

# Adaptive Methods for Nonconvex Optimization

## Problem Set

Grant Block, Duke Kwon, Priya Sapra

December 2020

### 1 Problem 1

One of the main results of YOGI is that it can be shown that the bound on the stationary condition decreases linearly with increased batch size.

- a. Implement YOGI in your HW6 autoencoder.
- b. Run your autoencoder with YOGI with minibatch sizes of 16, 32, 64, 128. Comment on results, what trends do you see?

### 2 Problem 2

The paper also stated that the optimal YOGI parameters are  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 10^{-3}$ . The paper does not provide much justification for these parameters.

- a. With a minibatch size of 128 in your autoencoder, run YOGI with those parameters, and some others of your choice
- b. Discuss your observations, and say whether you agree with the paper that those are the optimal parameters.