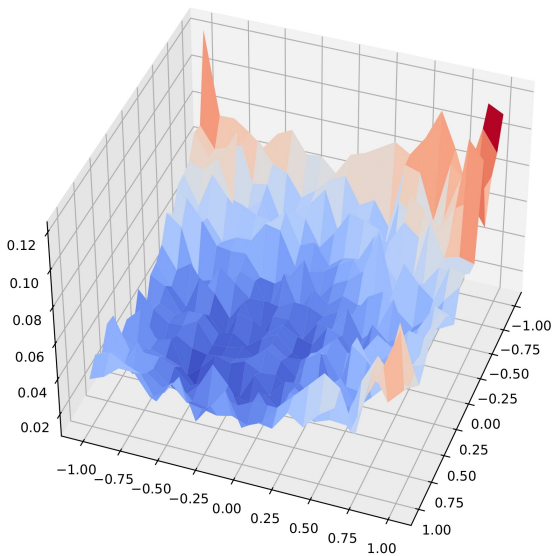
Neural Networks Loss Landscape Convergence in Different Low-Dimensional Spaces

Goal: Measure how the loss function changes as the training set size grows:

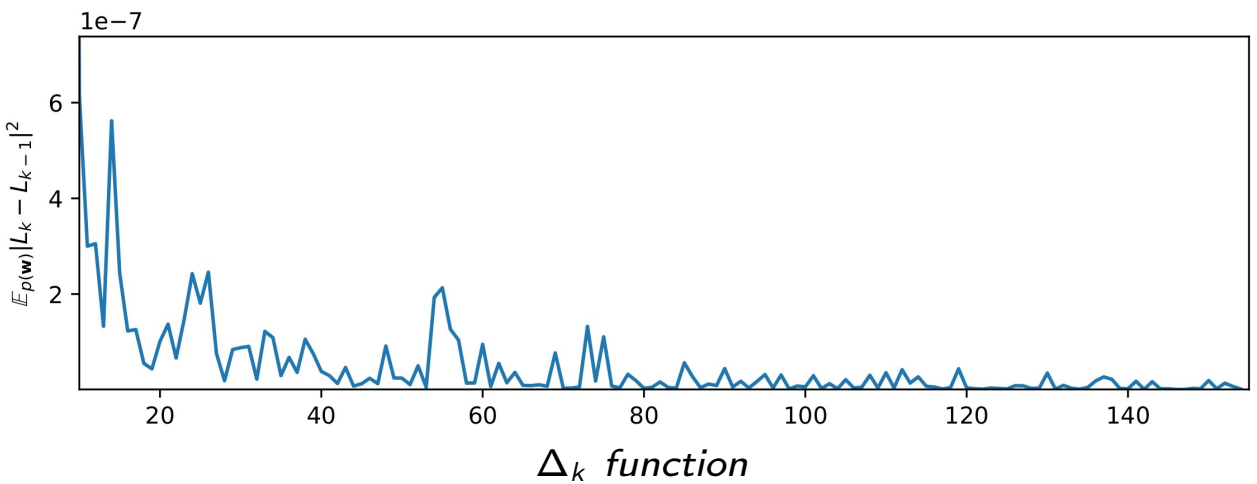
$$\Delta_k = \mathbb{E}\Big(\mathcal{L}_{k+1}(\mathbf{w}) - \mathcal{L}_k(\mathbf{w})\Big)^2.$$

Method:

- **Monte Carlo:** Generate points near the minimum according to $p(\mathbf{w})$ and average the differences.
- **Hessian Eigenvectors:** Use directions with the largest eigenvalues to focus on key curvature components.



Loss function



Δ_k *function*