



THE WEIZMANN INSTITUTE OF SCIENCE
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Machine Learning and Statistics Seminar

Room 290C ,Ziskind Building
on Wednesday, Jun 08, 2016
at 11:15

Daniel Soudry
Columbia University

No bad local minima: Data independent training error guarantees for multilayer
neural networks

Abstract:

We use smoothed analysis techniques to provide guarantees on the training loss of Multilayer Neural Networks (MNNs) at differentiable local minima. Specifically, we examine MNNs with piecewise linear activation functions, quadratic loss and a single output, under mild over-parametrization. We prove that for a MNN with one hidden layer, the training error is zero at every differentiable local minimum, for almost every dataset and dropout-like noise realization. We then extend these results to the case of more than one hidden layer. Our theoretical guarantees assume essentially nothing on the training data, and are verified numerically. These results suggest why the highly non-convex loss of such MNNs can be easily optimized using local updates (e.g., stochastic gradient descent), as observed empirically.