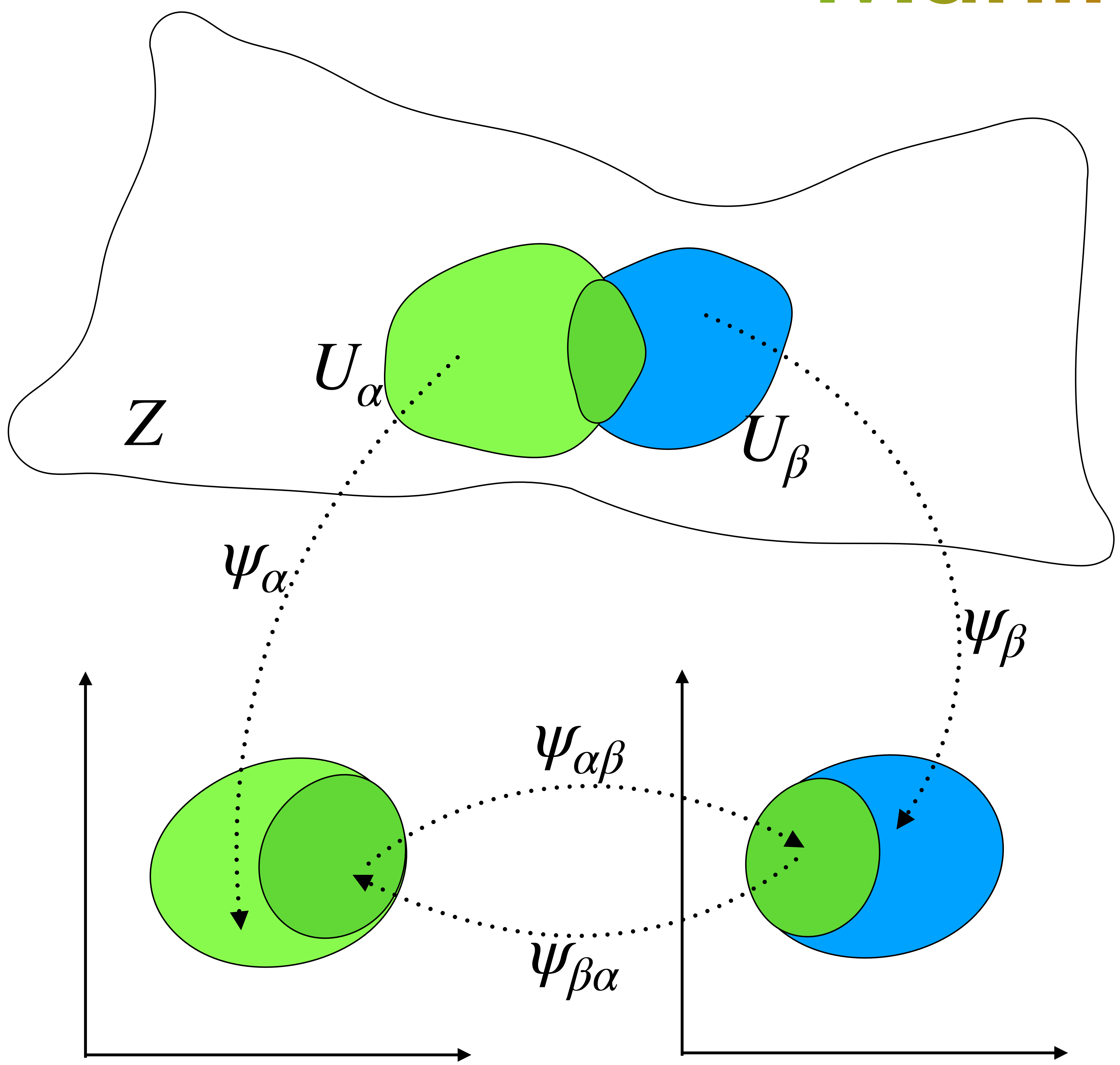
# State Representation Learning Using an Unbalanced Atlas

Li Meng, Morten Goodwin, Anis Yazidi, Paal Engelstad

# State Representation Learning (SRL)

- Reinforcement Learning (RL).
- Self-supervised learning (SSL).
- Learn representations from (unlabelled) data collected in RL.

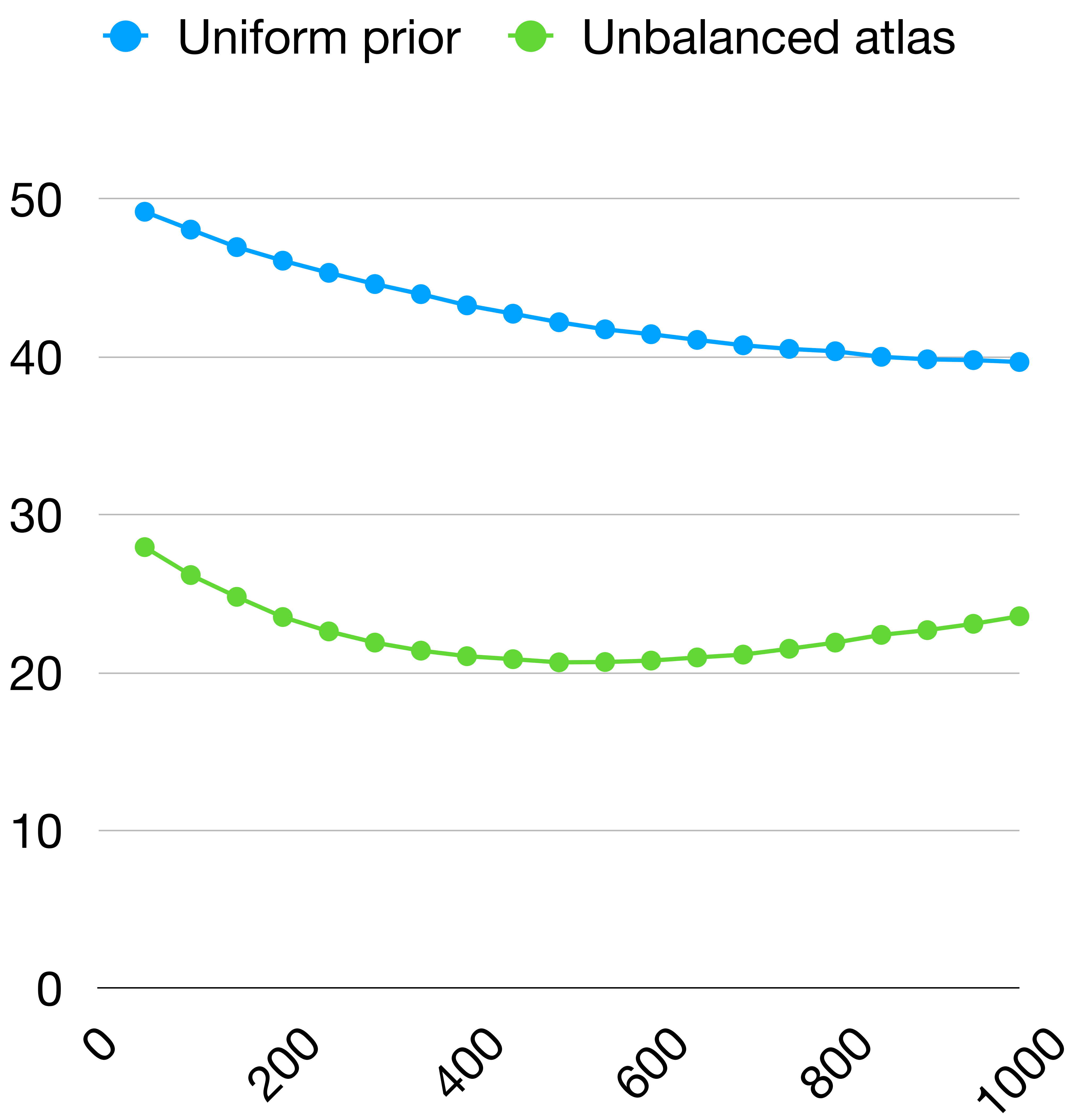
# Manifold




A manifold $Z$ embedded in a higher dimension. Two domains are denoted by $U_\alpha$ and $U_\beta$ in $Z$. $\psi_\alpha$ and $\psi_\beta$ are the corresponding charts that map them to a lower dimensional Euclidean space.

A manifold can be learned by finding an atlas that accurately describes the local structure in each chart.

# Unbalanced Atlas (UA)




- An atlas is a collection of the charts that together cover the entire manifold.
- Membership probability distribution is deliberately trained to deviate significantly from uniformity.

# Unbalanced Atlas (UA)

- $\text{Output}(x) = \sum q_i(f(x))\mathcal{I}(\psi_i(f(x)))$
- $\text{Output}(x) = \mathcal{I}(\psi_i(f(x)))$, where $i = \operatorname{argmax}_j q_j(f(x))$ at inference time.

# Deep InfoMax with UA

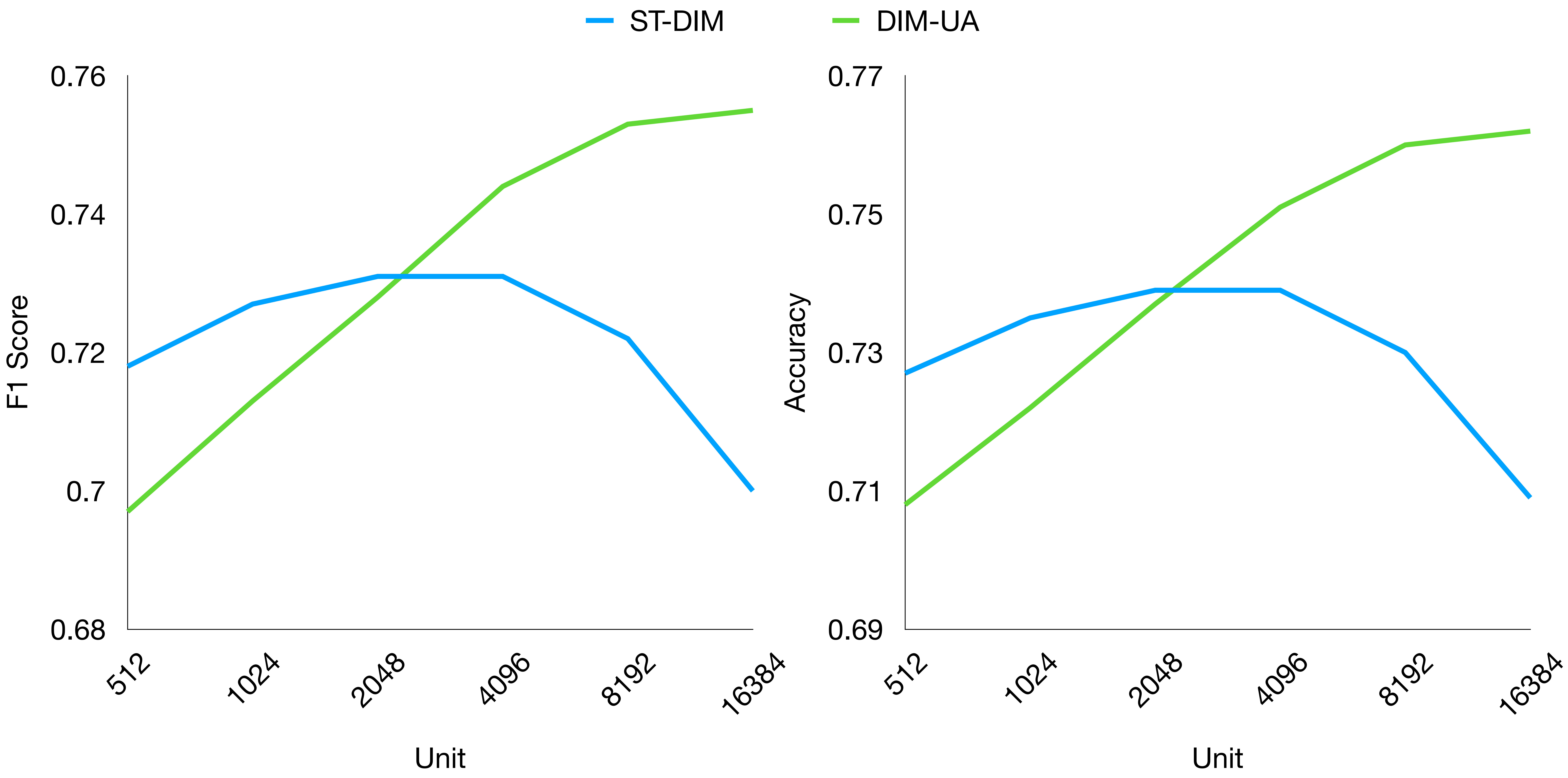
- $$\mathcal{L}_{GL} = \sum_{m=1}^{M} \sum_{n=1}^{N} -\log \frac{\exp(g_{m,n}(x_t, x_{t+1}))}{\sum_{x_{t*} \in X_{next}} \exp(g_{m,n}(x_t, x_{t*}))}$$
- $$\mathcal{L}_{LL} = \sum_{m=1}^{M} \sum_{n=1}^{N} -\log \frac{\exp(h_{m,n}(x_t, x_{t+1}))}{\sum_{x_{t*} \in X_{next}} \exp(h_{m,n}(x_t, x_{t*}))}$$
- $$\mathcal{L}_{Q} = -\frac{1}{2} \sum_{i=1}^{n} ((q_i(f(x_t)) - \frac{1}{n})^2 + (q_i(f(x_{t+1})) - \frac{1}{n})^2)$$

# Experiments

- 19 games of the AtariARI benchmark.
- 5 categories of state variables: agent , small object, other localizations, miscellaneous, and score/clock/lives/display.
- Evaluate the probe accuracy and F1 scores on the downstream linear probing tasks.

# Probe F1 scores of each game averaged across categories

| Game | VAE | CPC | ST-DIM | ST-DIM* | DIM-UA |
|---|---|---|---|---|---|
| Asteroids | 0.36 | 0.42 | 0.49 | 0.48 ± 0.005 | **0.5** ± 0.007 |
| Bowling | 0.50 | 0.90 | **0.96** | **0.96** ± 0.021 | **0.96** ± 0.018 |
| Boxing | 0.20 | 0.29 | 0.58 | 0.61 ± 0.008 | **0.64** ± 0.007 |
| Breakout | 0.57 | 0.74 | 0.88 | 0.88 ± 0.02 | **0.9** ± 0.016 |
| Demon Attack | 0.26 | 0.57 | 0.69 | 0.71 ± 0.01 | **0.74** ± 0.012 |
| Freeway | 0.01 | 0.47 | 0.81 | 0.3 ± 0.355 | **0.86** ± 0.02 |
| Frostbite | 0.51 | **0.76** | 0.75 | 0.73 ± 0.005 | 0.75 ± 0.004 |
| Hero | 0.69 | 0.90 | 0.93 | 0.93 ± 0.008 | **0.94** ± 0.004 |
| Montezuma Revenge | 0.38 | 0.75 | 0.78 | 0.81 ± 0.016 | **0.84** ± 0.014 |
| Ms Pacman | 0.56 | 0.65 | 0.72 | 0.74 ± 0.017 | **0.76** ± 0.011 |
| Pitfall | 0.35 | 0.46 | 0.60 | 0.69 ± 0.031 | **0.73** ± 0.029 |
| Pong | 0.09 | 0.71 | 0.81 | 0.78 ± 0.015 | **0.85** ± 0.004 |
| Private Eye | 0.71 | 0.81 | 0.91 | 0.91 ± 0.009 | **0.93** ± 0.009 |
| Qbert | 0.49 | 0.65 | 0.73 | 0.78 ± 0.026 | **0.79** ± 0.02 |
| Seaquest | 0.56 | 0.66 | 0.67 | 0.68 ± 0.007 | **0.69** ± 0.007 |
| Space Invaders | 0.52 | 0.54 | 0.57 | 0.59 ± 0.007 | **0.62** ± 0.013 |
| Tennis | 0.29 | 0.60 | 0.60 | 0.57 ± 0.018 | **0.64** ± 0.025 |
| Venture | 0.38 | 0.51 | **0.58** | 0.57 ± 0.014 | **0.58** ± 0.01 |
| Video Pinball | 0.45 | 0.58 | 0.61 | 0.6 ± 0.031 | **0.62** ± 0.023 |
| Mean | 0.41 | 0.63 | 0.72 | 0.7 ± 0.033 | **0.75** ± 0.013 |




DIM-UA continues to improve as the total number of units grows, whereas the performance of ST-DIM drops at the same time.

# Linear evaluation accuracy on CIFAR10

| Method | Head | Dimension | | |
|---|---|---|---|---|
| | | 256 | 512 | 1024 |
| SimCLR | - | $0.881 \pm 0.002$ | $0.883 \pm 0.002$ | $0.881 \pm 0.003$ |
| MSimCLR | 2 | $0.877 \pm 0.002$ | $0.878 \pm 0.001$ | $0.866 \pm 0.003$ |
| MSimCLR | 4 | $0.873 \pm 0.001$ | $0.873 \pm 0.001$ | $0.861 \pm 0.002$ |
| MSimCLR | 8 | $0.864 \pm 0.001$ | $0.859 \pm 0.005$ | $0.857 \pm 0.002$ |
| SimCLR-UA | 2 | $0.882 \pm 0.001$ | $0.884 \pm 0.001$ | $0.885 \pm 0.001$ |
| SimCLR-UA | 4 | $0.885 \pm 0.001$ | $0.884 \pm <0.001$ | $0.88 \pm 0.001$ |
| SimCLR-UA | 8 | $0.882 \pm <0.001$ | **0.886** $\pm 0.002$ | $0.876 \pm 0.005$ |

## Changing $\tau$

| $\tau$ | Linear scaling | Accuracy |
|---|---|---|
| 0.2 | | 0.791 |
| 0.2 | ✓ | 0.797 |
| 0.1 | | 0.791 |
| 0.1 | ✓ | 0.8 |
| 0.05 | | 0.799 |
| 0.05 | ✓ | 0.796 |
| 0.02 | | 0.**802** |
| 0.02 | ✓ | 0.785 |

Suggested Improvement: $\tau$ to regulate the $\mathscr{L}_Q$ loss, could be set smaller or set to 0 initially and gradually increased over time.

# Discussion

- UA helps improve the performance of both ST-DIM and SimCLR.
- UA also exhibits the potential of modeling a manifold using further higher dimensions.
- Future research may focus on representing a manifold using UA more efficiently.