Extending Homomorphism AE:
On Improving the Disentanglement (Idea 1)

cs18m20p100002

Recap: Loss Functions

$$\rho = \exp \circ \phi$$

$$\mathcal{L}^N_{rec}(\rho, h, d) = \sum_{t=1}^{N+1} \left\| o_t - d\left(\Big(\prod_{i \geq 1}^{t-1} \rho(g_i) \Big) h(o_1) \right) \right\|_2^2$$

$$\mathcal{L}^N_{pred}(\rho, h) = \sum_{t=2}^{N+1} \left\| h(o_t) - \Big(\prod_{i=1}^{t-1} \rho(g_i) \Big) h(o_1) \right\|_2^2$$

Recap: Loss Functions

$$\rho = \exp \circ \phi$$

$$\mathcal{L}_{rec}^{N}(\rho, h, d) = \sum_{t=1}^{N+1} \left\| o_t - d\left(\left(\prod_{i \geq 1}^{t-1} \rho(g_i) \right) h(o_1) \right) \right\|_2^2$$

$$\mathcal{L}_{pred}^{N}(\rho, h) = \sum_{t=2}^{N+1} \left\| h(o_t) - \left(\prod_{i=1}^{t-1} \rho(g_i) \right) h(o_1) \right\|_2^2$$

For Disentanglement:

$$\mathcal{L}_{sparse}(\rho) = \sum_t \sum_{i \geq 0} \sqrt{\sum_{j \geq i+1,\, k \leq i} \rho_{kj}(g_t)^2 + \rho_{jk}(g_t)^2}$$

$$\phi(\varphi(g)) = \begin{pmatrix} \phi_1(\varphi(g)) & 0 & \cdots & 0 \\ 0 & \phi_1(\varphi(g)) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \phi_1(\varphi(g)) \end{pmatrix} \quad (5)$$

While we do not prove that this block diagonal constraint leads to disentanglement, we show through experiments that

Proposal for Disentanglement

$$\rho = \exp \circ \phi$$

$$\mathcal{L}_{rec}^{N}(\rho, h, d) = \sum_{t=1}^{N+1} \left\| o_t - d\left(\left(\prod_{i \geq 1}^{t-1} \rho(g_i)\right) h(o_1)\right) \right\|_2^2 \qquad \mathcal{L}_{pred}^{N}(\rho, h) = \sum_{t=2}^{N+1} \left\| h(o_t) - \left(\prod_{i=1}^{t-1} \rho(g_i)\right) h(o_1) \right\|_2^2$$

- <u>If</u> we show that $\mathcal{L}_{rec}^{N}(\rho, h, d) + \mathcal{L}_{pred}^{N}(\rho, h)$ is strongly convex in terms of ϕ.
- <u>If</u> exp preserves the strong convexity.

Proposal for Disentanglement

$$\rho = \exp \circ \phi$$

$$\mathcal{L}_{rec}^{N}(\rho, h, d) = \sum_{t=1}^{N+1} \left\| o_t - d\left(\Big(\prod_{i \geq 1}^{t-1} \rho(g_i) \Big) h(o_1) \right) \right\|_2^2 \qquad \mathcal{L}_{pred}^{N}(\rho, h) = \sum_{t=2}^{N+1} \left\| h(o_t) - \Big(\prod_{i=1}^{t-1} \rho(g_i) \Big) h(o_1) \right\|_2^2$$

- If we show that $\mathcal{L}_{rec}^{N}(\rho, h, d) + \mathcal{L}_{pred}^{N}(\rho, h)$ is strongly convex in terms of ϕ.
- If exp preserves the strong convexity.
- Then we will get an equivalent weakly submodular optimization:

Restricted Strong Convexity
Implies Weak Submodularity*

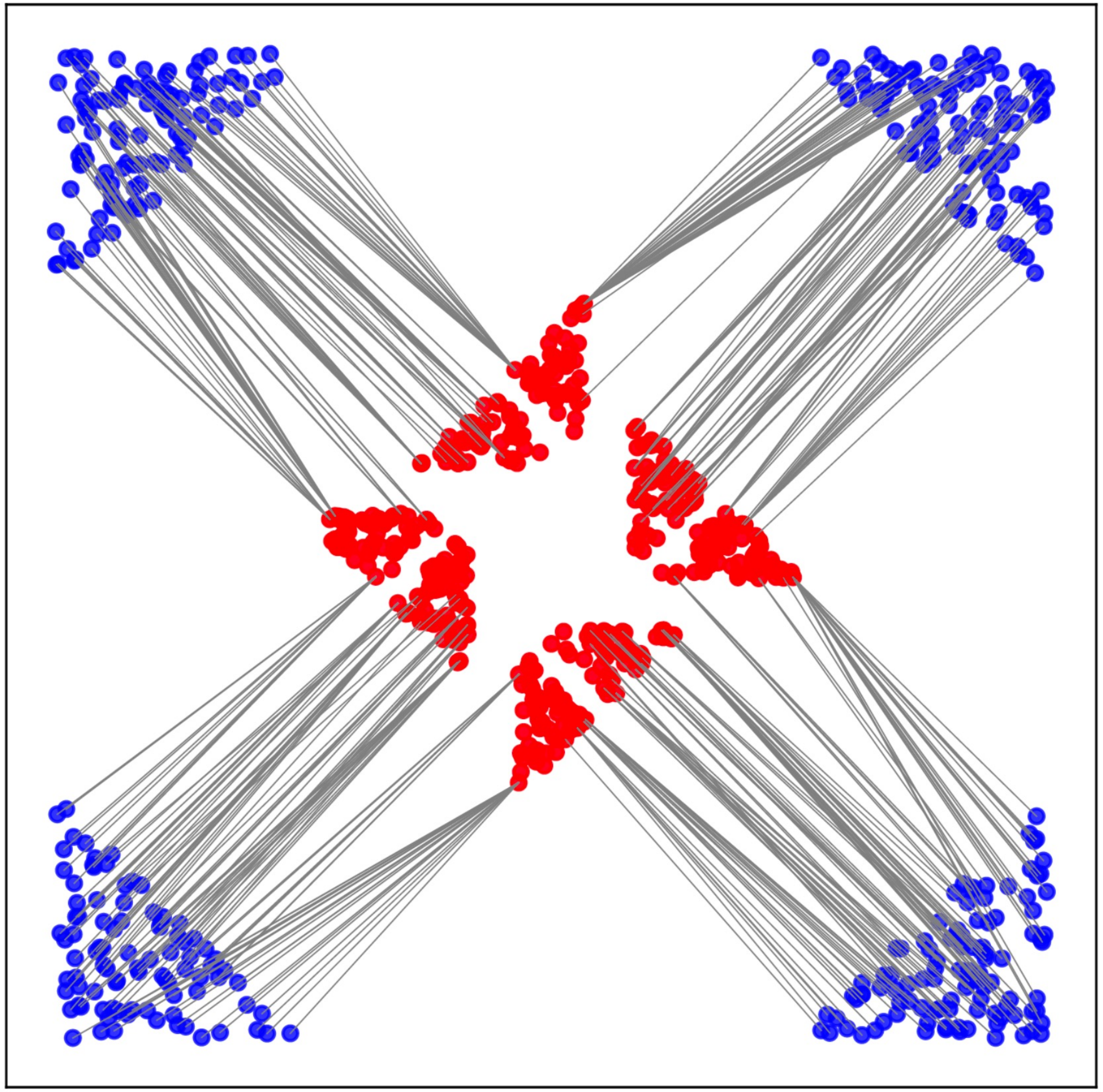
Ethan R. Elenberg[1], Rajiv Khanna[1], Alexandros G. Dimakis[1],
and Sahand Negahban[2]

- Greedy algos can enforce the block-diagonal constraint with a constant-factor approximation guarantee for its optimality.

Extending Homomorphism AE: Applications to Latent OT (Idea 2)

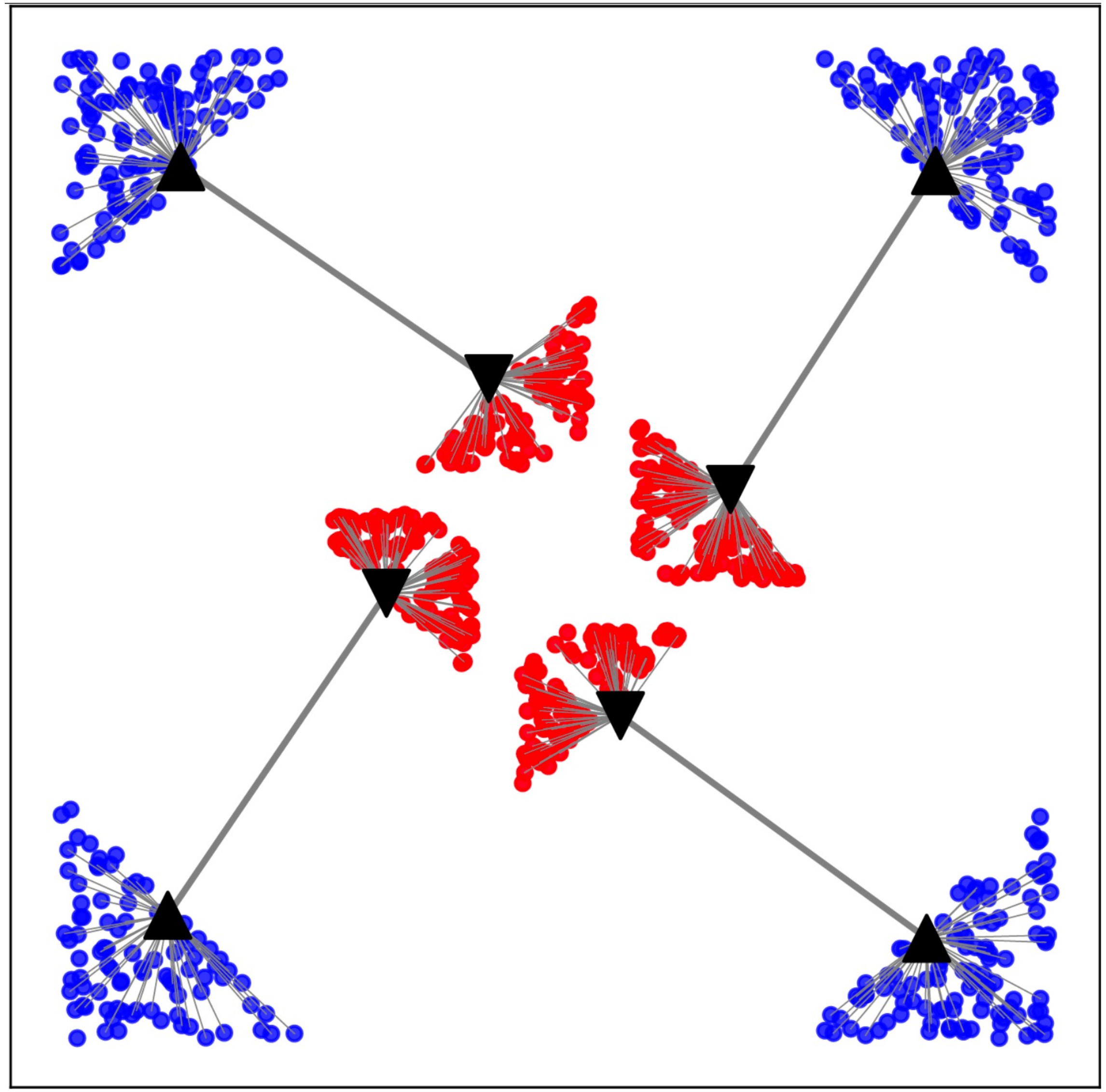
Brief Overview: Latent OT

OT alignments b/w source & target :

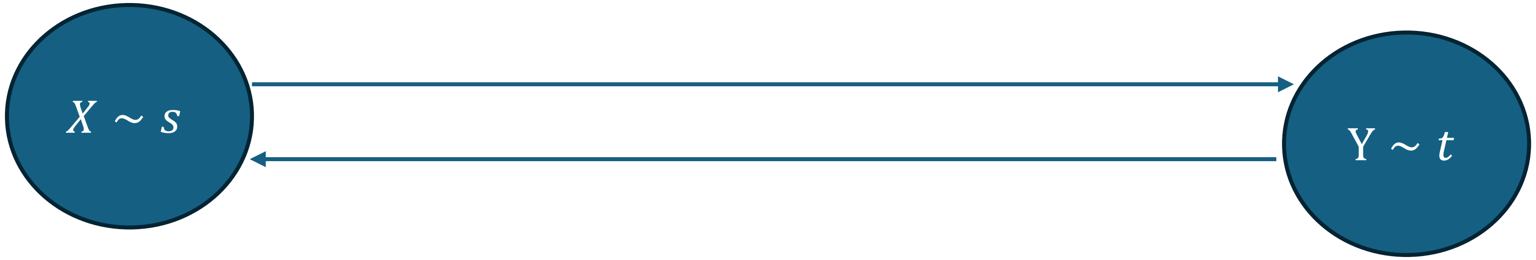


Brief Overview: Latent OT

LOT (ICML'21) alignments b/w source & target :

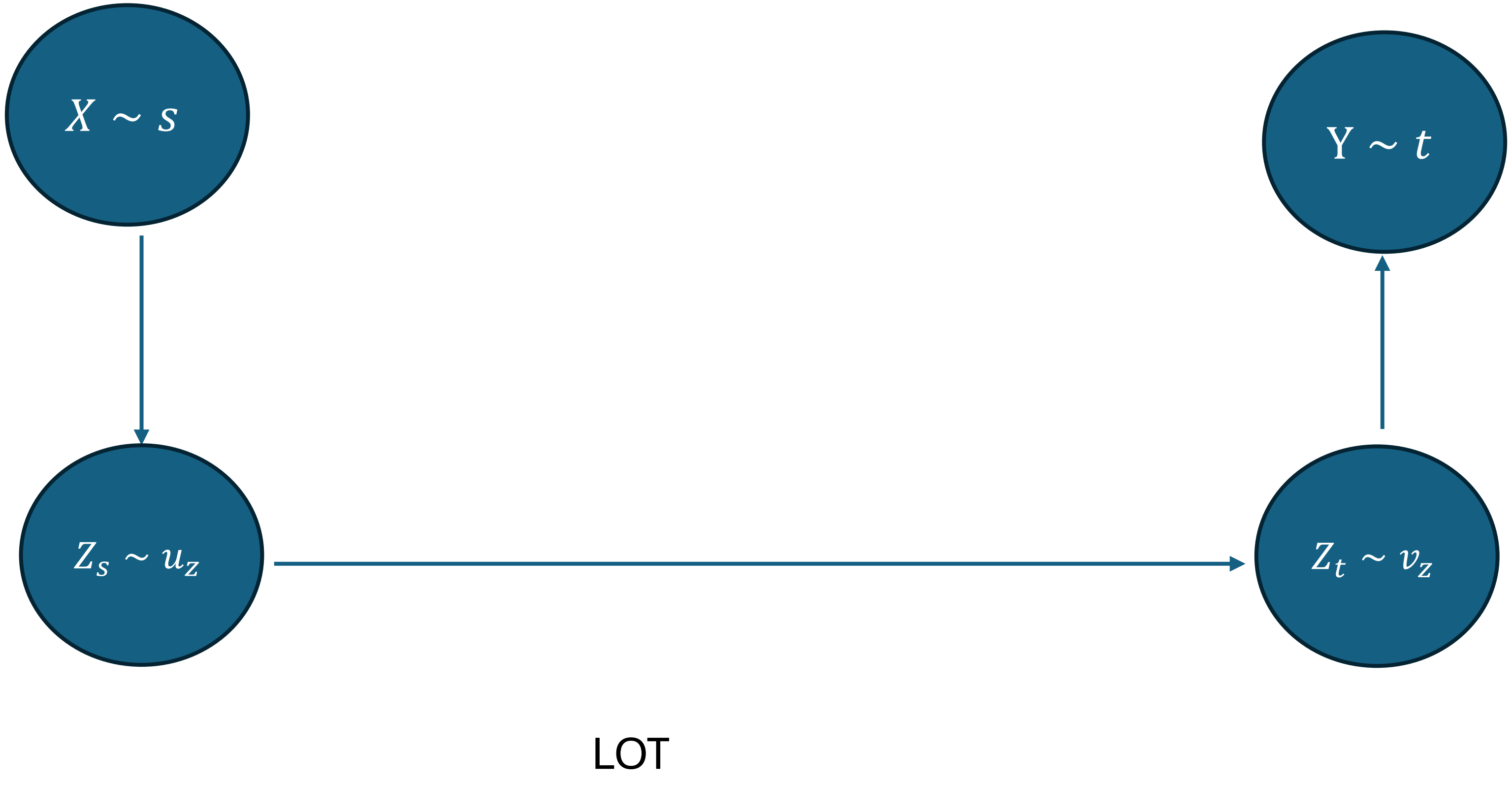


LOT & HAE (High-Level Connection)

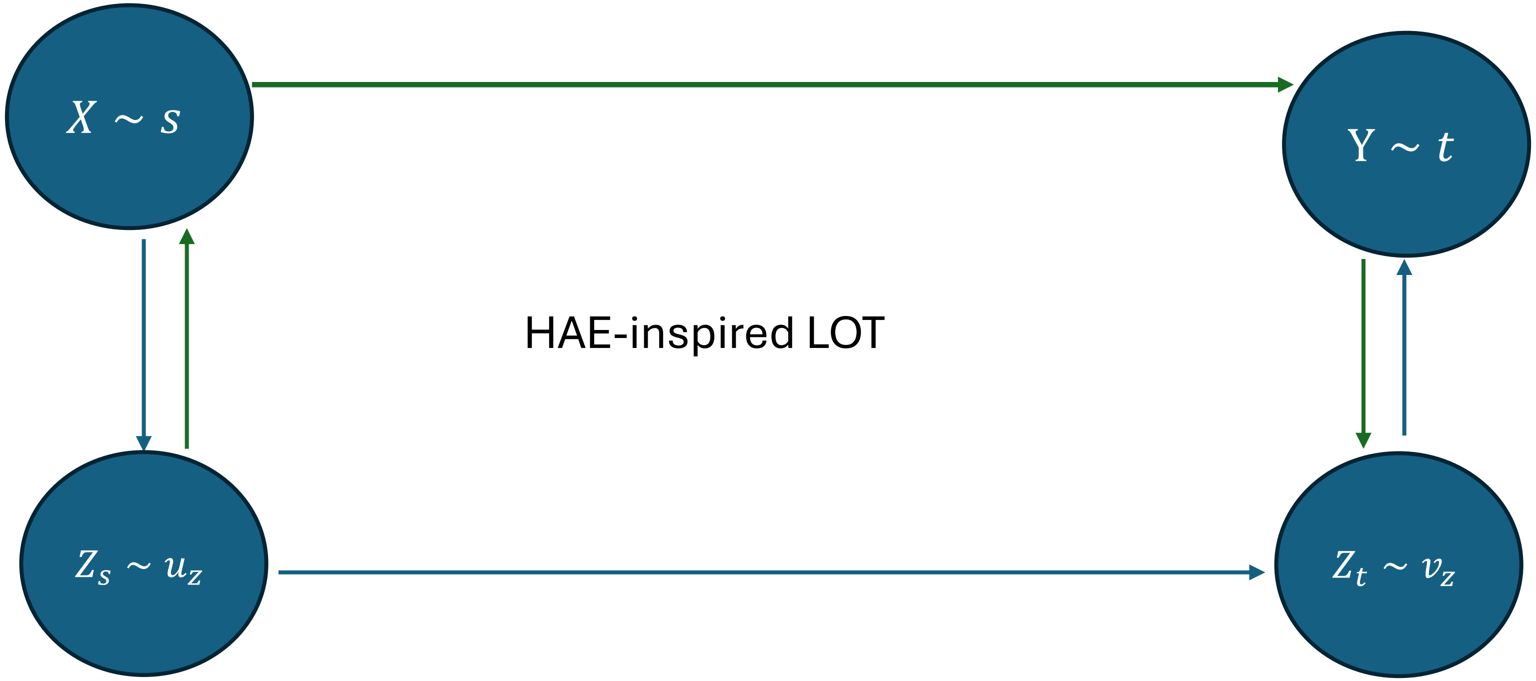


Usual OT

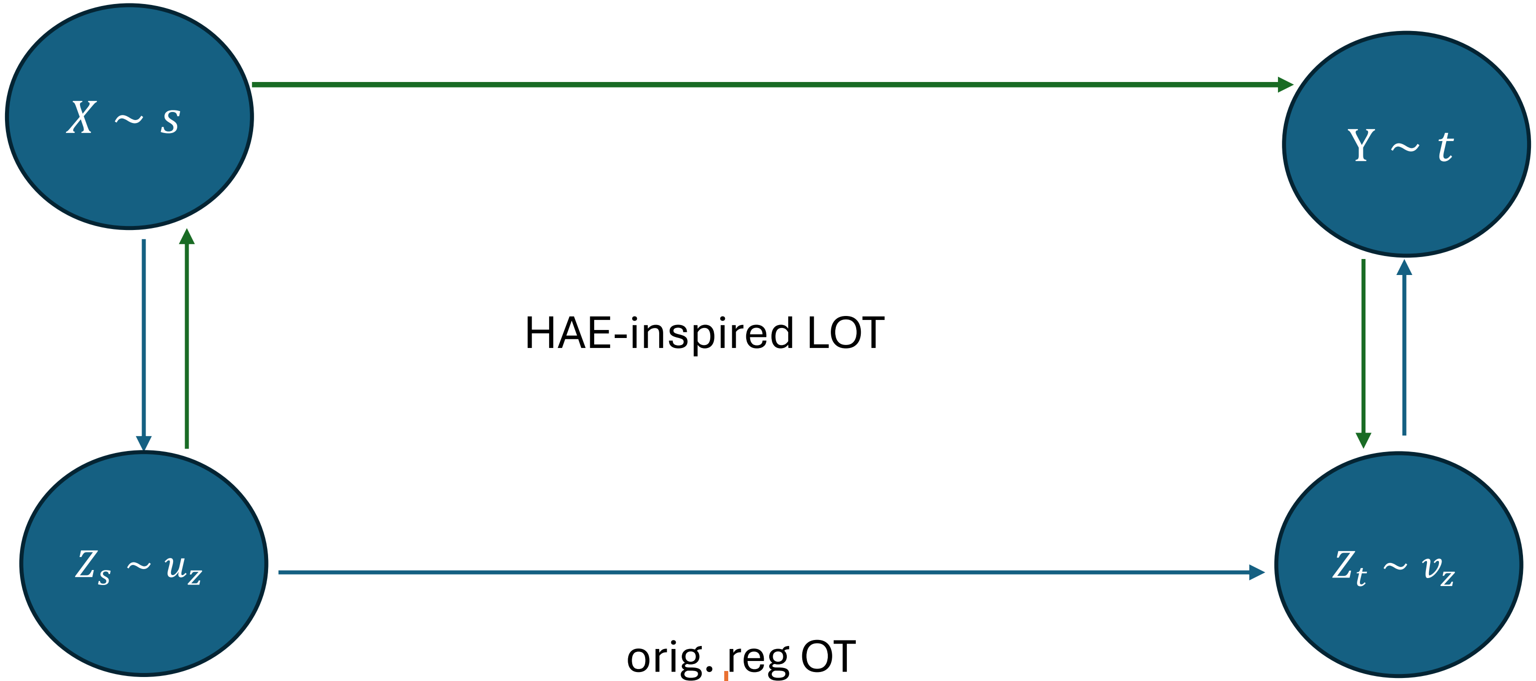
LOT & HAE (High-Level Connection)



LOT & HAE (High-Level Connection)



LOT & HAE (High-Level Connection)



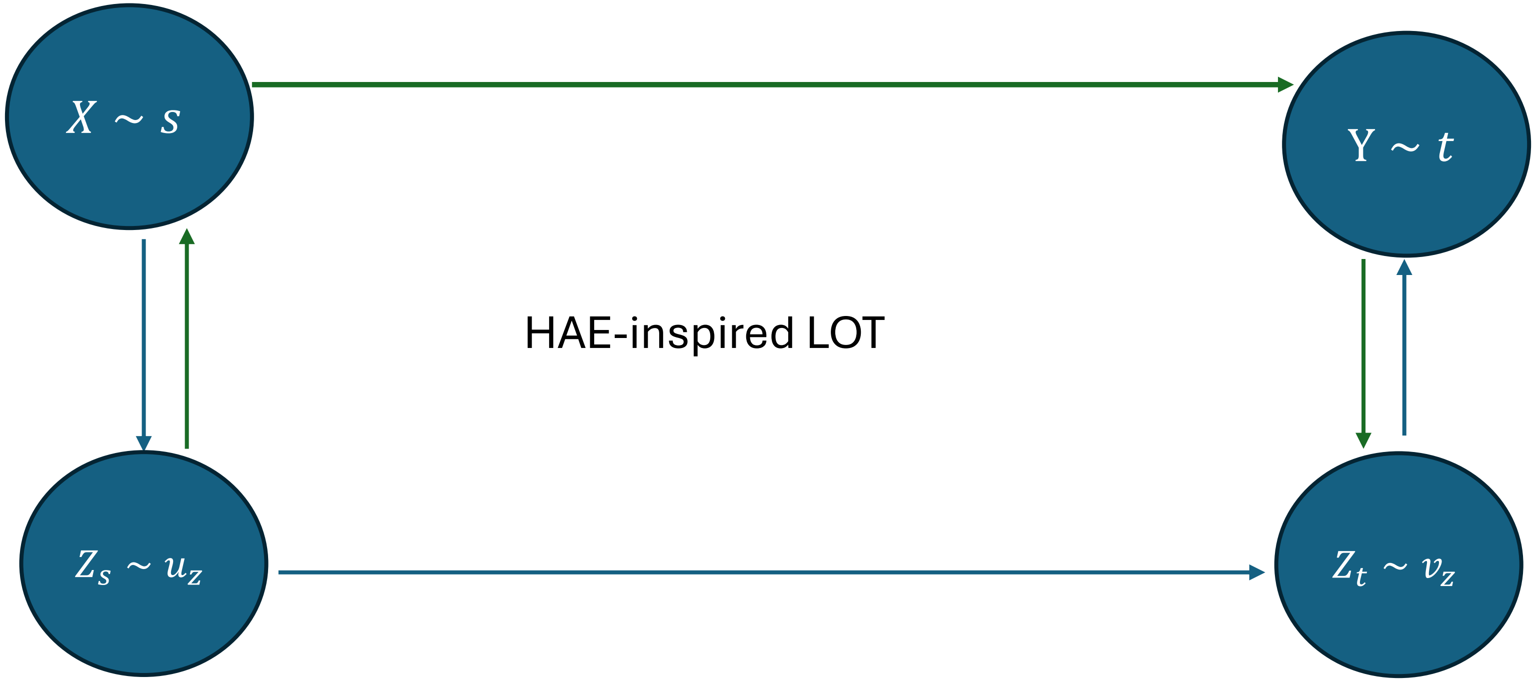
orig. reg OT

Why should disentanglement help LOT?

$$
\begin{aligned}
\varepsilon KL(\mathbf{P}\|\mathbf{K}) \leq & \varepsilon(KL(\mathbf{P}_x\|\mathbf{K}_x) + KL(\mathbf{P}_z\|\mathbf{K}_z) \\
& + KL(\mathbf{P}_y\|\mathbf{K}_y)) + \varepsilon(\mathbf{H}(\mathbf{u}_z) + \mathbf{H}(\mathbf{v}_z)),
\end{aligned}
$$

where $\mathbf{H}(\mathbf{a}) := -\sum_i \mathbf{a}_i \log \mathbf{a}_i$ *denotes the entropy.*

LOT & HAE (High-Level Connection)



Why should disentanglement help LOT?

$$\varepsilon KL(\mathbf{P}\|\mathbf{K}) \leq \varepsilon(KL(\mathbf{P}_x\|\mathbf{K}_x) + KL(\mathbf{P}_z\|\mathbf{K}_z) + KL(\mathbf{P}_y\|\mathbf{K}_y)) + \varepsilon(\mathbf{H}(\mathbf{u}_z) + \mathbf{H}(\mathbf{v}_z)), \quad \leq \mathrm{I}(Z_s; Z_t)$$

where $\mathbf{H}(\mathbf{a}) := -\sum_i \mathbf{a}_i \log \mathbf{a}_i$ *denotes the entropy.*