Influence of the Choice of Dataset and Learning Rate (Training Set Size: 256) Mode: FUMBRELLA Mode: VANILLA 0.8 0.7 accuracy 00 0.5 Dataset: qqp, Learning Rate: 1e-05 Dataset: qqp, Learning Rate: 1e-05 Dataset: qqp, Learning Rate: 5e-05 Dataset: qqp, Learning Rate: 5e-05 Dataset: sst2, Learning Rate: 1e-05 Dataset: sst2, Learning Rate: 1e-05 0.4 Dataset: sst2, Learning Rate: 5e-05 Dataset: sst2, Learning Rate: 5e-05 2.25 2.00 1.75 eval 1.50 le 1.25 1.00 0.75 0.50 Step Step 0.7 0.6 0.5 train_st2_loss 0 0 0 w 0 0 0.2 0.1 0.0 100 700 0 100 200 300 400 500 600 700 800 200 300 400 500 600 800