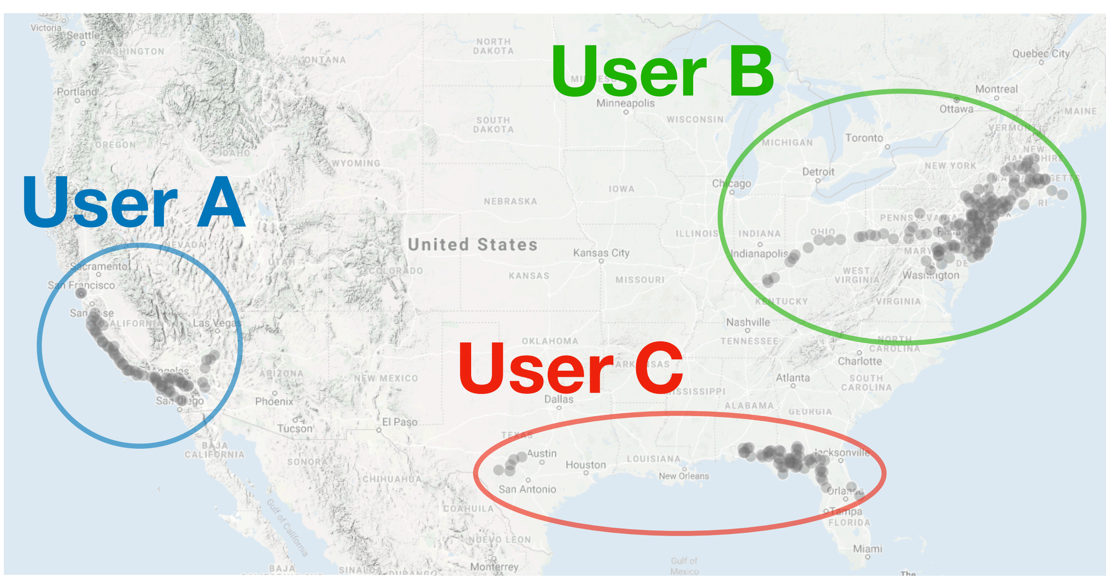
Measuring the Effects of Non-Identical Data Distribution for Federated Visual Classification

Tzu-Ming Harry Hsu, Hang Qi, Matthew Brown

stmharry@{google.com,mit.edu}, hangqi@google.com, mtbr@google.com

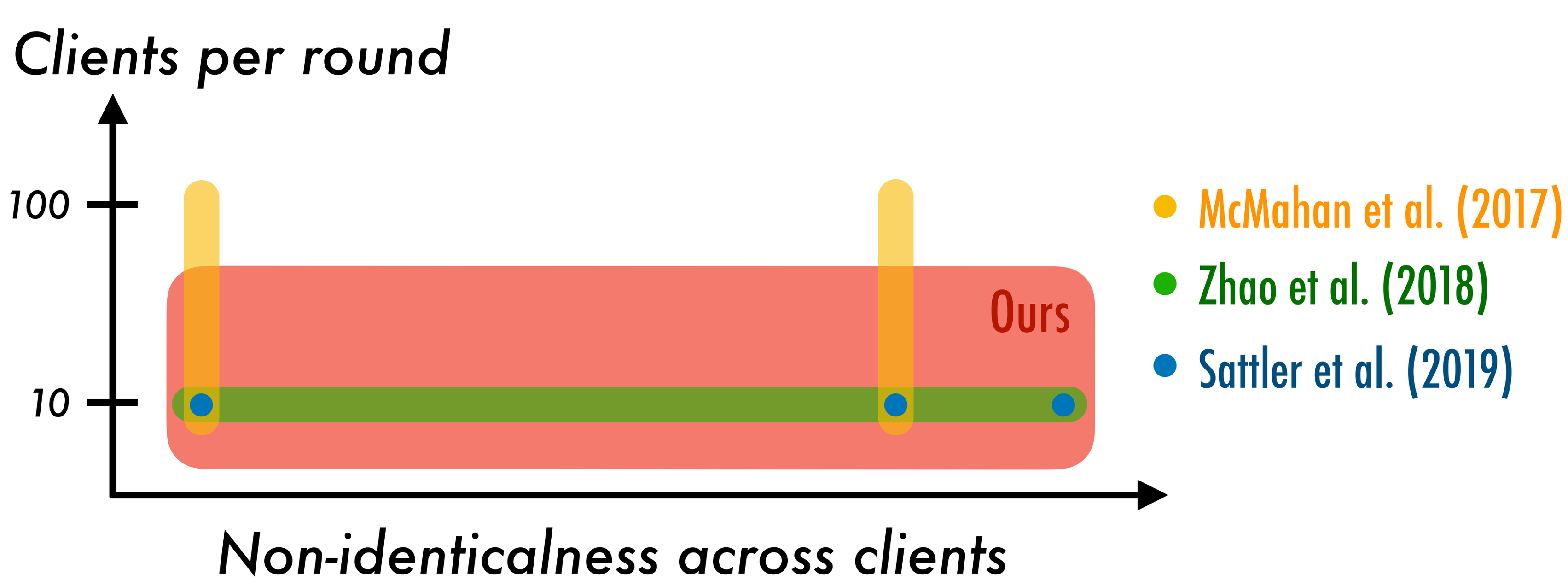
Motivation & Contributions

- In real-world Federated Learning, each participating user has very different and non-identical distribution.



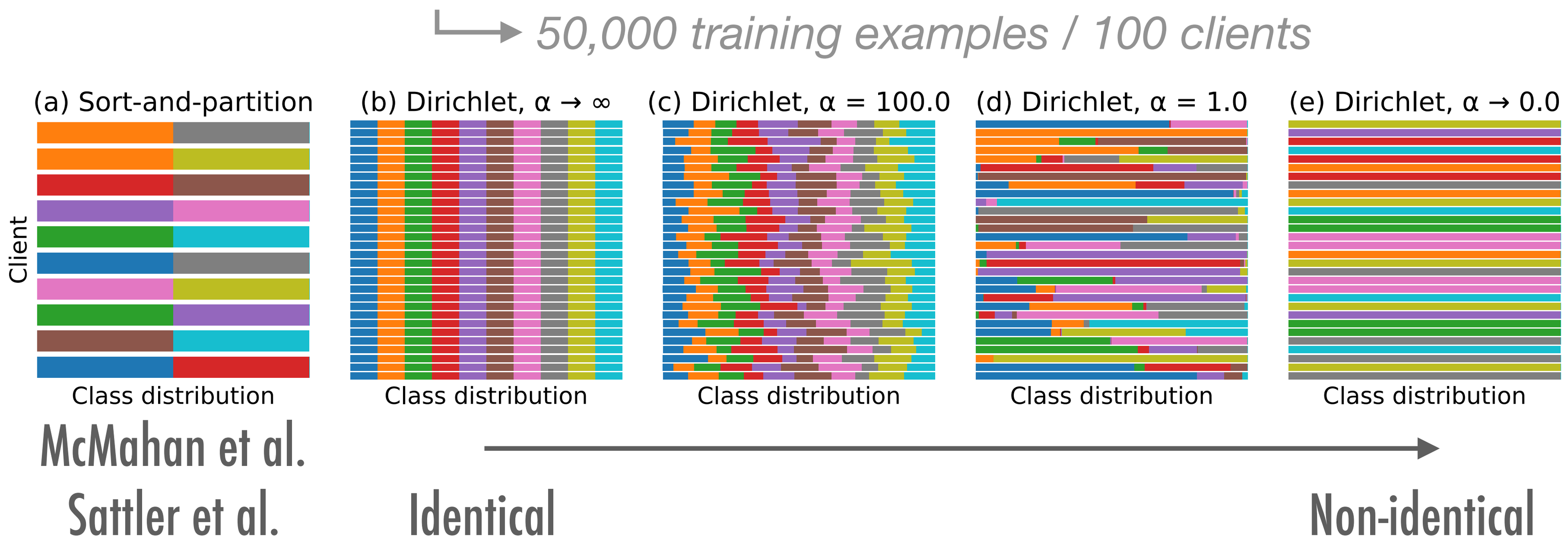
In iNaturalist dataset, users take photos in various geolocations, leading to different species pools

- A realistic and practical setting for Federated Learning is needed to study the effect of ***non-identical data*** and the ***pool size of clients***.



Synthetic Non-Identical Data

- We synthesize 100 federated learning clients from CIFAR-10 by
  1. Draw class-marginal distribution from Dirichlet distribution $\boldsymbol{q} \sim \mathrm{Dir}(\alpha \boldsymbol{p})$, where p is uniform.
     Dirichlet concentration parameter
  2. Assign $(500 \times \boldsymbol{q}_c)$ examples from class c for all classes.
     50,000 training examples / 100 clients



Methods

- Federated Averaging (FedAvg) updates the weights via
  1. Select a fraction C of all clients to report.
  2. Locally train clients with their respective data for E local epochs and yield local model updates $\{\Delta \boldsymbol{w}_k\}_{k=1}^{K}$.
     weight update for client k
  3. Update server weights by $\boldsymbol{w} \leftarrow \boldsymbol{w} - \Delta \boldsymbol{w}$, where

$$\Delta \boldsymbol{w} = \sum_{k=1}^{K} \frac{n_k}{n} \Delta \boldsymbol{w}_k$$

examples for client k

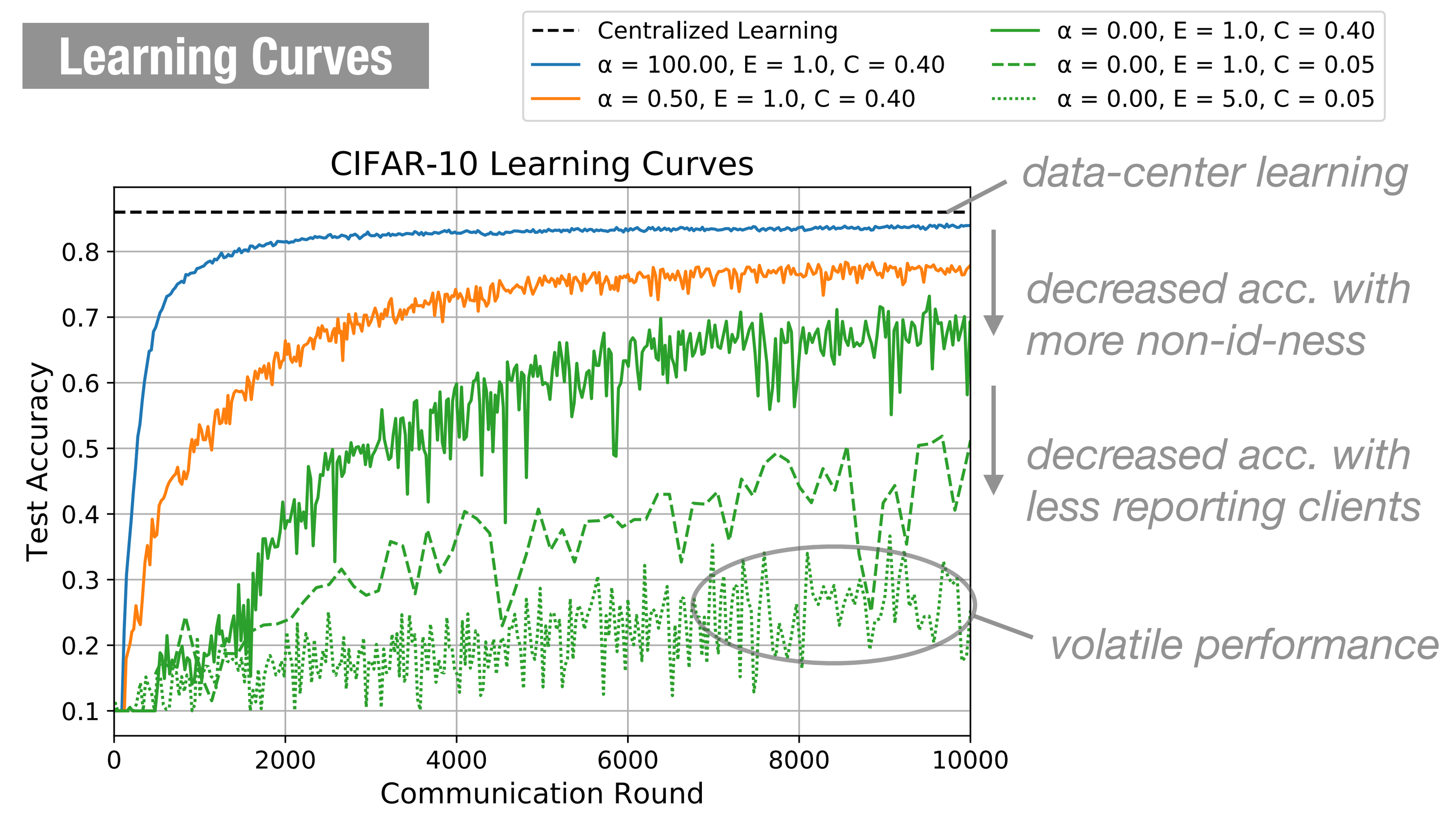
total # examples

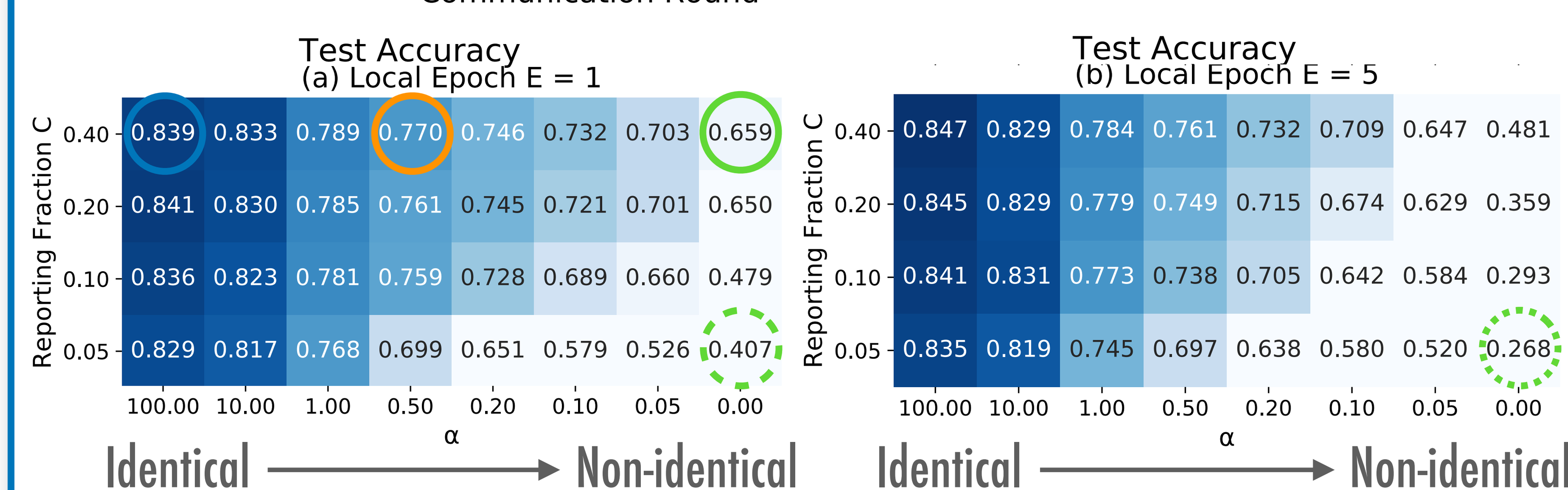
- Federated Averaging with Momentum (FedAvgM) updates server weights by $\boldsymbol{w} \leftarrow \boldsymbol{w} - \boldsymbol{v}$, where $\boldsymbol{v} \leftarrow \beta \boldsymbol{v} + \Delta \boldsymbol{w}$.

Results

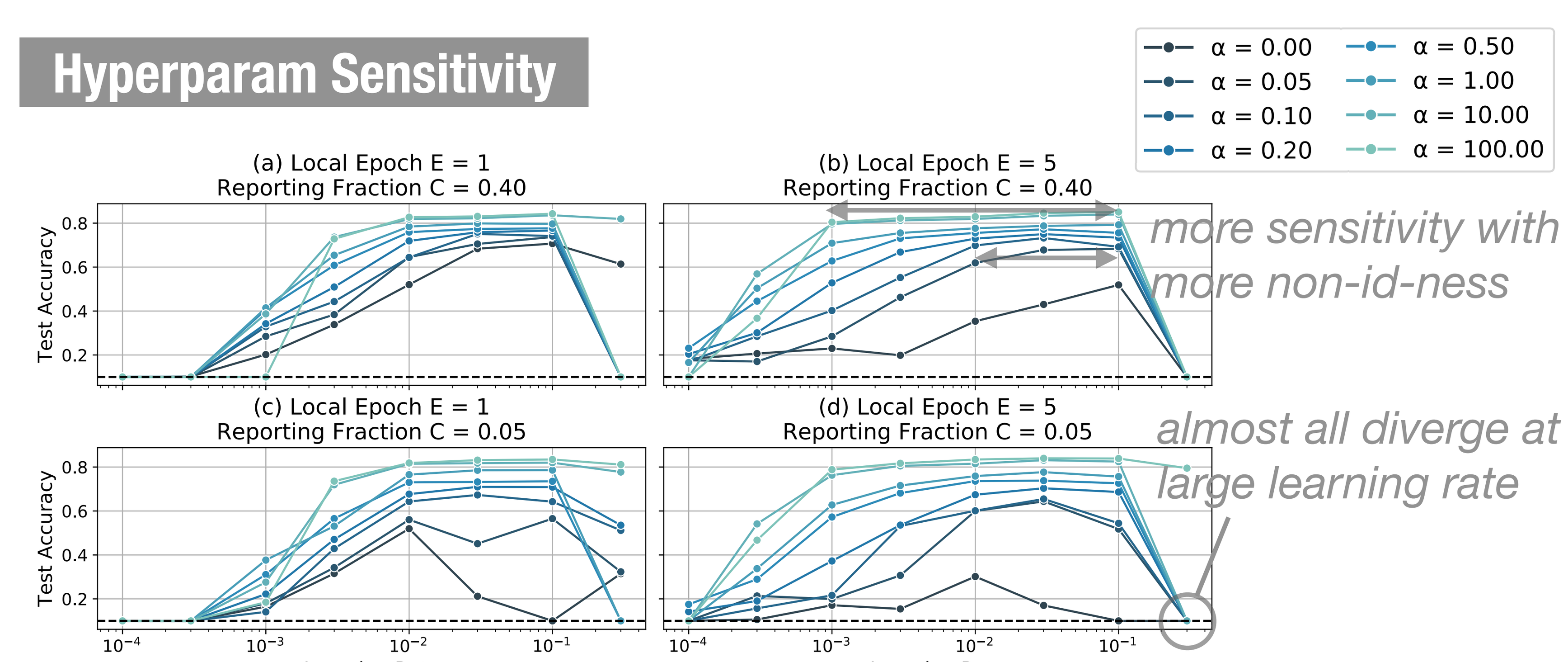
Non-Identical Data on FedAvg

Learning Curves



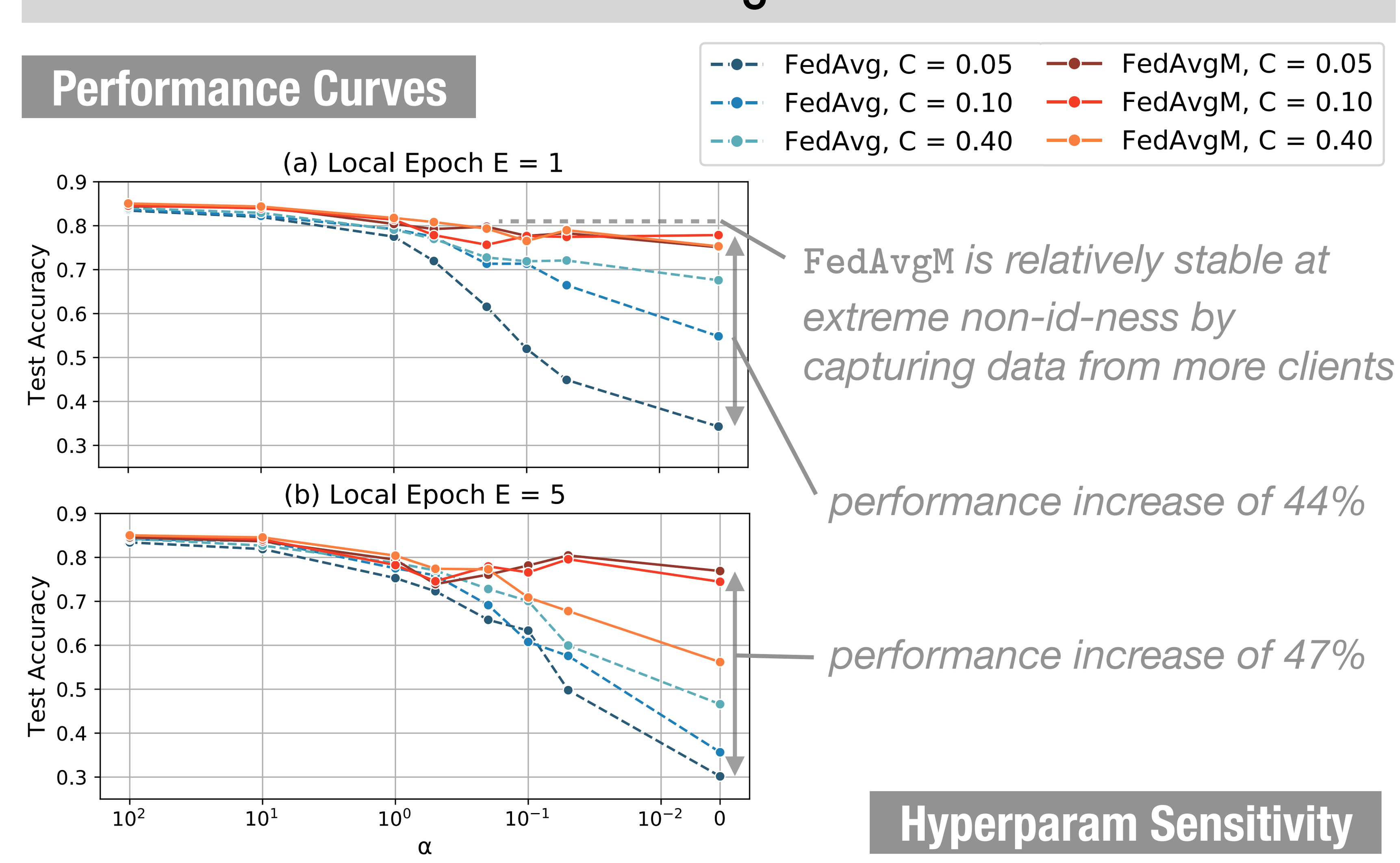


Hyperparam Sensitivity

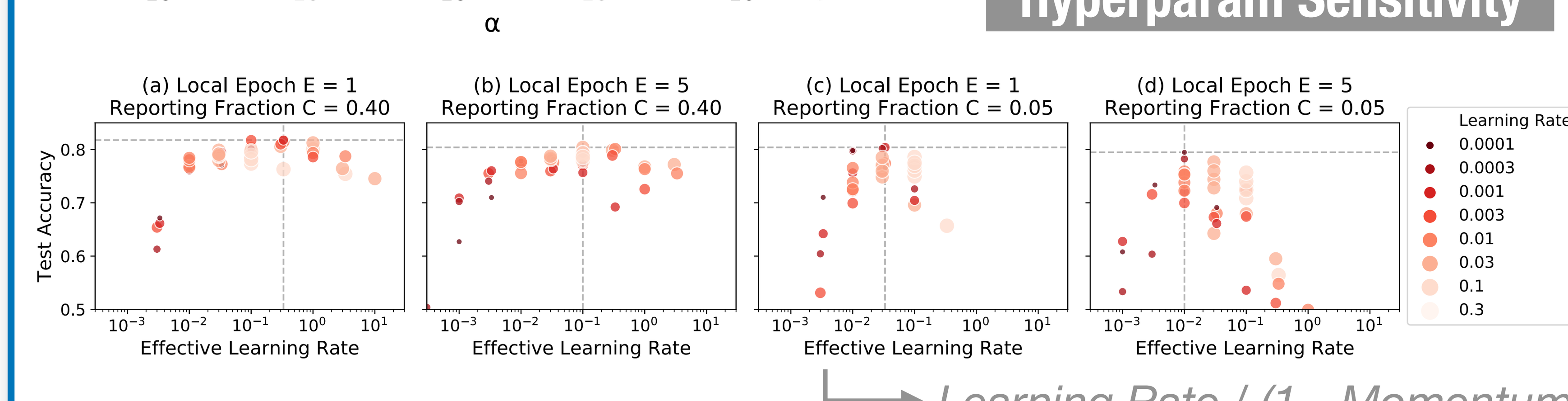


FedAvgM

Performance Curves



Hyperparam Sensitivity



References

McMahan et al. Communication-efficient learning of deep networks from decentralized data. AISTATS 2017.

Zhao et al. Federated learning with non-IID data. arXiv, 2018.

Sattler et al. Robust and communication-efficient federated learning from non-IID data. arXiv 2019.