

1135-68-1051

**Yunkai Zhang\*** (zhang.yunkai98@gmail.com), **Yu Ma, Zhaoqi Li** and **Catalina Marie Vajiac**. *Exploration of Numerical Precision in Deep Neural Networks*.

Reduced numerical precision is a common technique to lower computational cost in various Deep Neural Networks (DNNs). While it has been observed that DNNs are resilient to small errors and noise, there exists no general result capable of predicting the sensitivity to reduced precision for a given DNN system architecture. In this project, we emulate arbitrary bit-width using a specified floating-point representation and truncating the remainder after a certain number of bits. This truncation is applied to the neural network after every batch. We show results on two representative networks, MNIST and CIFAR-10. In these, we explore the impact of several model parameters and their impact on the network's training accuracy. We then present a preliminary theoretical investigation of the error scaling in both forward and backward propagations. We end with a discussion of the implications of these results as well as the potential for generalization to other network architectures. (Received September 18, 2017)