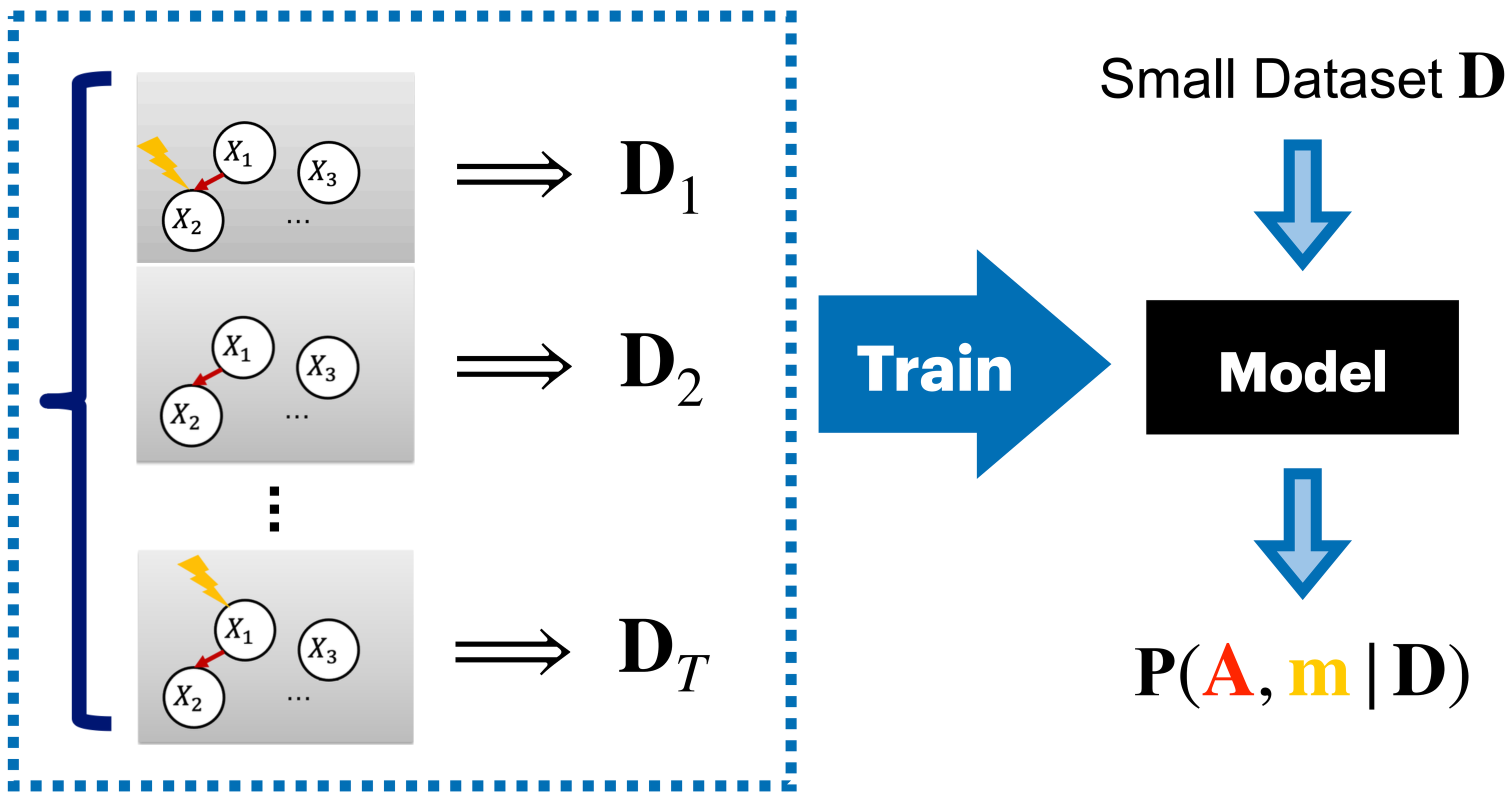
0.054

+



Intervention Targets

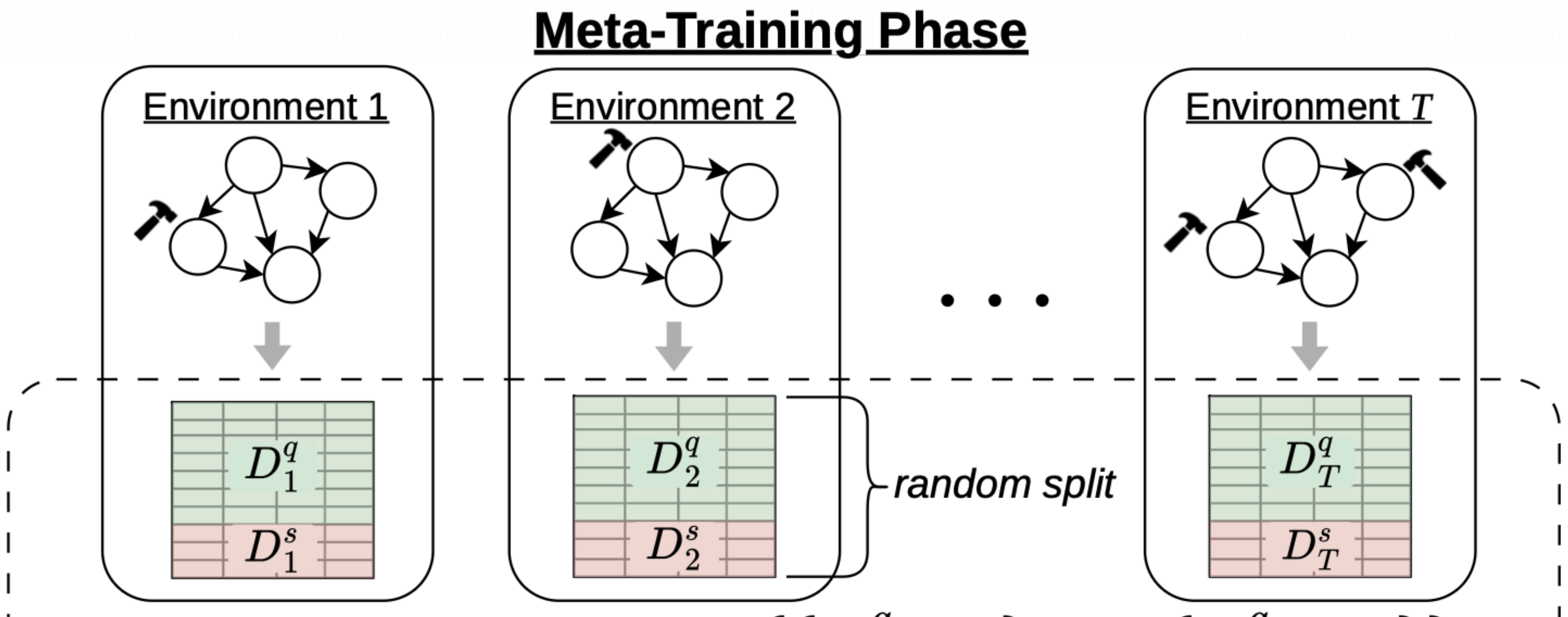
X_1	X_2	X_3
0.7	0.1	0.2



Meta-Learning for *Ca*usal *D*iscovery with Unknown *I*nterventions

Meta-Learning Setup

Meta-Training Phase



Meta-training Data: $\mathcal{D}_{\text{train}} = \{\{D_1^q, D_1^s\}, \ldots, \{D_T^q, D_T^s\}\}$

meta-learn

MetaCaDI Model

Shared Graph Posterior

$\sim \mathrm{P}_{\Upsilon,\Phi}(\mathbf{A} \mid \mathcal{D}_{\text{train}})$

Rapid adaptation mechanism for $D'_{t'}$

Meta-test Phase

Meta-test Data: $\mathcal{D}_{\text{test}} = \{D'_1, D'_2, \ldots, D'_{T'}\}$

D'_1 D'_2 $\cdots$ $D'_{T'}$

$\mathrm{P}_{\Upsilon,\Phi}(\mathbf{m}_1 \mid D'_1)$ $\mathrm{P}_{\Upsilon,\Phi}(\mathbf{m}_2 \mid D'_2)$ $\mathrm{P}_{\Upsilon,\Phi}(\mathbf{m}_{T'} \mid D'_{T'})$

Intervention Target Posteriors

Model Components

Likelihood Model: Masked Auto-encoder

$$X_i = (1 - m_i) \cdot f_i^{\text{obs}}(\mathbf{A}_{:,i} \circ \mathbf{X}) + m_i \cdot f_i^{\text{intv}}(\mathbf{A}_{:,i} \circ \mathbf{X}) + \epsilon_i$$

Differentiable DAG Sampler

Adjacency Matrix $\mathbf{A} = \Pi^{\top} U \Pi$

$U \sim \mathrm{P}_{\phi}(U)$ Upper Triangular Matrix

$\Pi \sim \mathrm{P}_{\psi}(\Pi)$ Permutation Matrix

Intervention Predictor

Sample $m_i \in \{0,1\}$ indicating if the i^{th} variable was intervened on.

$m_i \sim \mathrm{P}(m_i \mid \mathbf{A}, D'_t)$

Task Adaptation: Closed-form solution for the last layer of f_i^{intv}

$$f_i^{\text{intv}}(\mathbf{A}_{:,i} \circ \mathbf{X}) = \mathbf{w}_i^{\top} h(\mathbf{A}_{:,i} \circ \mathbf{X})$$

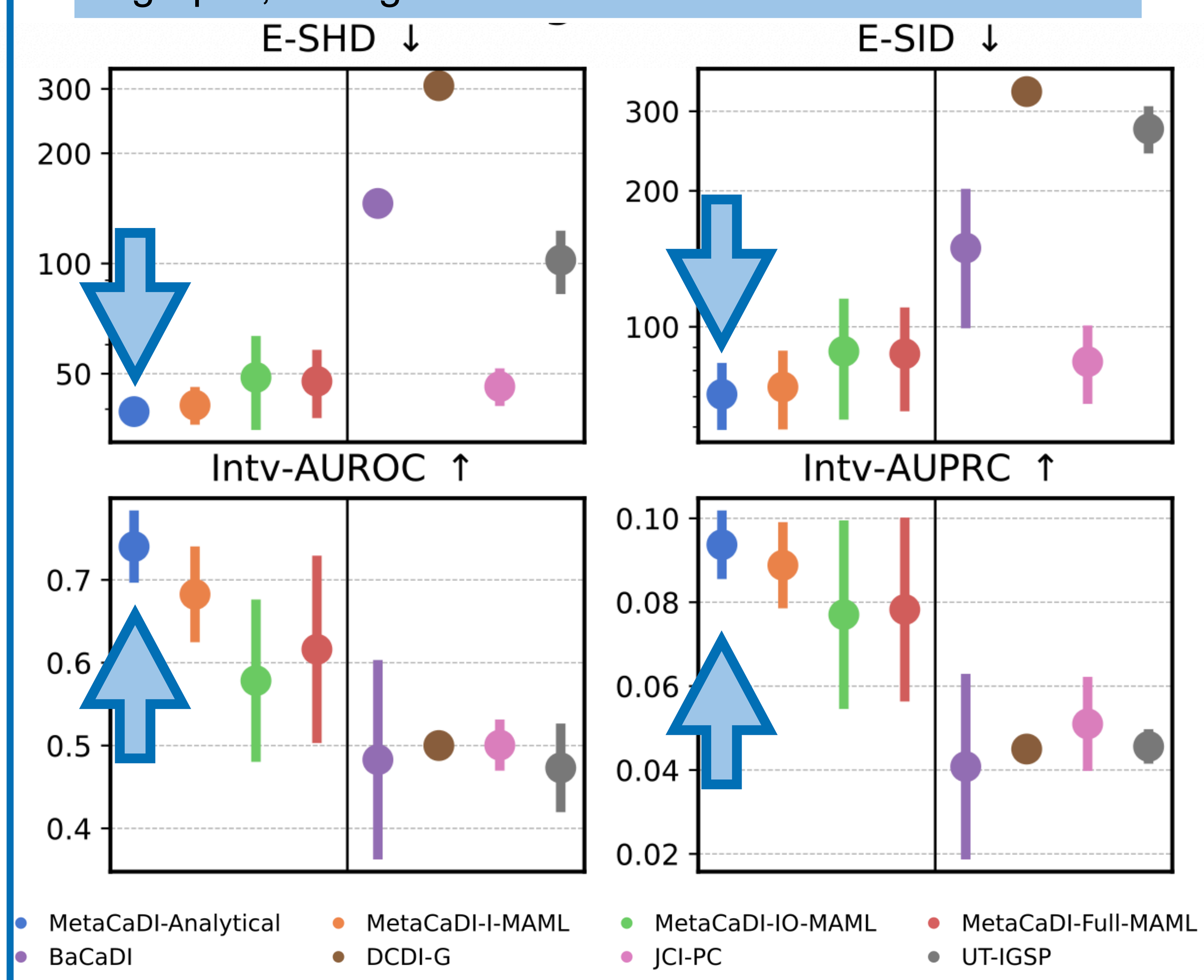
$$\boxed{\hat{\mathbf{w}}_i = (\mathbf{H}_i^{\top}\mathbf{H}_i + \lambda \mathbf{I})^{-1}\mathbf{H}_i^{\top}(D'_t)_{:,i}}$$

Analytical Adaptation (No Gradients!)
This closed-form solution adapts the model to the new task. All the heavy lifting is done during meta-training.

Results

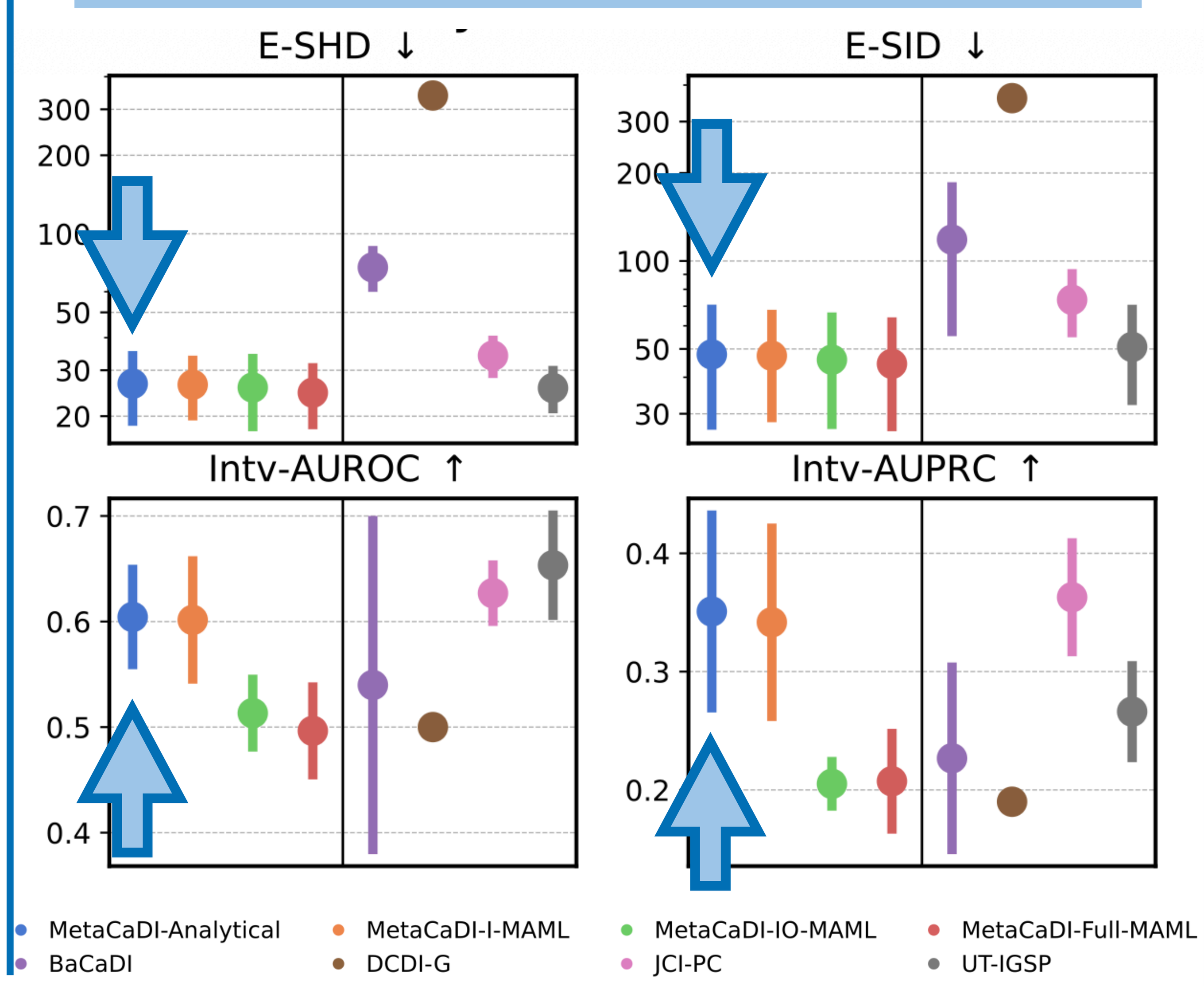
SERGIO Dataset

Realistic gene expression simulator, scale-free graphs, hard gene knockouts.



Synthetic Dataset

Non-linear (Gaussian Process), Erdős-Rényi graphs, soft interventions.



Contact

Hans Jarett Ong | Nara Institute of Science and Technology (NAIST)
Email: ong.hans_jarett.ol5@naist.ac.jp | **Website:** hansjarettong.com

Full Paper →