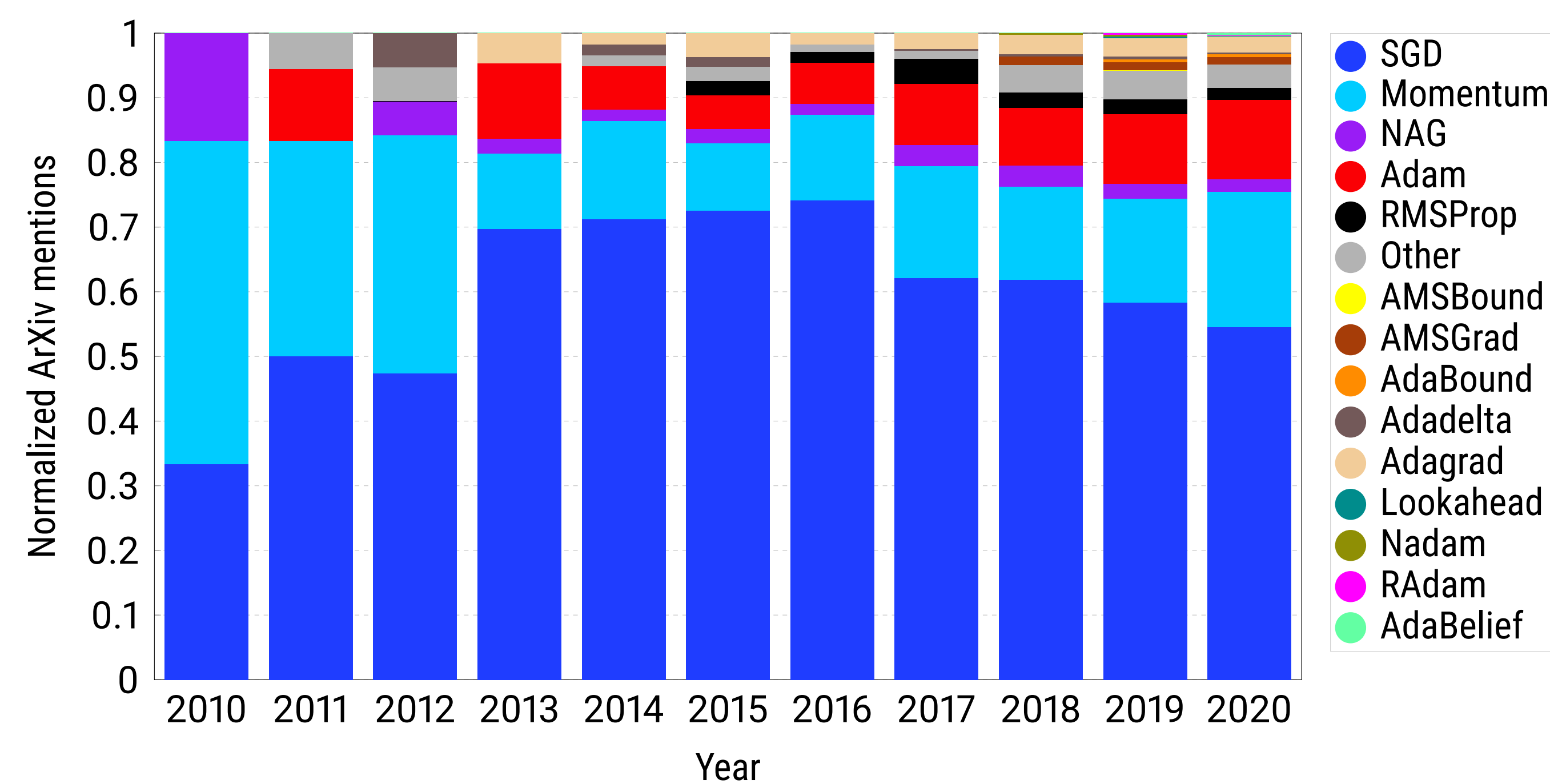


Backpropagation Beyond the Gradient

Felix Dangel
University of Tübingen

Deep learning is stuck with the gradient. But rich information beyond the gradient is affordable.

Problem: Deep Learning is Stuck with the Gradient



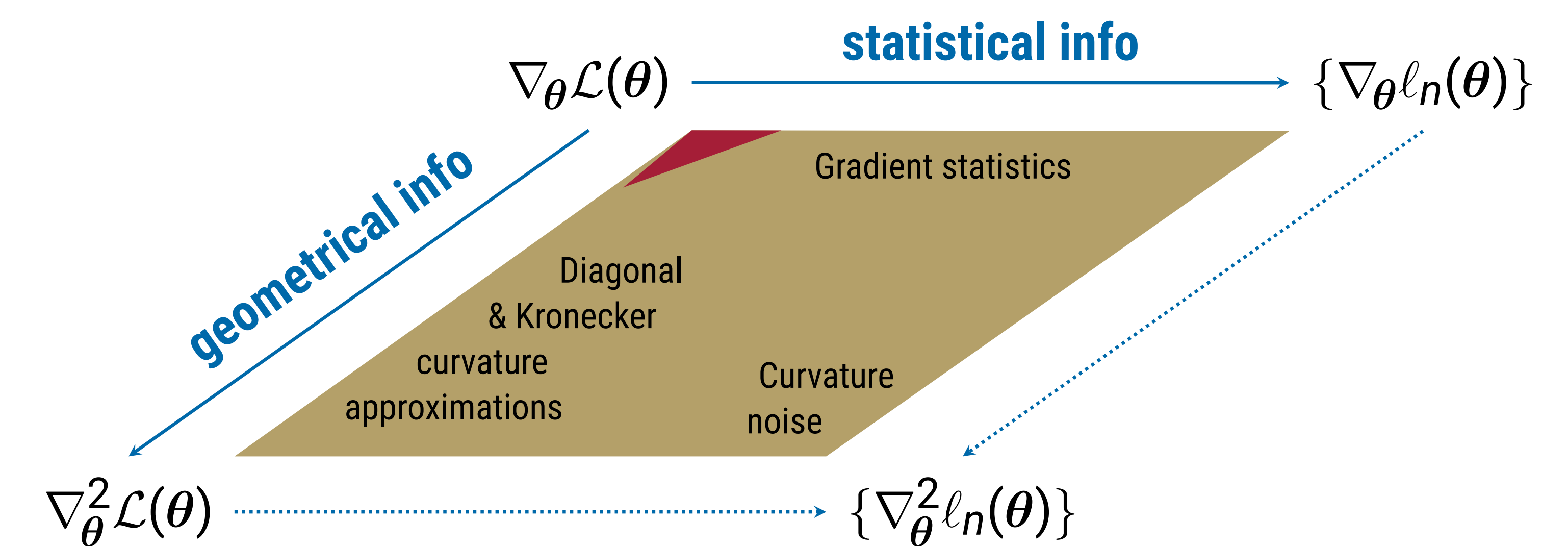
Despite efforts by the community, there is currently no method that clearly dominates the competition.

[...] tuning helps about as much as trying other optimizers.

[Schmidt et al., ICML 2021]

Remedy: Higher-order Information

The loss is an average over per-sample losses: $\mathcal{L}(\theta) = \frac{1}{N} \sum_n \ell(\mathbf{f}_\theta(\mathbf{x}_n), \mathbf{y}_n)$.



We built tools that make it efficiently accessible. This simplifies experimentation and enables novel research.



ICLR 2020

BackPACK: Efficiently Compute More Than Just the Gradient During a Backward Pass

All quantities are extensions of the standard backward pass.

```

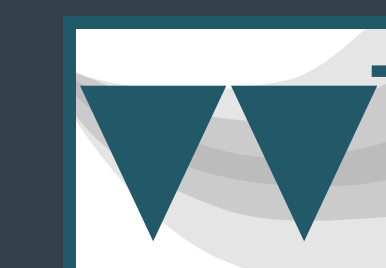
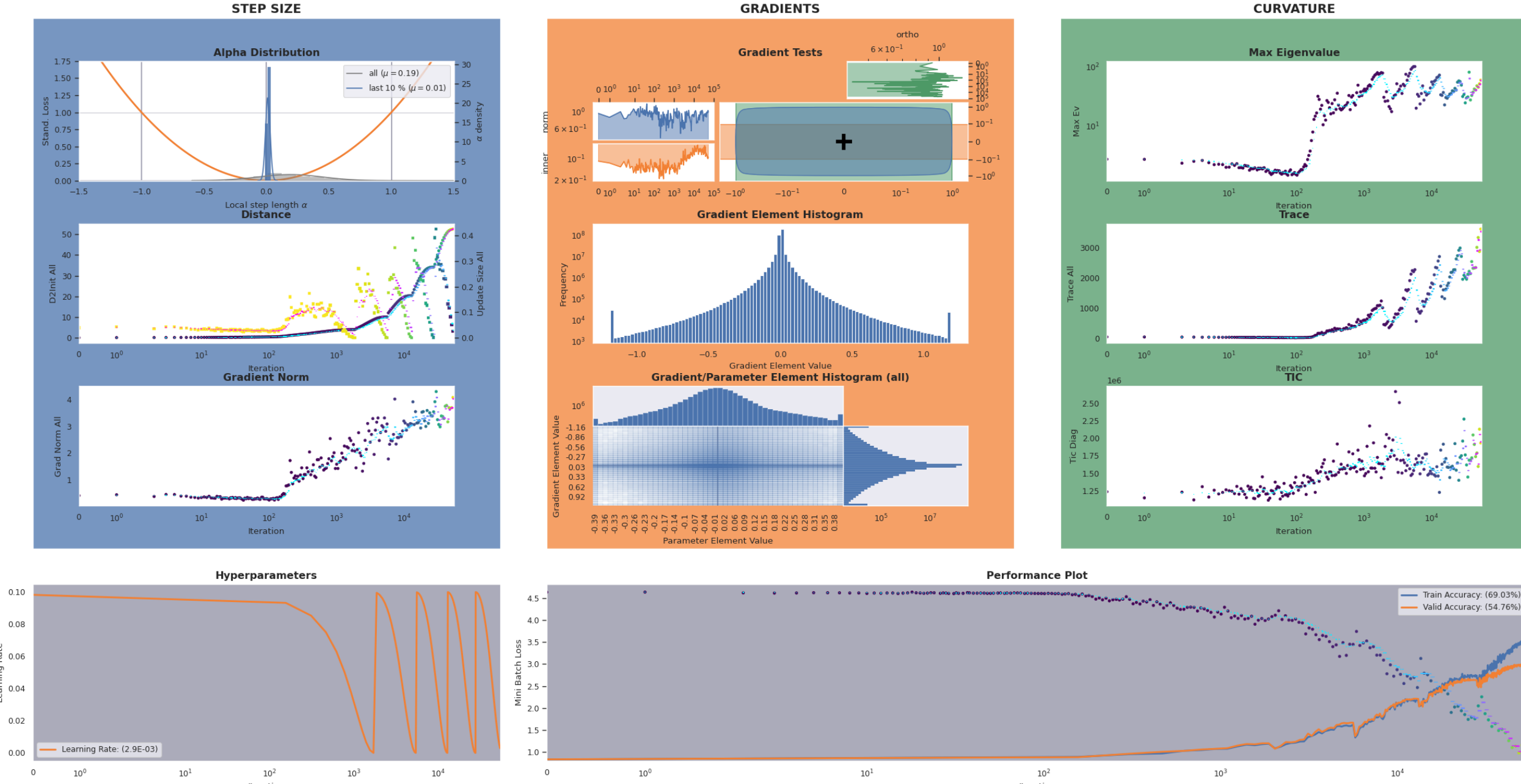
1 X, y = load_mnist_data()
2 model = extend(Linear(784, 10))
3 loss_func = extend(CrossEntropyLoss())
4 loss = loss_func(model(X), y)
5
6 with backpack(DiagHessian()):
7     loss.backward()
8
9 for param in model.parameters():
10     print(param.grad)
11     print(param.diag_h)
12
13 
```



NeurIPS 2021

Cockpit: Figure out What Is Happening Inside Your Neural Network During Training

A collection of instruments based on BackPACK's statistics.



pre-print

ViViT: Richer Access to Curvature Information (Including Curvature Noise)

BackPACK add-on to obtain more precise curvature information.

