Mean capacity with standard deviation per epoch of training for different data sizes, activation functions, and layers amount 167965 Configuration 160000 layers amount, data size 150000 1, 20000 140000 -2, 20000 4, 20000 130000 1,50000 120000 -2,50000 110000 4,50000 100000 Capacity 1, 100000 85972 4, 100000 2, 100000 70000 activation function 60000 **RReLU** 50000 softmax 34908 20000 10000 0 50 100 150 200 250 300 350 400 **Epochs**