Soft Separation and Distillation: Toward Global Uniformity in Federated Unsupervised Learning

Hung-Chieh Fang, Hsuan-Tien Lin, Irwin King, Yifei Zhang

ICCV OCT 19-23, 2025 HONOLULU HAWAII

Background

Unsupervised Representation Learning

Alignment: make similar samples closer

Uniformity: keep maximal information

Federated Learning

- non-iid data dist. across clients
- clients cannot share features or raw data

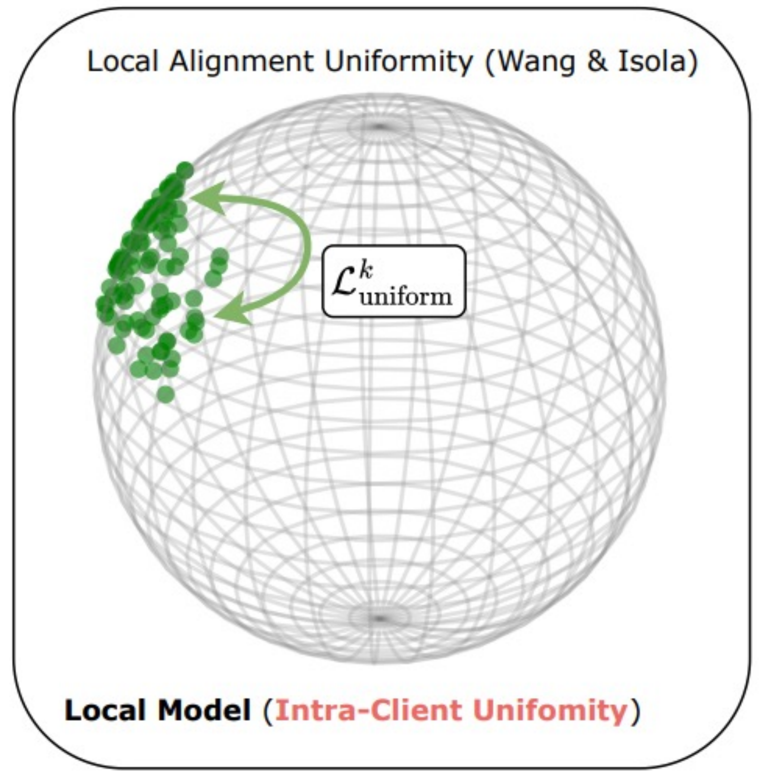
Global Uniformity?

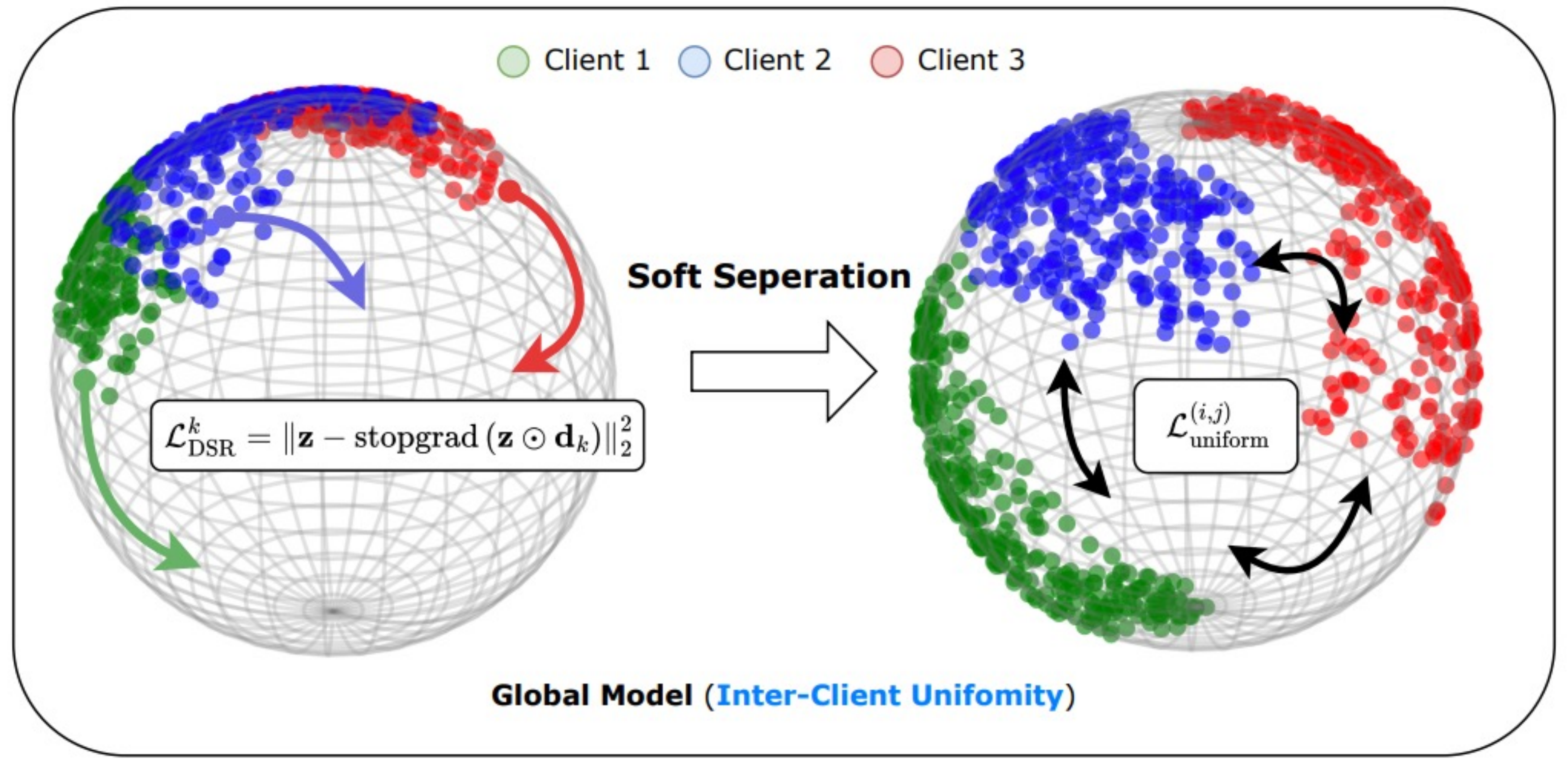
Limited Inter-client Uniformity

$$\mathcal{L}_{\text{uniform}} = -\log\Bigg(\underbrace{\sum_{k=1}^{K} \mathbb{E}_{\mathbf{z},\mathbf{z}' \overset{\text{i.i.d}}{\sim} p_k(\mathbf{z})}\left[e^{-t\|\mathbf{z}-\mathbf{z}'\|_2^2}\right]}_{\text{intra-client } \mathcal{L}^k_{\text{uniform}}} + \underbrace{\sum_{i\neq j} \mathbb{E}_{\mathbf{z}\sim p_i(\mathbf{z}),\mathbf{z}'\sim p_j(\mathbf{z})}\left[e^{-t\|\mathbf{z}-\mathbf{z}'\|_2^2}\right]}_{\text{inter-client } \mathcal{L}^{(i,j)}_{\text{uniform}}}\Bigg),$$

Problem: Under *non-iid* setting, local optimization fail to achieve global (inter-client) uniformity.

Soft Separation and Distillation



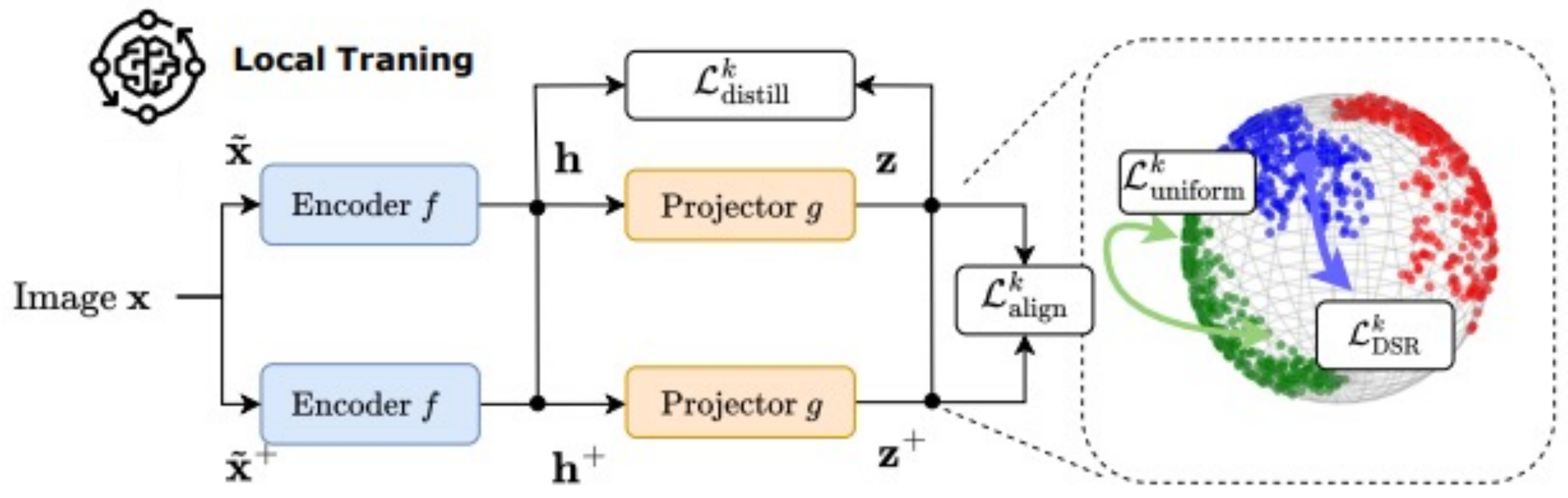


Dimensional-Scaled Regularization

$$\mathcal{L}^k_{\text{DSR}} = \mathbb{E}_{\mathbf{z}\sim p_k(\mathbf{z})}\left[\|\mathbf{z} - \text{stopgrad}(\mathbf{z}\odot\mathbf{d}_k)\|_2^2\right],$$

Key idea: assign client-specific subspaces, encouraging representations to spread toward diverse directions

Projector Distillation

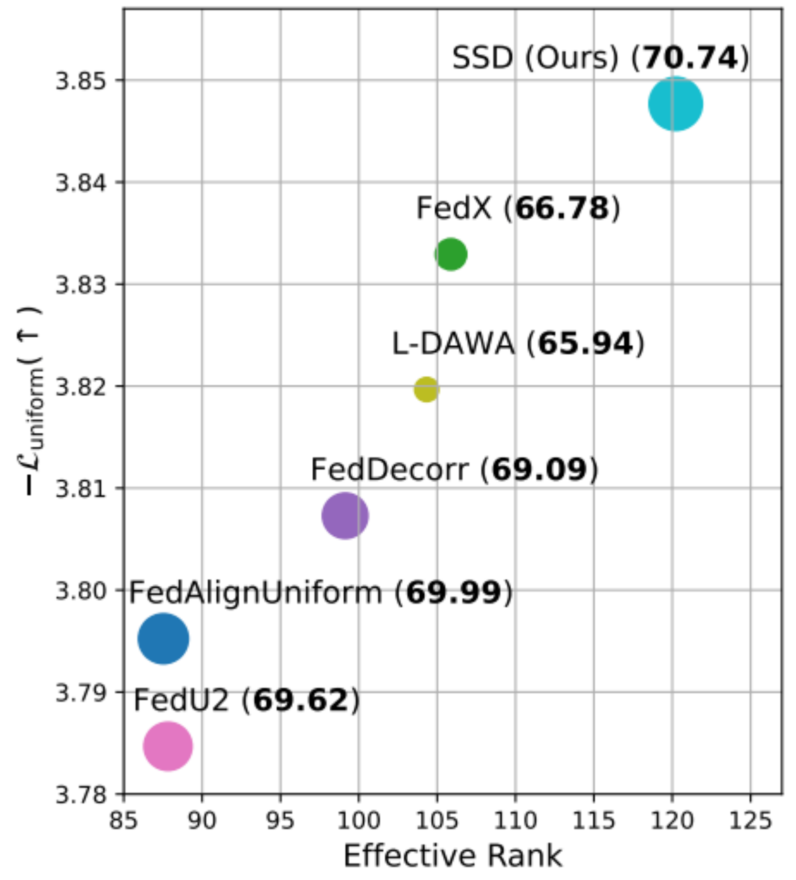


$$\mathcal{L}^k_{\text{distill}} = \mathbb{E}_{x\sim p_k(\mathbf{x})}\left[D_{\text{KL}}\left(\sigma(\mathbf{h})\|\sigma(\mathbf{z})\right)\right],$$

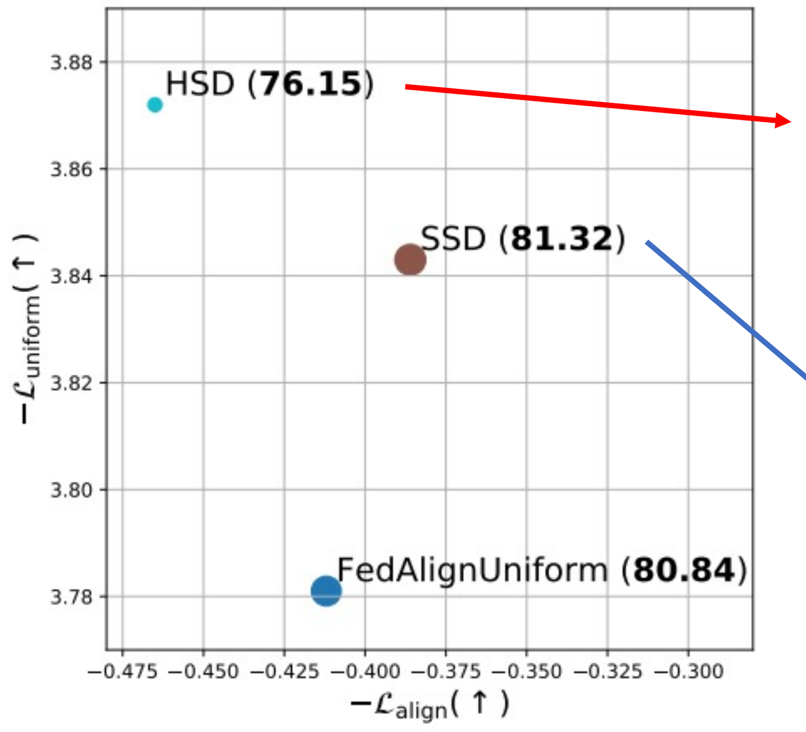
Why? Empirically, DSR enhances uniformity at the *embedding* level, but does not transfer to *representation* level.

Experiments

Transfer Learning



Soft vs. Hard Separation



Hard sep increases uniformity, but at the cost of alignment.

Soft sep (ours) increases both uniformity and alignment.

Why not remove the projector?

	Projector	LP	$-\mathcal{L}_{\text{uniform}}(\uparrow)$
FedAlignUniform	✗	73.16	3.72
+ DSR	✗	76.14 (+2.98)	3.77 (+0.05)
FedAlignUniform	✓	80.84	3.79
+ DSR	✓	81.05 (+0.21)	3.81 (+0.02)