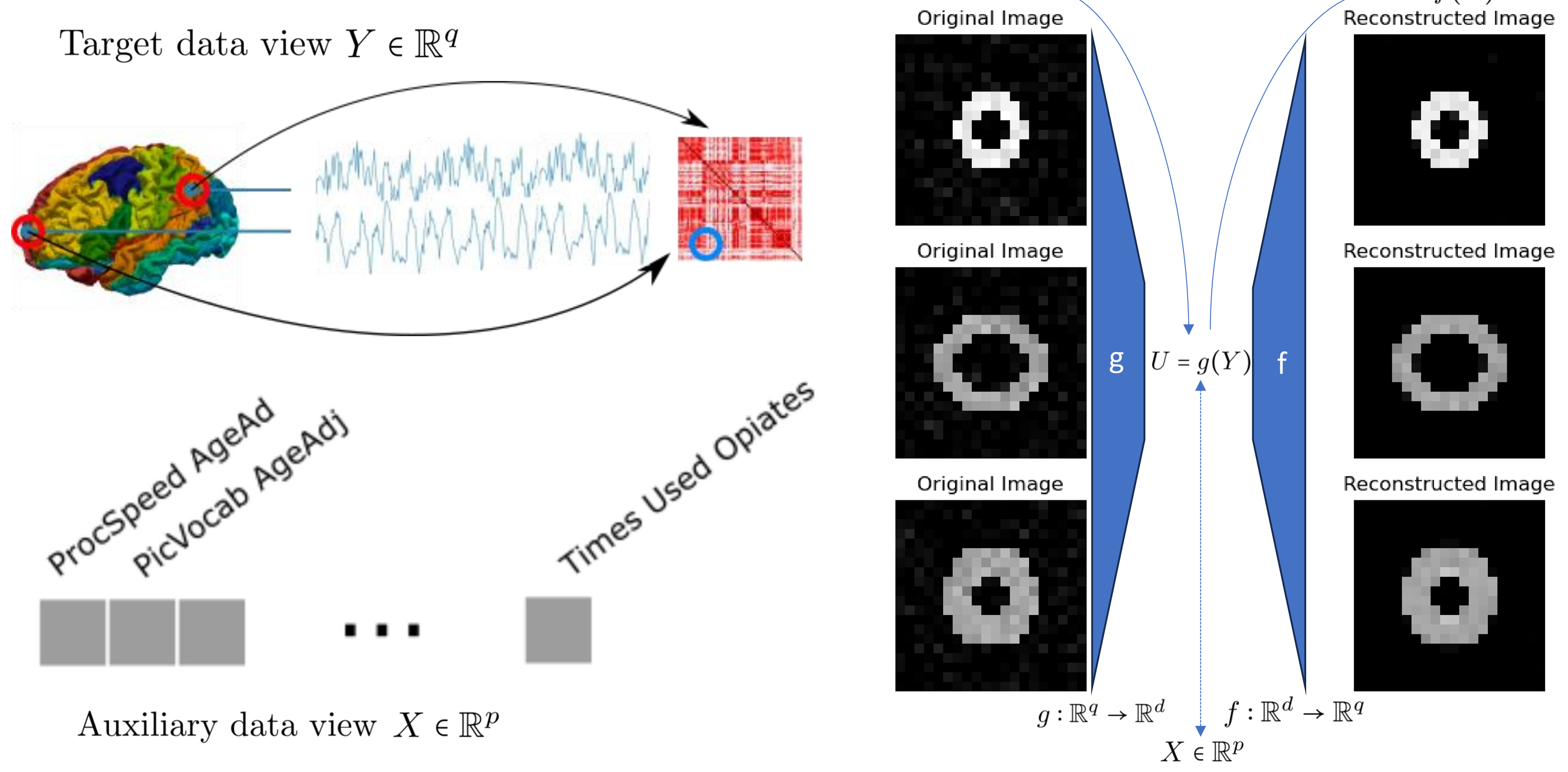
Supervised Disentanglement via Canonical Correlation Analysis

James Buenfil[1], Eardi Lila[2]

[1]Department of Statistics, University of Washington
[2]Department of Biostatistics, University of Washington

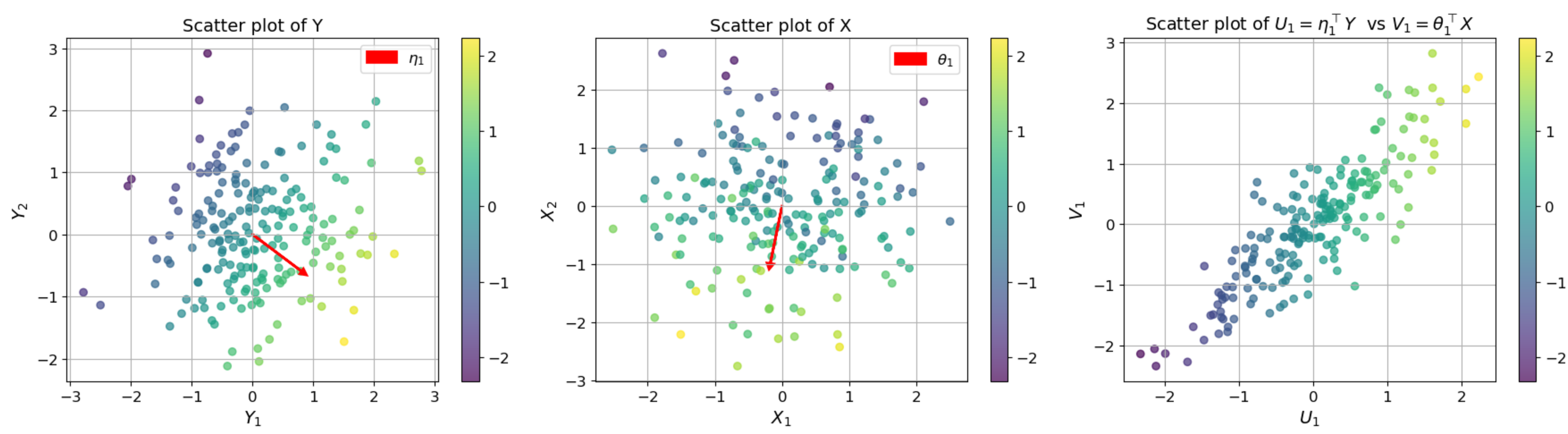
Supervised Disentanglement

We observe i.i.d. pairs (X_i, Y_i) for $i = 1, \ldots N$



Canonical Correlation Analysis (CCA)

$$(\eta, \theta) = \underset{\eta \in \mathbb{R}^q, \theta \in \mathbb{R}^p}{\arg\max} \operatorname{Corr}(\eta^\top Y, \theta^\top X) \quad \text{s.t.} \quad \operatorname{Var}(\eta^\top Y) = \operatorname{Var}(\theta^\top X) = 1.$$

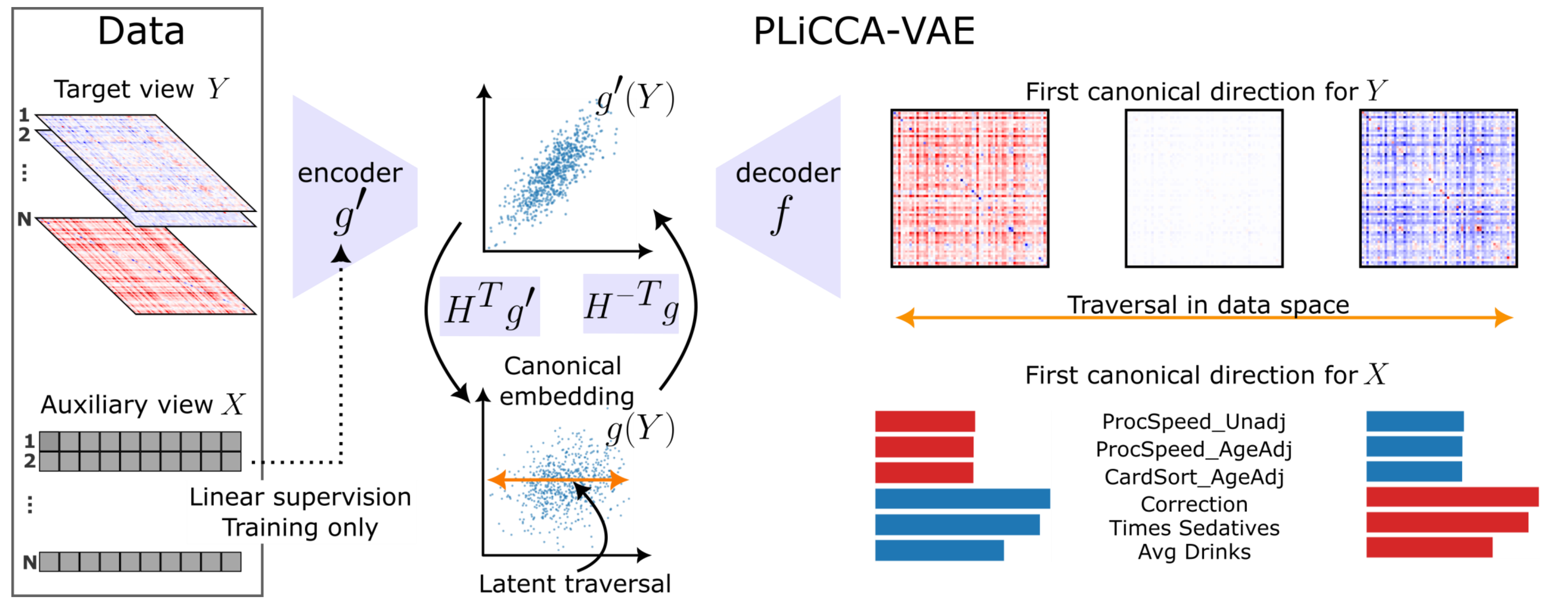


Estimating *multiple* canonical vectors:

$$\max_{\{\eta_i, \theta_i\}_{i=1}^d} \sum_{i=1}^d \operatorname{Corr}(\eta_i^\top Y, \theta_i^\top X) \quad \text{s.t.} \quad \operatorname{Corr}(\eta_i^\top Y, \eta_j^\top Y) = \operatorname{Corr}(\theta_i^\top X, \theta_j^\top X) = \delta_{ij}.$$

Canonical vectors, $H \equiv [\eta_1 \ldots \eta_d] \in \mathbb{R}^{q \times d}$ and $T \equiv [\theta_1, \ldots \theta_d] \in \mathbb{R}^{p \times d}$.

$$\max_{H \in \mathbb{R}^{q \times d}, T \in \mathbb{R}^{p \times d}} \mathbb{E}\left[(H^\top Y)^\top (T^\top X)\right] \quad \text{s.t.} \quad \Sigma_{H^\top Y} = \Sigma_{T^\top X} = I_d.$$



Partially Linear invertible CCA (PLiCCA)

Replace $H^\top Y$ with $g(Y)$, where $g: \mathbb{R}^q \to \mathbb{R}^d$:

$$\max_{\substack{g: \mathbb{R}^q \to \mathbb{R}^d, T \in \mathbb{R}^{p \times d} \\ g \in \mathcal{C}}} \mathbb{E}\left[g(Y)^\top (T^\top X)\right] \quad \text{s.t.} \quad \Sigma_{g(Y)} = \Sigma_{T^\top X} = I_d.$$

$$T = [\theta_1, \ldots, \theta_d]$$
$$g(Y) = (g_1(Y), \ldots, g_d(Y))^\top$$
$$\mathcal{C} \subseteq \{g: \mathbb{R}^q \to \mathbb{R}^d : \exists f: \mathbb{R}^d \to \mathbb{R}^q \text{ s.t. } \mathbb{E}\left[\|Y - f(g(Y))\|_2^2\right] < \varepsilon\}$$

Theorem (Regression formulation of PLiCCA)

Finding (g, T) that solves

$$\underset{\substack{g: \mathbb{R}^q \to \mathbb{R}^d,\ T \in \mathbb{R}^{p \times d} \\ \Sigma_{g(Y)} = \Sigma_{T^\top X} = I_d,\ g \in a^{-1/2}\mathcal{C}}}{\text{maximize}} \sum_{i=1}^d \mathbb{E}\left[g_i(Y)\, \theta_i^\top X\right]^2$$

is equivalent to finding (g', B) that solves

$$\underset{\substack{g' \in \mathcal{C},\ B \in \mathbb{R}^{p \times d} \\ \Sigma_{g'(Y)} \succeq a I_d}}{\text{minimize}} \mathbb{E}\left[\|g'(Y) - B^\top X\|_2^2\right].$$

Methodology - Conditional VAE

Introduce a latent $Z \in \mathbb{R}^d$ and specify:

$$Y \mid Z \sim \mathcal{N}(f(Z), I_d)$$
$$Z \mid X \sim \mathcal{N}(B^\top X, I_d)$$
$$q(z \mid y) = \mathcal{N}(g(y), I_d)$$

Determines the joint distribution $p(Y, Z \mid X)$.

Evidence Lower Bound (ELBO)

Rather than maximize the likelihood $p(Y \mid X)$ directly, more tractable to minimize

$$\min_{g, f, B} \underbrace{\mathbb{E}\left[\|g(Y) - B^\top X\|_2^2\right]}_{\text{regression}} + \beta_{\text{VAE}} \underbrace{\mathbb{E}\left[\mathbb{E}_{q(z|Y)}\left[\|Y - f(z)\|_2^2\right]\right]}_{\text{reconstruction}}$$

Results – Human Connectome Project Dataset

Functional connectivity		
Method	Total correlation	Reconstruction error
DCCA	1.6879 (0.0971)	–
DCCA-NOI	0.1933 (0.0855)	–
DCCA-SDL	1.8628 (0.0396)	–
DCCAE	0.9466 (0.0879)	0.000251 (0.000009)
DVCCA	0.9912 (0.1294)	0.015607 (0.008683)
PLiCCA-VAE	1.1440 (0.1822)	0.000538 (0.000116)
PLiCCA-NF	1.0753 (0.2969)	0.000458 (0.000008)

Cortical thickness		
Method	Total correlation	Reconstruction error
DCCA	0.5500 (0.0951)	–
DCCA-NOI	0.5793 (0.1560)	–
DCCA-SDL	0.6149 (0.0902)	–
DCCAE	0.8319 (0.0296)	0.2038 (0.0104)
DVCCA	0.8823 (0.0538)	0.8028 (0.1475)
PLiCCA-VAE	1.1195 (0.1507)	0.1489 (0.0013)
PLiCCA-NF	0.9536 (0.1329)	0.1753 (0.0508)

PLiCCA

$$\min_{g, f, B} \mathbb{E}\left[\|g(Y) - B^\top X\|_2^2\right] + \beta\, \mathbb{E}\left[\|Y - f(g(Y))\|_2^2\right]$$

s.t. $\Sigma_{g(Y)} \succeq a I_d$

Conditional VAE - Proxy for PLiCCA

$$\min_{g, f, B} \mathbb{E}\left[\|g(Y) - B^\top X\|_2^2\right] + \beta_{\text{VAE}}\, \mathbb{E}\left[\mathbb{E}_\varepsilon\left[\|Y - f(g(Y) + \varepsilon)\|_2^2\right]\right]$$

with $\varepsilon \sim \mathcal{N}(0, I_d)$

Theorem

Fix positive constants δ and σ^2_{enc}. Suppose g and f are such that the reconstruction error $\mathbb{E}\left[\mathbb{E}_{q(z|Y)}\left[\|Y - f(z)\|_2^2\right]\right] < \delta$, and suppose that we model $q(z|y) \sim \mathcal{N}\left(g(y), \sigma^2_{\text{enc}} I_d\right)$. Then,

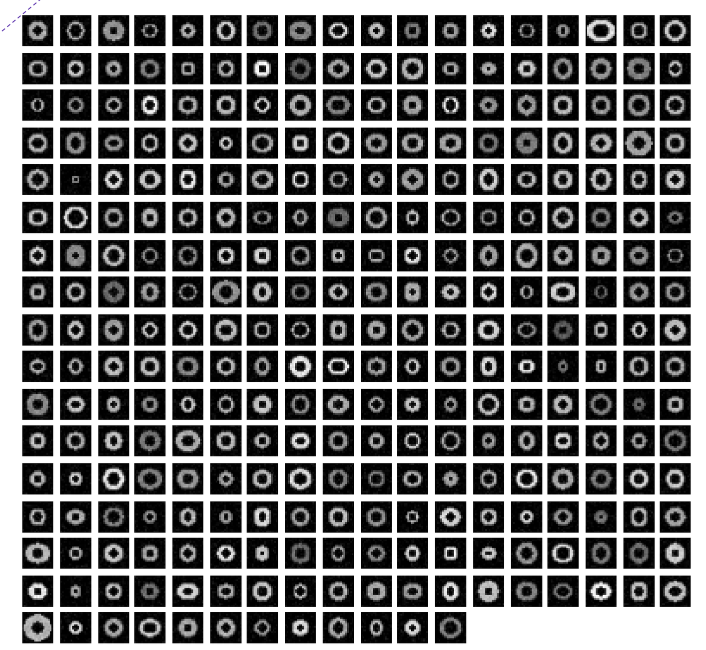
$$\operatorname{tr}\left(\Sigma_{g(Y)}\right) \geq \sigma^2_{\text{enc}} C(\delta), \tag{3}$$

where $C(\delta)$ is non-increasing in δ.

Takeaway: the proxy problem relaxes the lower bound on $\Sigma_{g(Y)}$ to a lower bound on $\operatorname{tr}\left(\Sigma_{g(Y)}\right)$.

Results

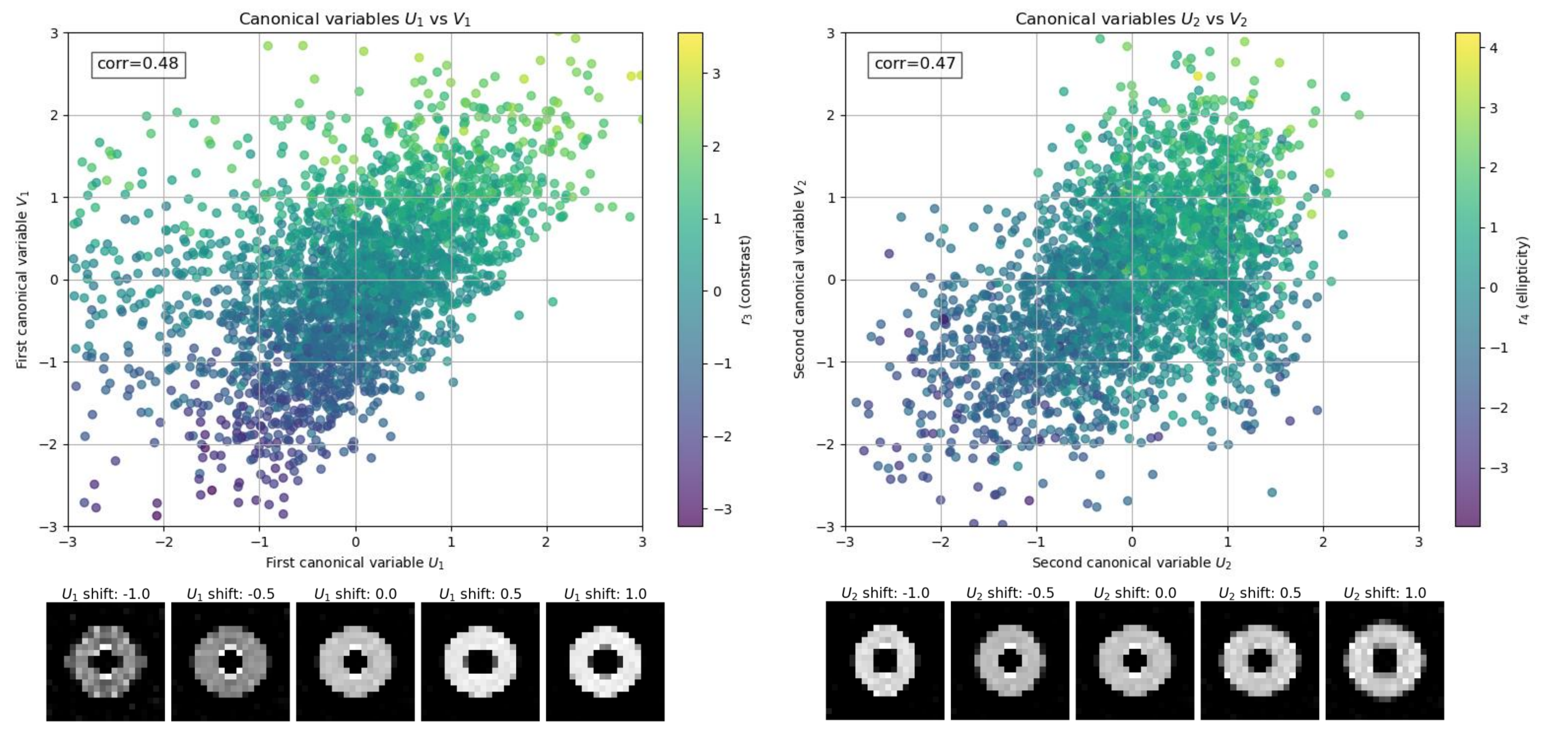
Results – Synthetic Ring Dataset



Images are 20×20 pixels: $Y \in \mathbb{R}^{400}$.

The rings are parameterized by four parameters:
r_1 = "radius of the hole,"
r_2 = "width of the ring,"
r_3 = "contrast," which scales the images by a constant, effectively decreasing or increasing the contrast of the image, and
r_4 = "ellipticity," compressing the ring along the x-axis.

$X \in \mathbb{R}^{30}$ contains Gaussian noise except through X_3 and X_4 which are correlated with r_3 and r_4, with correlations $\gamma_1 = 0.9$ and $\gamma_2 = 0.7$ respectively.



PLiCCA has successfully learned that the contrast of the image r_3 is correlated with X_3, and that the ellipticity of the ring r_4 is correlated with X_4.

Method	Validation correlation	Validation reconstruction error
DCCA	0.680 (0.033)	–
DCCA-NOI	0.680 (0.037)	–
DCCA-SDL	0.350 (0.031)	–
DCCAE	0.040 (0.009)	0.02297 (0.00047)
DVCCA	0.151 (0.021)	0.10463 (0.00652)
PLiCCA-VAE	0.959 (0.024)	0.00796 (0.00017)
PLiCCA-NF	0.898 (0.033)	0.00794 (0.00012)