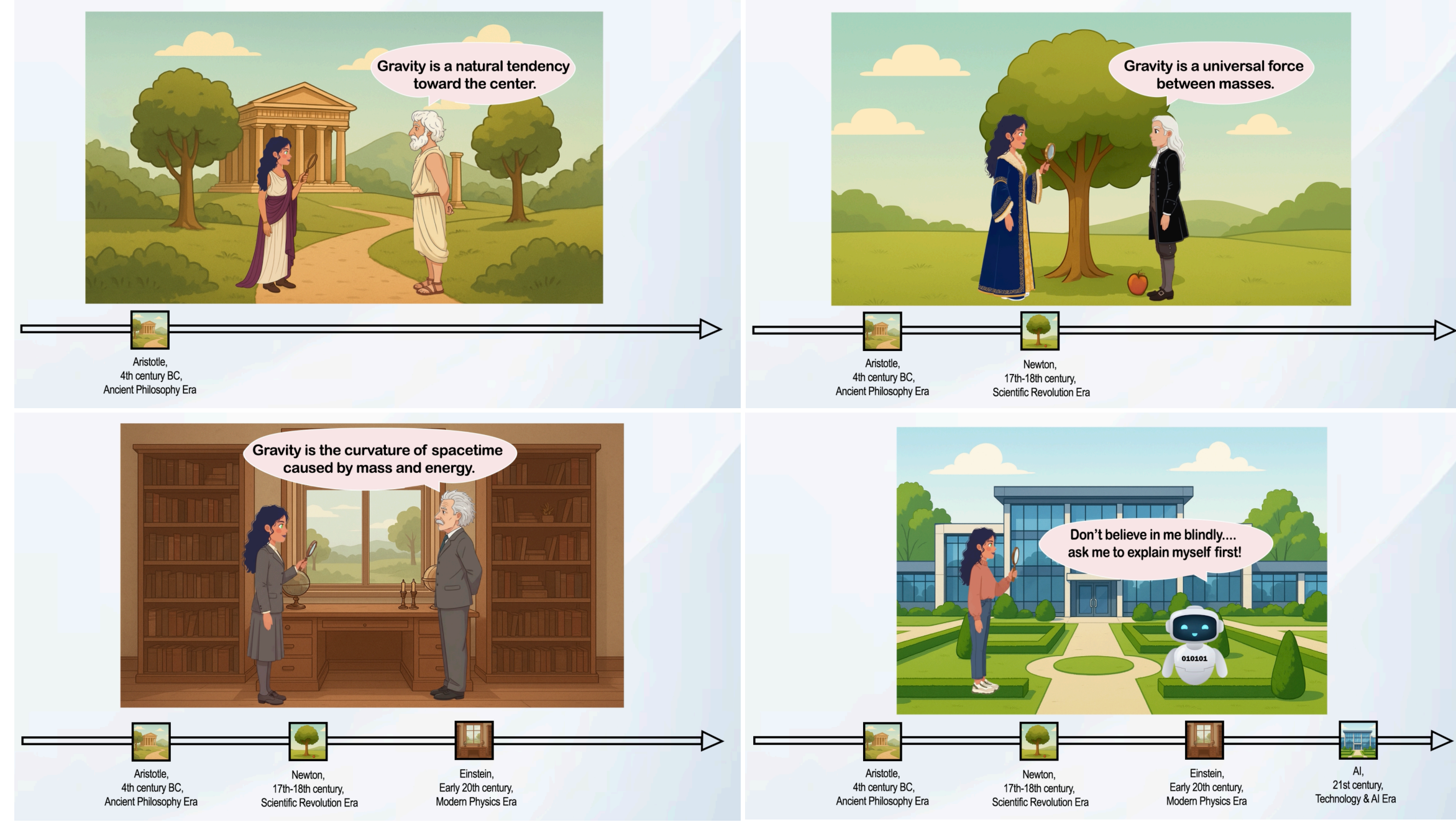


Cost-Free Personalization via Information-Geometric Projection in Bayesian Federated Learning

Nour Jamoussi, Giuseppe Serra, Photios A. Stavrou, Marios Kountouris



Motivation



Explainability: a Lens for Trust!

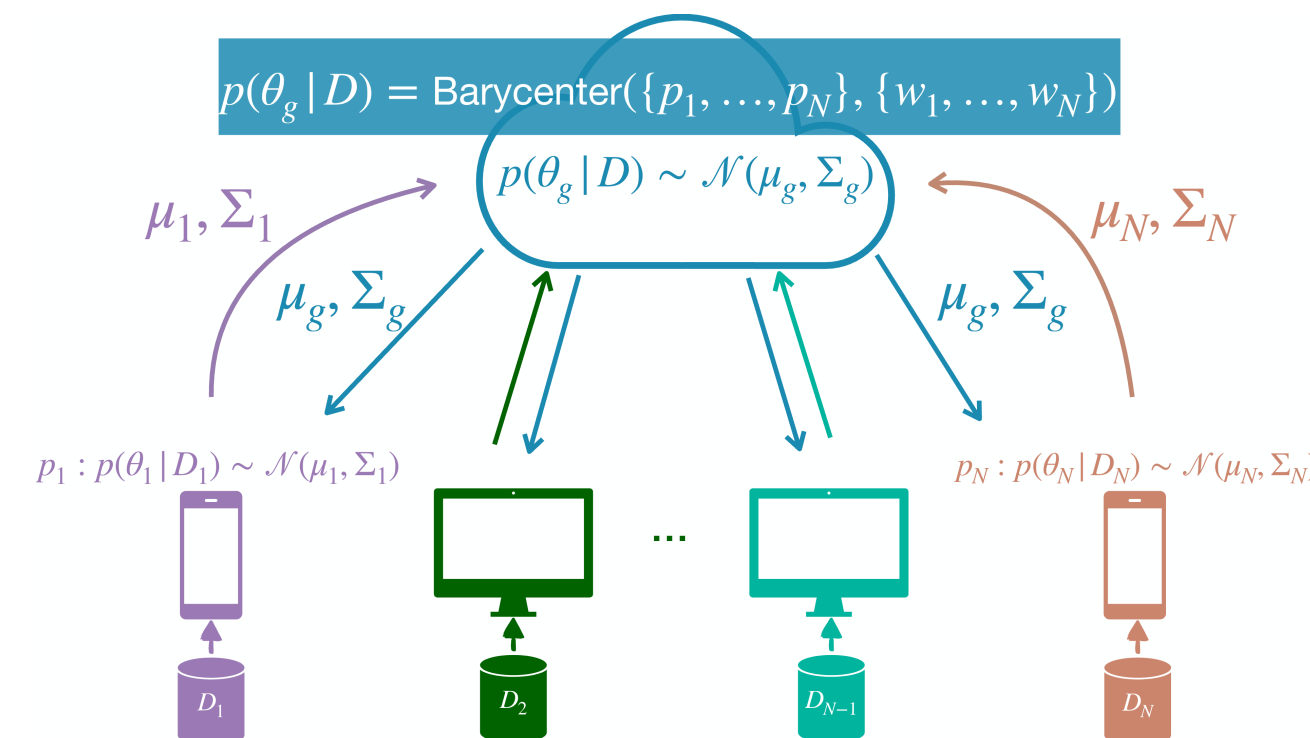
Background

Federated Learning

- Privacy-preserving
- Continual Learning

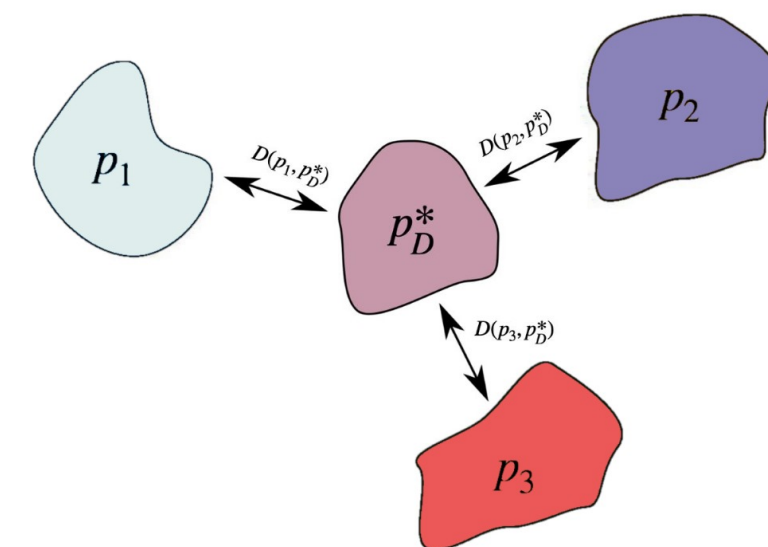
Challenges

- Statistical Heterogeneity → **Personalization**
- Uncertainty Quantification → **Variational Learning**

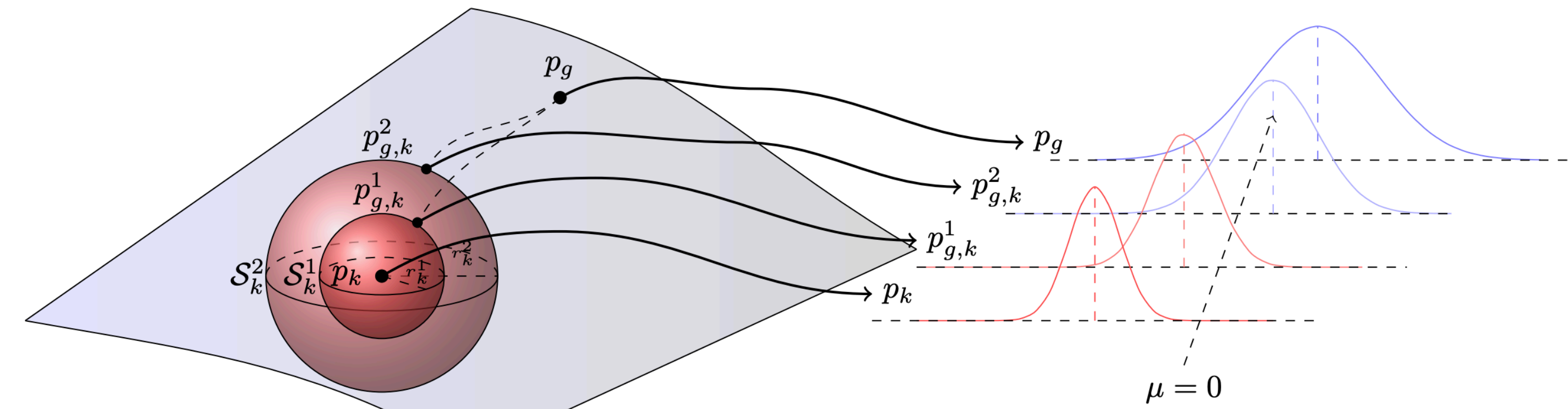


D-Barycenter: Given distributions $\{p_k\}_{k=1}^N \subseteq \mathcal{M}$ with weights $\{w_k\}_{k=1}^N$, the D -barycenter is:

$$p_D^*(\{p_k\}_{k=1}^N, \{w_k\}_{k=1}^N) = \arg \min_{q \in \mathcal{M}} \sum_{k=1}^N w_k D(q || p_k).$$



Main Contribution



Personalization via Information-Geometric Projection

(Theorem) Projection $\leftrightarrow$ Barycenter

Projection on S_k : $p_{g,k} = \arg \min_{p \in S_k} D(p || p_g)$

Equivalent to a weighted barycenter:

$$p_{g,k} = p_D^*(\{p_g, p_k\}, \{w_g, w_k\}),$$

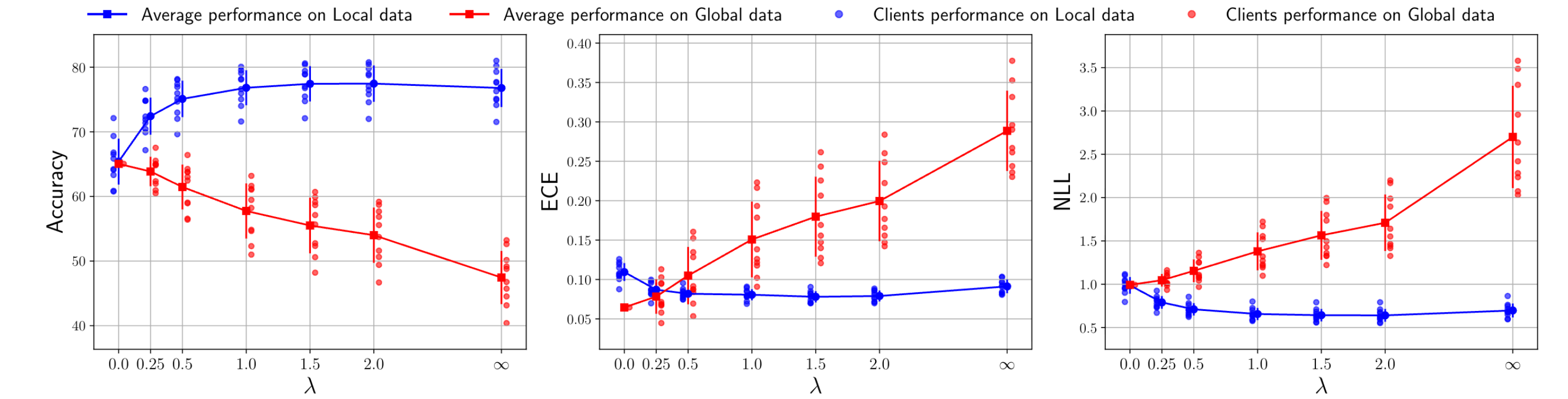
$$\text{with } w_g = \frac{1}{\lambda + 1}, \quad w_k = \frac{\lambda}{\lambda + 1}, \quad \lambda \in [0, \infty)$$

Remark. Larger radius $r_k \Leftrightarrow$ smaller λ

Experimental Setting

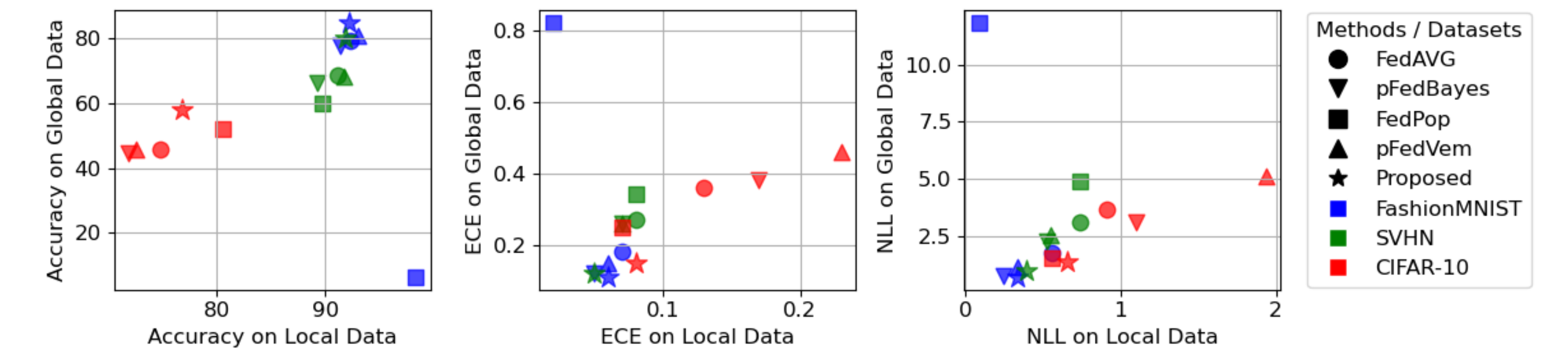
- **IVON for Variational Learning**
- Label Shift: Dirichlet distribution ($\beta = 0.5$)
- Wasserstein Barycenter / Reverse KL Barycenter for Personalization

Trade-off Performance on Local and on Global Data



Effect of λ on performance under local vs. global data distributions (CIFAR-10)

- $\lambda \rightarrow 0$: Global model
- $\lambda \rightarrow \infty$: Local model
- As λ increases:
 - **Global performance deteriorates** (red)
 - **Local performance improves and remains stable over a range of λ** (blue)



Local vs. Global Performance Trade-offs of Personalized Models — SOTA Comparison

- Accuracy: best in the top-right region
- ECE & NLL: best in the bottom-left region
- **Proposed method ($\lambda = 1$) achieves a superior balance across datasets**

Summary

- Proposed information-geometric projection for personalization in Bayesian FL
- Proved equivalence between projection and barycenter formulations
- Demonstrated stable global-local trade-off across datasets

References

Shen, Yuesong, et al. "Variational Learning is Effective for Large Deep Networks." (2024).
Jamoussi, Nour, et al. "Information-Geometric Barycenters for Bayesian Federated Learning." (2024).

nour.jamoussi@eurecom.fr

[linkedin.com/in/nour-jamoussi](https://www.linkedin.com/in/nour-jamoussi)