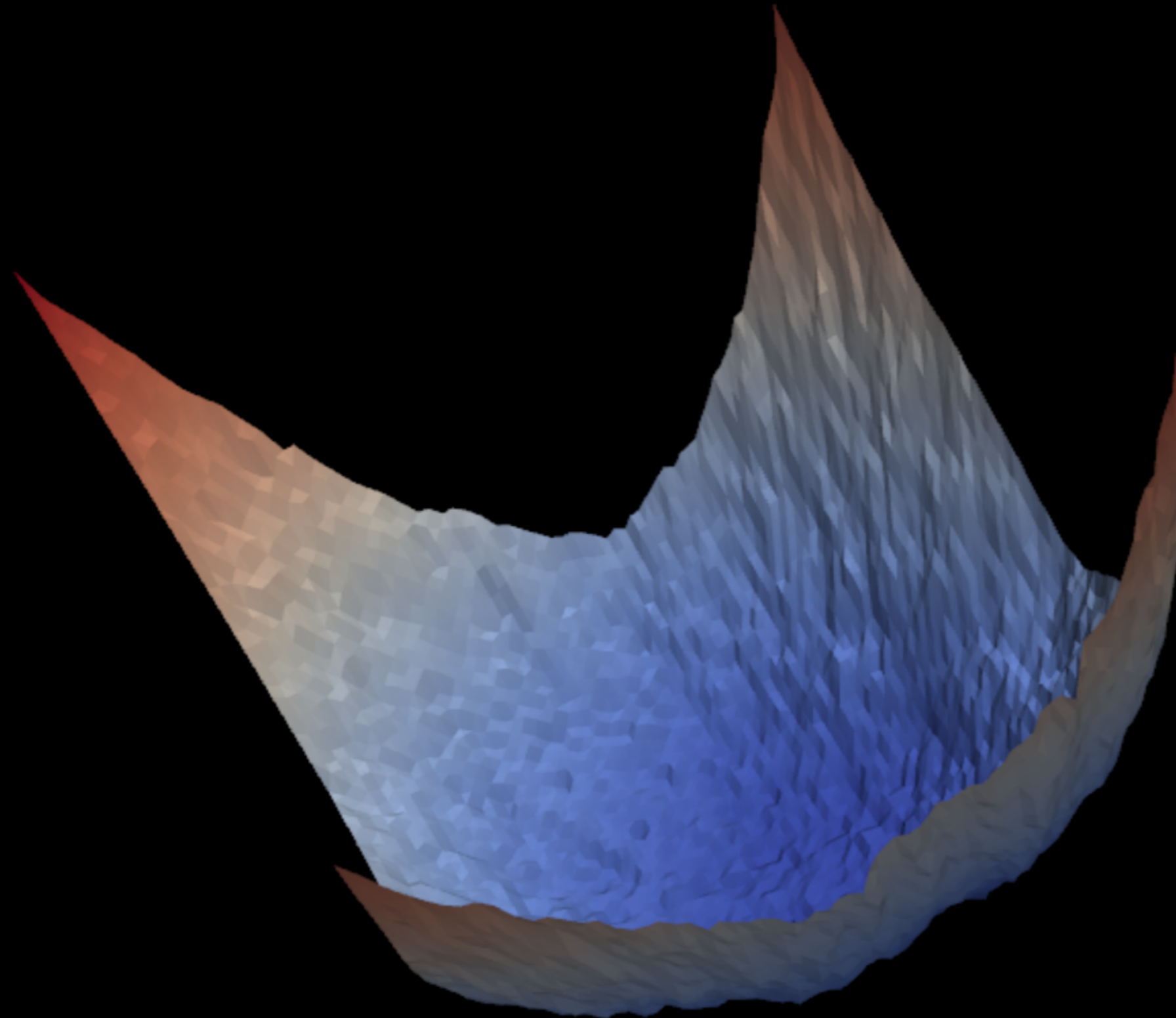


**Richard Cepka. Loss Landscape.** Deep learning is well known for its non-convex and highly dimensional nature, but DNNs often converge (to local minima) and find a good solution. This is also possible due to some 'tricks', e.g. BatchNorm makes the loss landscape significantly smoother. In this picture can be seen a visualization of 51 789 578 dimensional loss landscape (at CIFAR-10) at the initialization phase of ResNet50. ResNet50 is a network topology that incorporates these 'tricks', like BatchNorm.



4.7

CrossEntropyLoss  
5.3

5.9