BERT (large) GLUE MNLI on 1 GPU accuracy 8.0 0.7 0.6 Validation 0.5 0.4 4bs 1gpu 1vn (baseline) 32bs 1gpu 8vn 25000 15000 5000 35000 45000 Step