synchronous Batch Size: 32, Bias: only label 8 0.990 10 0.985 **—** 20 **一 25** --- 30 0.980 35 40 Accuracy 45 0.975 50 55 0.970 0.965 0.960 10^{1} Communication Cost (GB)

synchronous Batch Size: 64, Bias: only label 8 0.990 10 **—** 15 0.985 **→** 20 **—** 25 30 0.980 **—** 35 40 Accuracy 45 0.975 50 55 0.970 0.965 0.960 10^1 Communication Cost (GB)

synchronous Batch Size: 128, Bias: only label 8 0.985 0.980 Accuracy 0.970 60 0.965 0.960 10^{1} Communication Cost (GB)

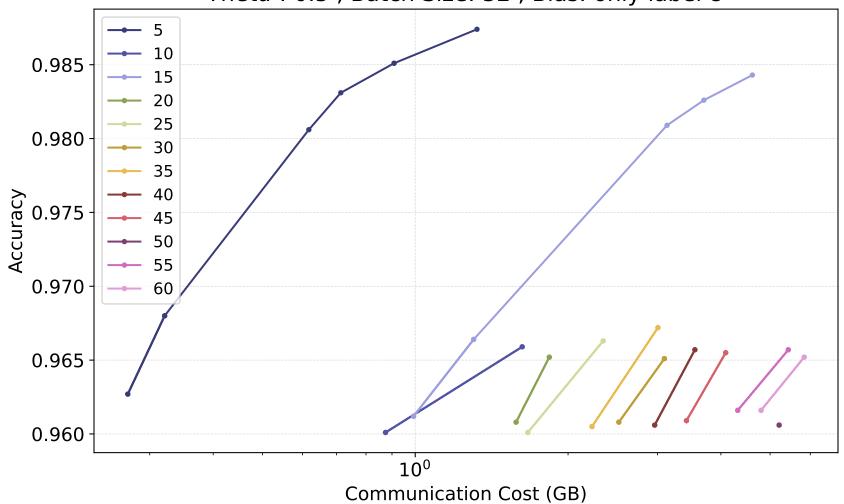
synchronous Batch Size: 256, Bias: only label 8 0.990 10 15 0.985 20 25 30 0.980 35 40 Accuracy 45 0.975 50 55 ---- 60 0.970 0.965 0.960 10^{1} Communication Cost (GB)

gm Theta: 0.5, Batch Size: 32, Bias: only label 8 0.990 10 0.985 20 30 0.980 35 40 Accuracy 45 0.975 50 55 60 0.970 0.965 0.960 10⁰ 10^{1} Communication Cost (GB)

naive Theta: 0.5, Batch Size: 32, Bias: only label 8 0.990 10 0.985 0.980 Accuracy 0.975 55 60 0.970 0.965 0.960 100 Communication Cost (GB)

linear

Theta: 0.5, Batch Size: 32, Bias: only label 8



sketch Theta: 0.5, Batch Size: 32, Bias: only label 8 10 0.985 20 0.980 30 35 Accuracy 0.75 40 45 50 55 0.970 60 0.965 0.960 100 Communication Cost (GB)

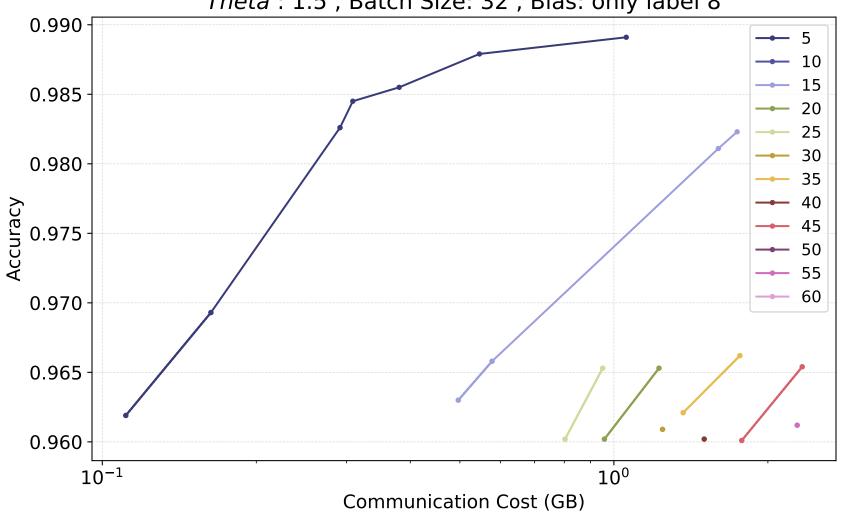
gm Theta: 1.5, Batch Size: 32, Bias: only label 8 0.990 10 0.985 0.980 Accuracy 0.975 60 0.970 0.965 0.960 100 Communication Cost (GB)

naive Theta: 1.5, Batch Size: 32, Bias: only label 8 10 0.985 **—** 20 25 0.980 30 35 Accuracy 40 45 50 55 0.970 **6**0 0.965 0.960 100

Communication Cost (GB)

linear

Theta: 1.5, Batch Size: 32, Bias: only label 8



sketch Theta: 1.5, Batch Size: 32, Bias: only label 8 10 0.985 20 25 0.980 30 35 Accuracy 2.60 2.60 40 45 50 55 0.970 **—** 60 0.965 0.960 100 Communication Cost (GB)

gm Theta: 3.0, Batch Size: 32, Bias: only label 8 0.985 10 0.980 20 25 Accuracy 0.975 30 35 40 0.970 50 55 0.965 60 0.960 10^{-1} 100 Communication Cost (GB)

naive Theta: 3.0, Batch Size: 32, Bias: only label 8 0.985 10 **—** 15 **→** 20 0.980 **3**0 **—** 35 Accuracy 97.00 97. 40 45 50 55 **←** 60 0.965 0.960 100 10^{-1} Communication Cost (GB)

linear Theta: 3.0, Batch Size: 32, Bias: only label 8 **→** 10 0.985 **—** 20 **—** 25 0.980 --- 30 **—** 35 **-** 40 Accuracy 0.975 45 50 55 0.970 **6**0 0.965 0.960 10^{-1} 100 Communication Cost (GB)

sketch Theta: 3.0, Batch Size: 32, Bias: only label 8 0.990 10 0.985 **→** 20 **—** 25 ---- 30 0.980 **—** 35 **-** 40 Accuracy 45 0.975 50 55 ---- 60 0.970 0.965 0.960 100 Communication Cost (GB)

gm Theta: 5.0, Batch Size: 32, Bias: only label 8 0.985 **—** 20 25 0.980 30 35 40 Accuracy 0.975 45 50 55 0.970 **60** 0.965 0.960 10^{-1} 100 Communication Cost (GB)

naive Theta: 5.0, Batch Size: 32, Bias: only label 8 10 0.985 **-** 15 **—** 20 **—** 25 0.980 ---- 30 - 35 - 40 Accuracy 0.975 45 50 55 **—** 60 0.970 0.965 0.960 10^{-1} Communication Cost (GB)

linear Theta: 5.0, Batch Size: 32, Bias: only label 8 10 0.985 **—** 15 **—** 20 **—** 25 0.980 30 35 40 Accuracy 0.975 45 50 55 0.970 -- 60 0.965 0.960 2×10^{-1} 10^{-1} 3×10^{-1} 4×10^{-1} Communication Cost (GB)

sketch Theta: 5.0, Batch Size: 32, Bias: only label 8 10 0.985 **—** 15 **→** 20 0.980 ---- 30 **—** 35 Accuracy 0.052 - 40 45 50 55 0.970 **60** 0.965 0.960 100 Communication Cost (GB)

gm Theta: 7.0, Batch Size: 32, Bias: only label 8 0.990 10 0.985 0.980 Accuracy 0.975 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)

naive Theta: 7.0, Batch Size: 32, Bias: only label 8 0.985 **→** 20 **←** 25 0.980 ---- 30 **—** 35 40 Accuracy 0.975 45 50 55 **—** 60 0.970 0.965 0.960 10^{-1} 2×10^{-1} 3×10^{-1} 4×10^{-1} Communication Cost (GB)

linear Theta: 7.0, Batch Size: 32, Bias: only label 8 ---- 10 0.985 0.980 ---- 30 **—** 35 Accuracy 0.975 45 50 55 0.970 - 60 0.965 0.960 3×10^{-1} 4×10^{-1} 2×10^{-1} 6×10^{-2} 10^{-1} Communication Cost (GB)

sketch Theta: 7.0, Batch Size: 32, Bias: only label 8 0.985 0.980 Accuracy 0.970 0.965 0.960 3×10^{-1} 4×10^{-1} 2×10^{-1} 6×10^{-1} 100 Communication Cost (GB)



naive Theta: 10.0, Batch Size: 32, Bias: only label 8 0.985 0.980 O.975 O.970 0.965 0.960 2×10^{-1} 10^{-1} Communication Cost (GB)

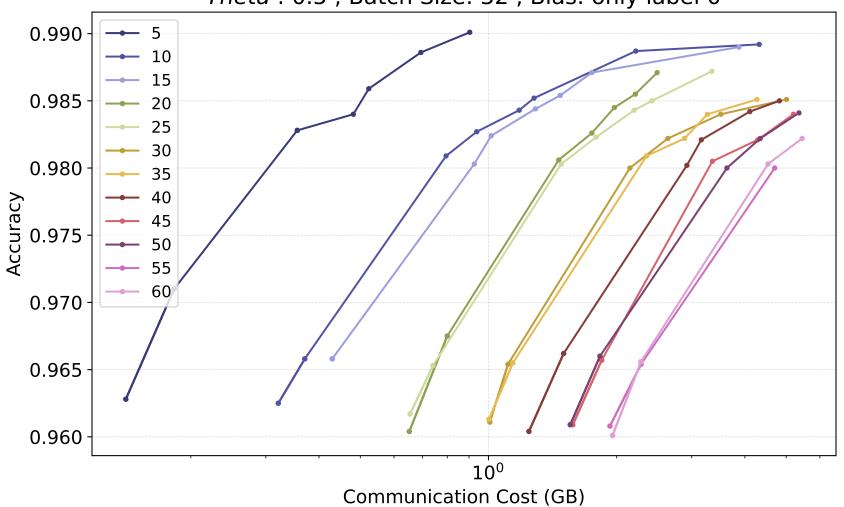
linear Theta: 10.0, Batch Size: 32, Bias: only label 8 0.985 0.980 25 30 35 Accuracy Accuracy 2.00 40 45 50 55 0.970 **—** 60 0.965 3×10^{-1} 2×10^{-1} 10^{-1} Communication Cost (GB)

sketch Theta: 10.0, Batch Size: 32, Bias: only label 8 0.980 Accuracy contract contract contracts 0.970 0.965 4×10^{-1} 6×10^{-1} 3×10^{-1} Communication Cost (GB)

gm Theta: 0.5, Batch Size: 32, Bias: only label 0 0.990 0.985 0.980 10 Accuracy 0.975 20 25 30 0.970 35 40 **–** 45 0.965 **–** 50 55 -- 60 0.960 100 10^{1} Communication Cost (GB)

naive Theta: 0.5, Batch Size: 32, Bias: only label 0 0.990 0.985 0.980 Accuracy 0.975 20 25 30 0.970 35 40 **-** 45 0.965 **–** 50 55 **→** 60 0.960 100 Communication Cost (GB)

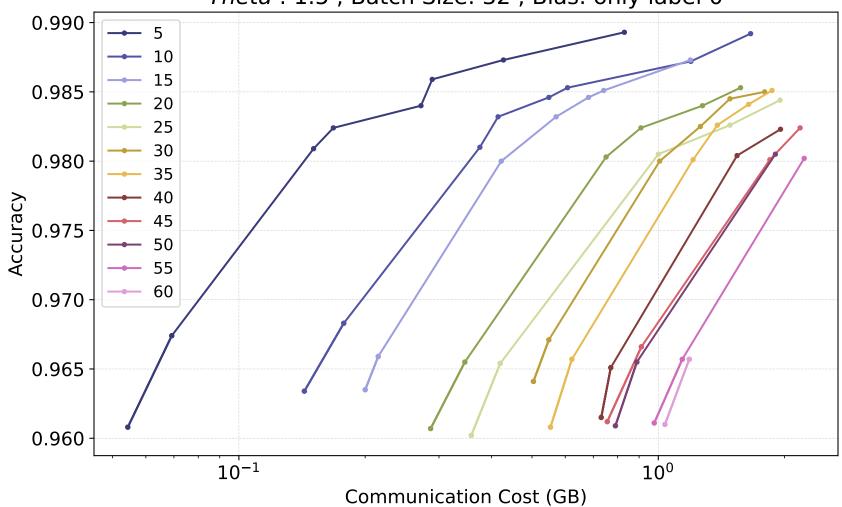
linear *Theta*: 0.5, Batch Size: 32, Bias: only label 0



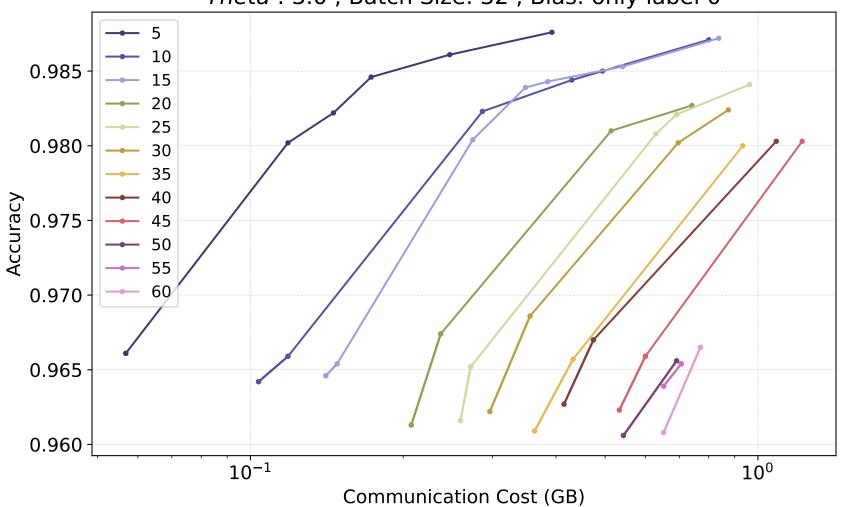
naive Theta: 1.5, Batch Size: 32, Bias: only label 0 0.990 10 0.985 0.980 Accuracy 0.975 55 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)

linear

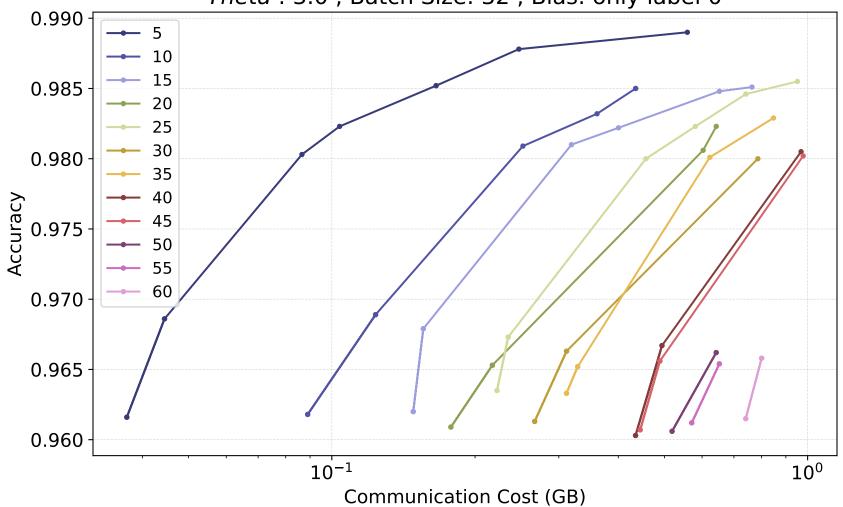
Theta: 1.5, Batch Size: 32, Bias: only label 0



naive *Theta*: 3.0, Batch Size: 32, Bias: only label 0



linear Theta: 3.0, Batch Size: 32, Bias: only label 0



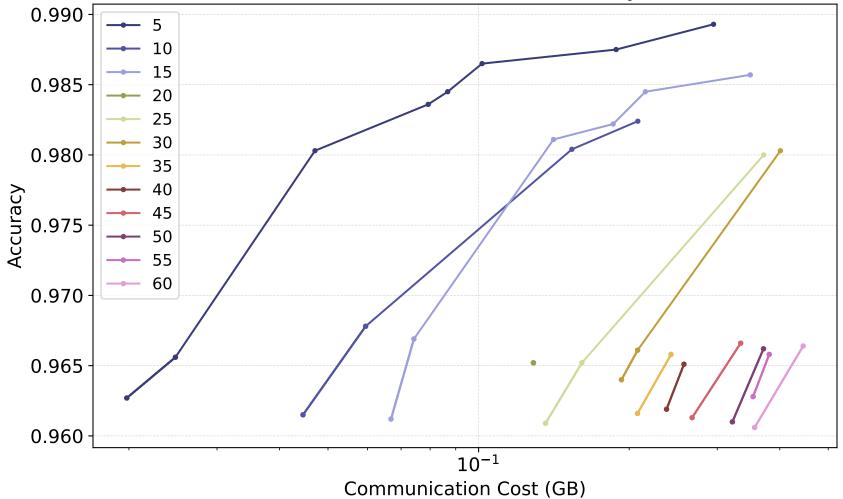
naive Theta: 5.0, Batch Size: 32, Bias: only label 0 0.985 0.980 Accuracy 970.0 970.0 970.0 50 0.970 0.965 0.960 10^{-1} Communication Cost (GB)

linear Theta: 5.0, Batch Size: 32, Bias: only label 0 0.990 10 0.985 0.980 Accuracy 0.975 55 60 0.970 0.965 0.960 10^{-1} Communication Cost (GB)

naive Theta: 7.0, Batch Size: 32, Bias: only label 0 0.985 0.980 Accuracy 0.052 0.970 60 0.965 0.960 10^{-1} Communication Cost (GB)

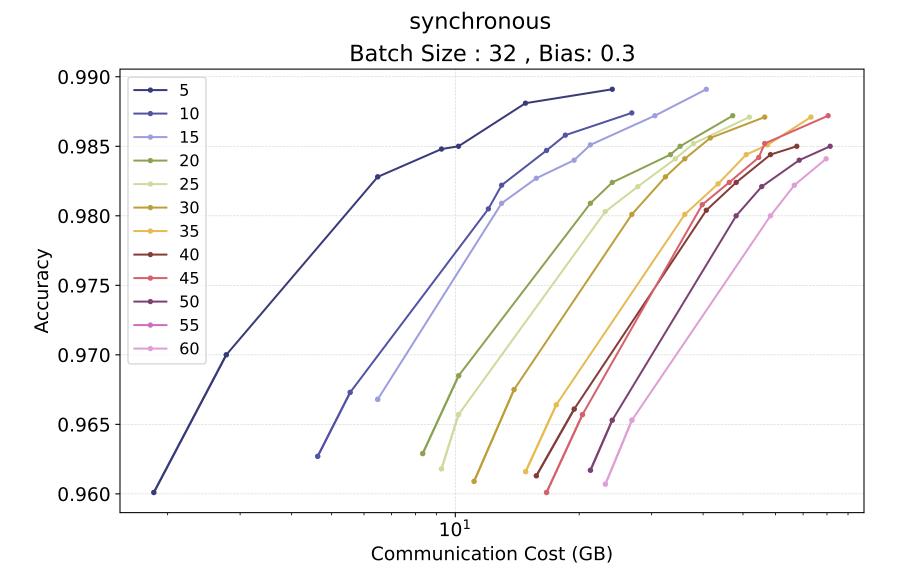
linear

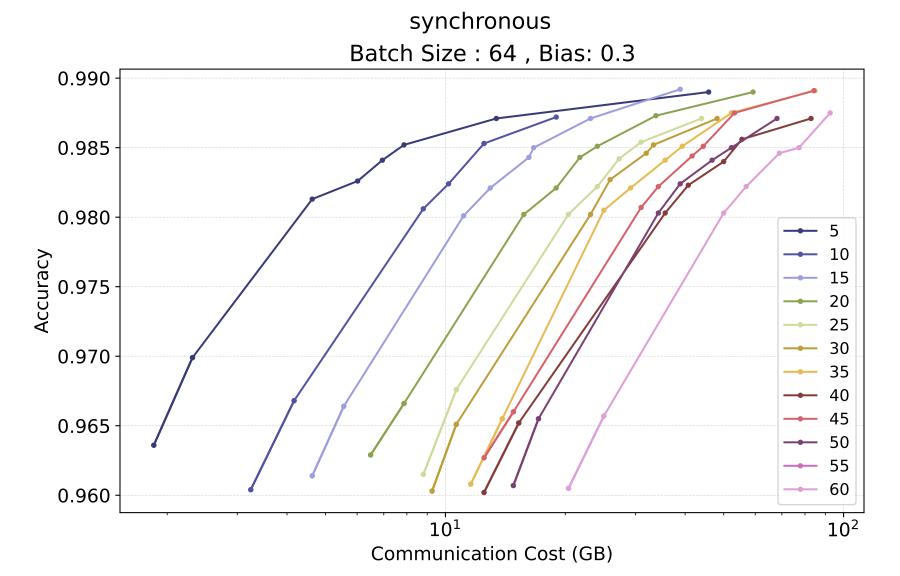
Theta: 7.0, Batch Size: 32, Bias: only label 0

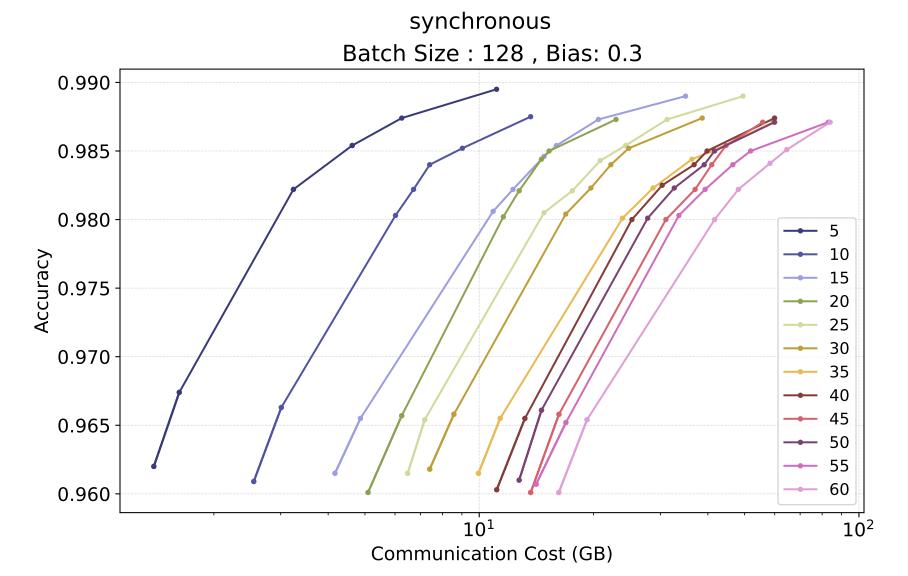


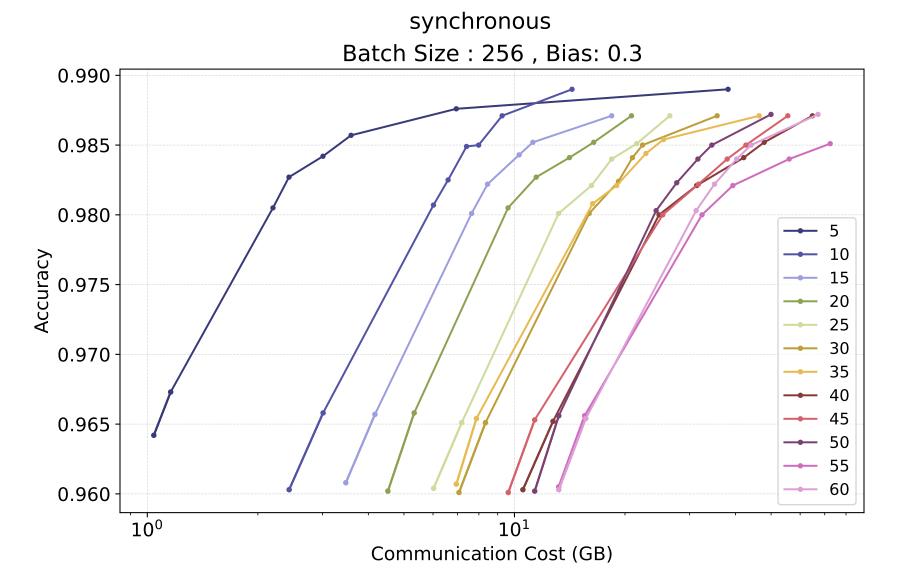
naive Theta: 10.0, Batch Size: 32, Bias: only label 0 0.985 **→** 20 0.980 **→** 30 35 Accuracy 0.052 45 50 55 0.970 **6**0 0.965 0.960 10 -1 Communication Cost (GB)

linear Theta: 10.0, Batch Size: 32, Bias: only label 0 **—** 10 0.985 20 25 0.980 30 35 40 Accuracy 0.975 45 50 55 0.970 60 0.965 0.960 10^{-1} Communication Cost (GB)

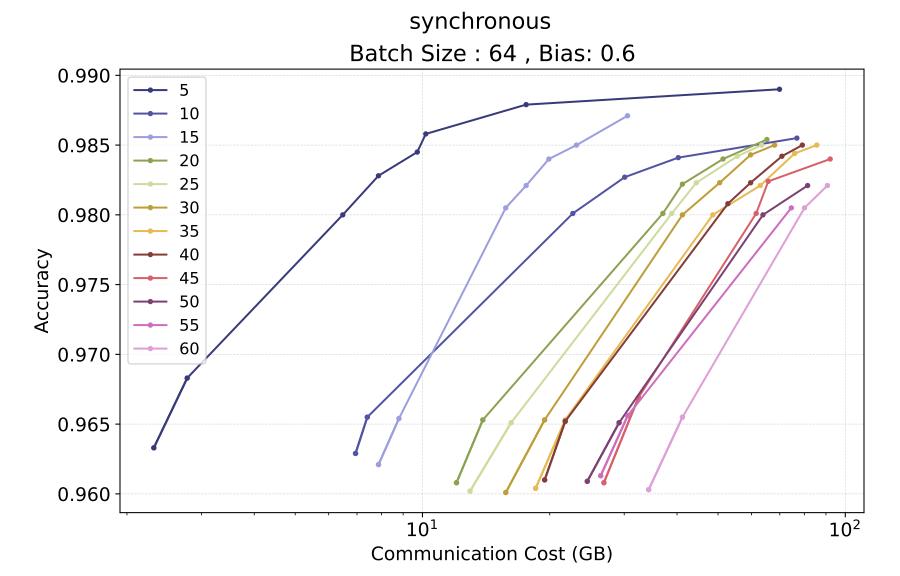




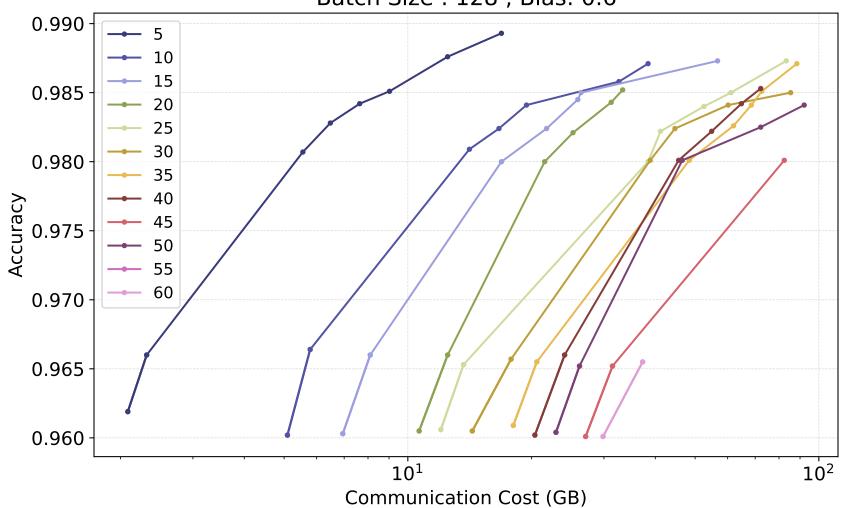




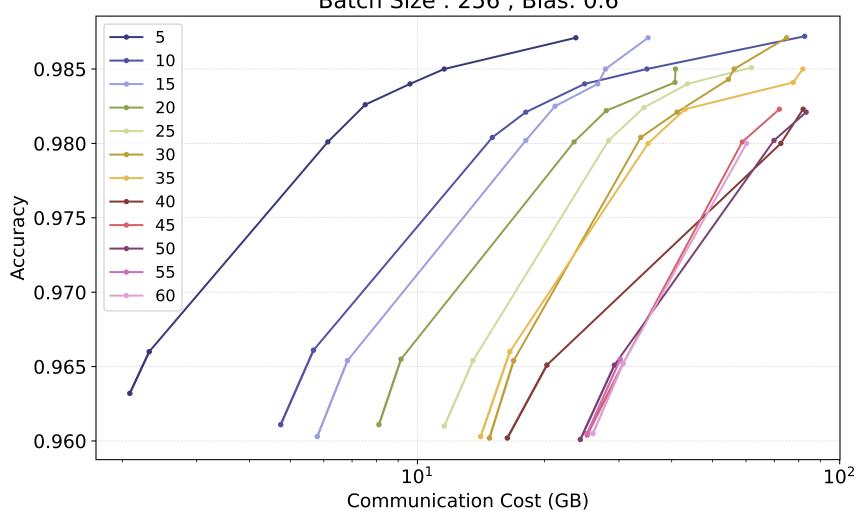
synchronous Batch Size: 32, Bias: 0.6 0.990 10 0.985 30 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10^{1} 10² Communication Cost (GB)



synchronous Batch Size: 128, Bias: 0.6



synchronous Batch Size : 256 , Bias: 0.6



gm Theta: 0.5, Batch Size: 32, Bias: 0.6 0.985 0.980 35 Accuracy 6.00 0.970 60 0.965 0.960 10° Communication Cost (GB)

naive Theta: 0.5, Batch Size: 32, Bias: 0.6 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 100 Communication Cost (GB)

linear Theta: 0.5, Batch Size: 32, Bias: 0.6 0.990 10 0.985 0.980 0.975 50 55 60 0.970 0.965 0.960 100 Communication Cost (GB)

Accuracy

sketch Theta: 0.5, Batch Size: 32, Bias: 0.6 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10⁰ Communication Cost (GB)

gm Theta: 1.5, Batch Size: 32, Bias: 0.6 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)

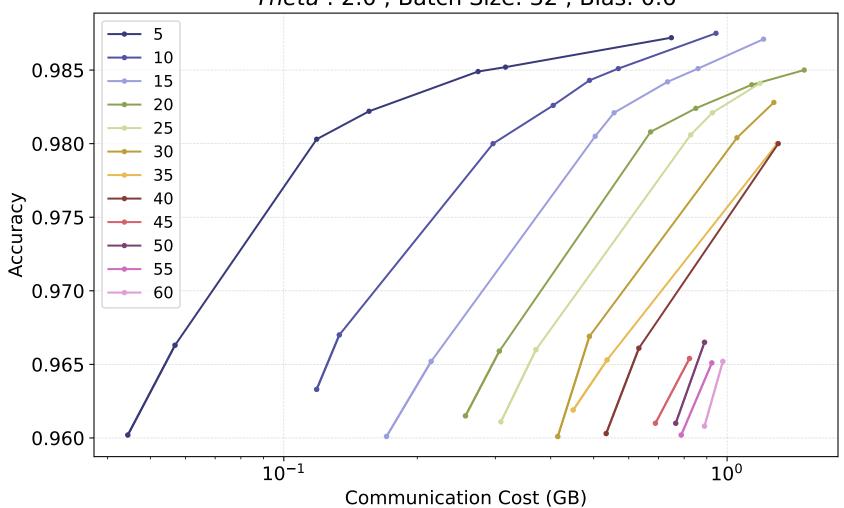
naive Theta: 1.5, Batch Size: 32, Bias: 0.6 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)

linear Theta: 1.5, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 10 15 20 25 0.970 30 35 40 0.965 45 50 55 60 0.960 10^{-1} 100 Communication Cost (GB)

sketch Theta: 1.5, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 2.60 2.60 50 55 0.970 60 0.965 0.960 10^{-1} 100 Communication Cost (GB)

naive Theta: 2.0, Batch Size: 32, Bias: 0.6 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)

linear *Theta*: 2.0, Batch Size: 32, Bias: 0.6



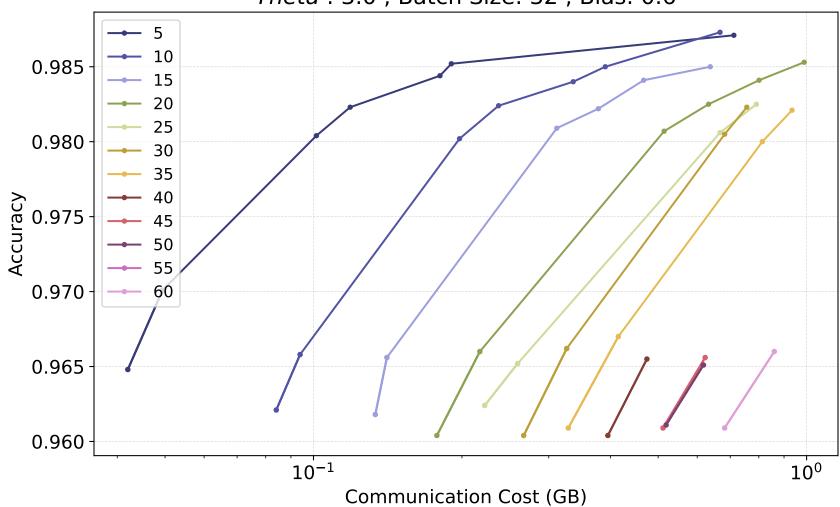
sketch Theta: 2.0, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 6.00 50 55 0.970 60 0.965 0.960 10^{-1} 100 Communication Cost (GB)



naive Theta: 3.0, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 0.052 50 0.970 0.965 0.960 10-1 100 Communication Cost (GB)

linear

Theta: 3.0, Batch Size: 32, Bias: 0.6



sketch Theta: 3.0, Batch Size: 32, Bias: 0.6 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)

gm Theta: 5.0, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 2.60 2.60 50 55 0.970 60 0.965 0.960 10^{-1} 100 Communication Cost (GB)

naive Theta: 5.0, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 2.60 2.60 50 55 0.970 60 0.965 0.960 10^{-1} Communication Cost (GB)

linear Theta: 5.0, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 0.975 50 55 0.970 60 0.965 0.960 10^{-1} Communication Cost (GB)

sketch Theta: 5.0, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 0.75 **-** 10 20 25 0.970 30 - 35 40 0.965 45 50 55 60 0.960 10^{-1} 100 Communication Cost (GB)

gm Theta: 7.0, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 0.975 50 0.970 0.965 0.960 10-1 Communication Cost (GB)

naive Theta: 7.0, Batch Size: 32, Bias: 0.6 10 0.985 **—** 15 20 25 0.980 30 35 Accuracy 40 45 50 55 0.970 **—** 60 0.965 0.960 10^{-1} Communication Cost (GB)

linear Theta: 7.0, Batch Size: 32, Bias: 0.6 0.990 10 0.985 0.980 Accuracy 0.975 50 60 0.970 0.965 0.960 10^{-1} Communication Cost (GB)

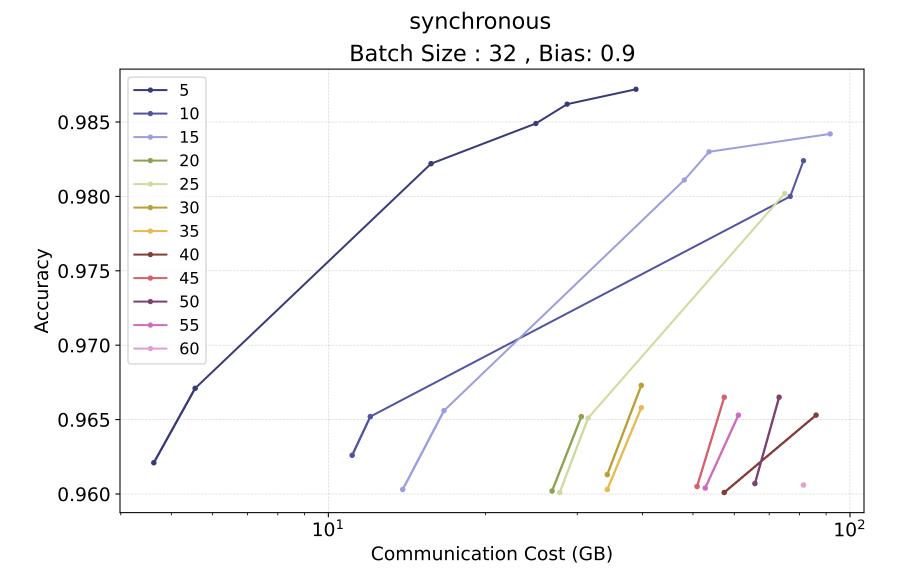
sketch Theta: 7.0, Batch Size: 32, Bias: 0.6 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)

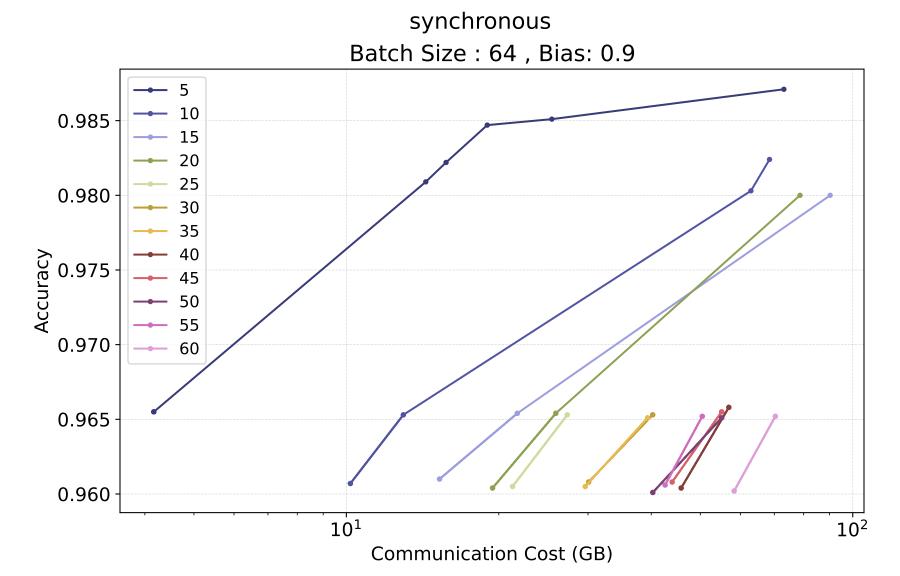
gm Theta: 10.0, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 26.0 26.0 50 0.970 0.965 0.960 10⁻¹ Communication Cost (GB)

naive Theta: 10.0, Batch Size: 32, Bias: 0.6 10 0.985 15 20 25 0.980 30 35 Accuracy 40 45 50 55 0.970 **←** 60 0.965 0.960 10^{-1} Communication Cost (GB)

linear Theta: 10.0, Batch Size: 32, Bias: 0.6 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10-1 Communication Cost (GB)

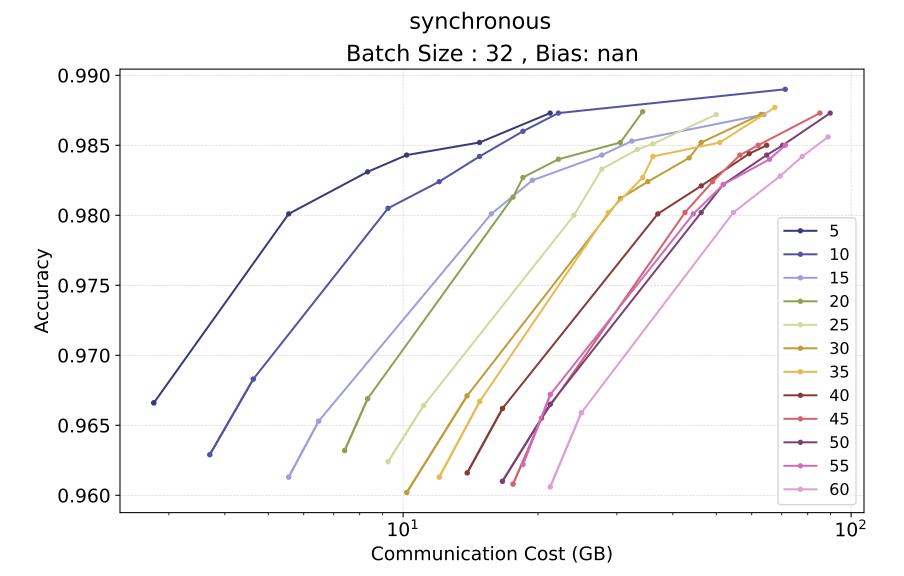
sketch Theta: 10.0, Batch Size: 32, Bias: 0.6 0.985 0.980 Accuracy 0.970 60 0.965 0.960 10 -1 100 Communication Cost (GB)



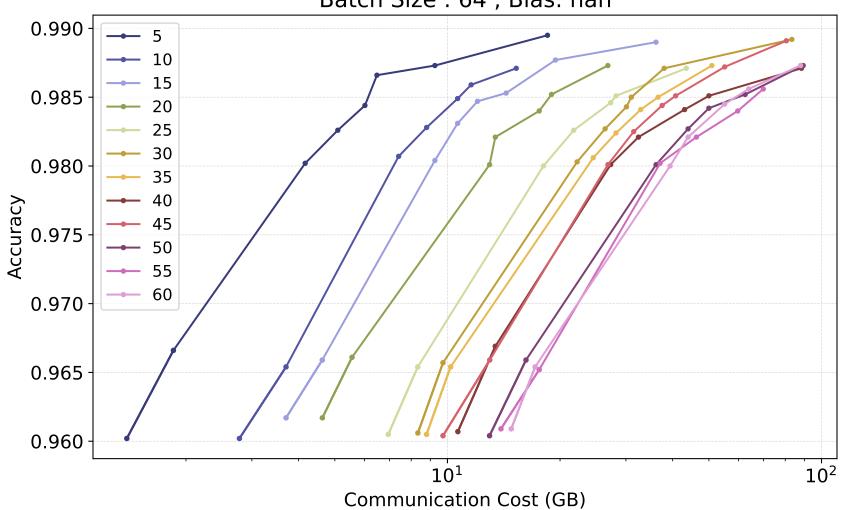


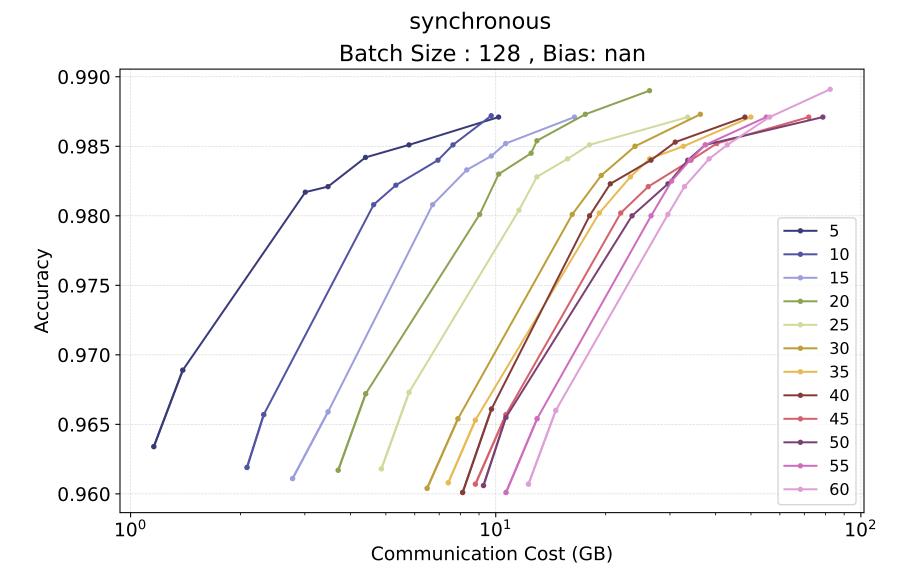
synchronous Batch Size: 128, Bias: 0.9 0.985 0.980 Accuracy 0.970 60 0.965 0.960 10^{1} Communication Cost (GB)

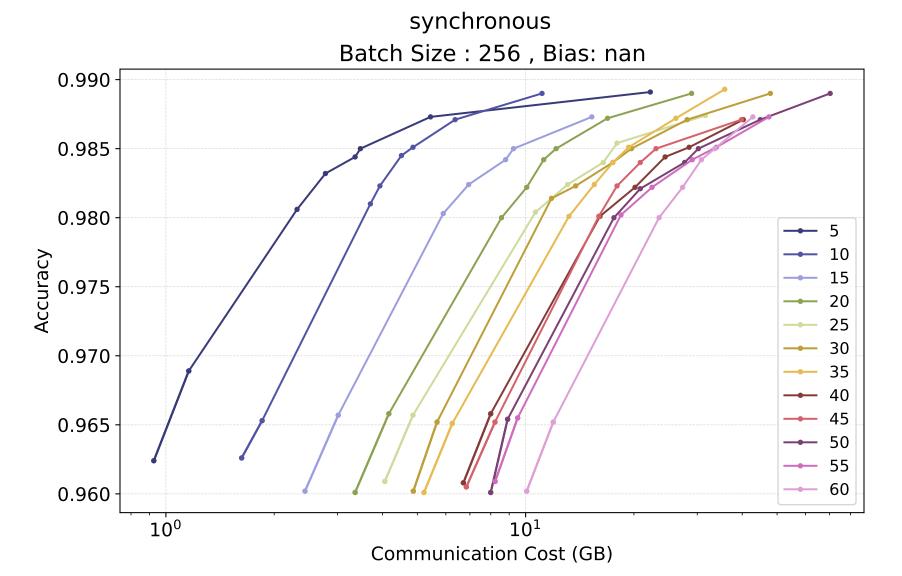
synchronous Batch Size: 256, Bias: 0.9 0.985 10 0.980 Accuracy 0.970 60 0.965 0.960 10^{1} Communication Cost (GB)

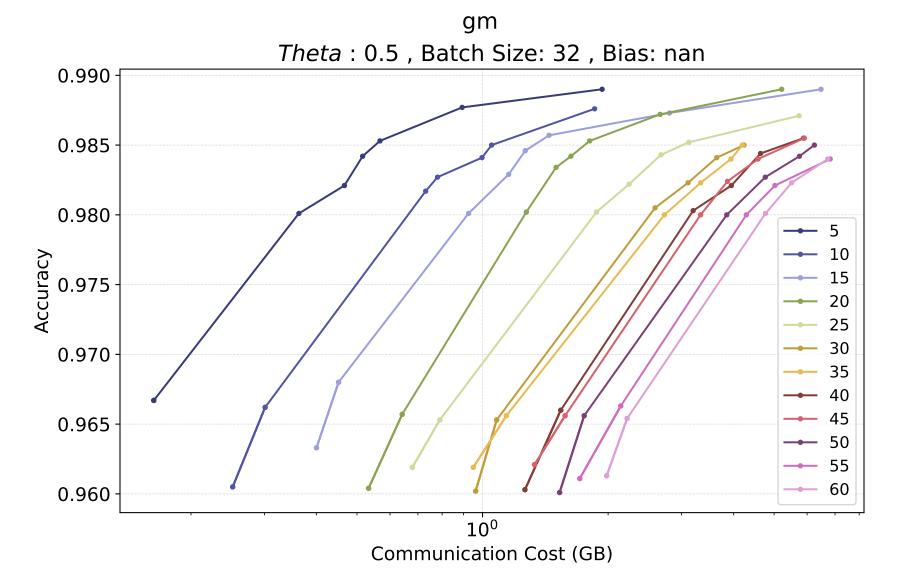


synchronous Batch Size : 64 , Bias: nan









naive Theta: 0.5, Batch Size: 32, Bias: nan 0.990 0.985 0.980 Accuracy 0.975 20 25 30 0.970 35 40 45 0.965 **–** 50 55 ---- 60 0.960 100 Communication Cost (GB)

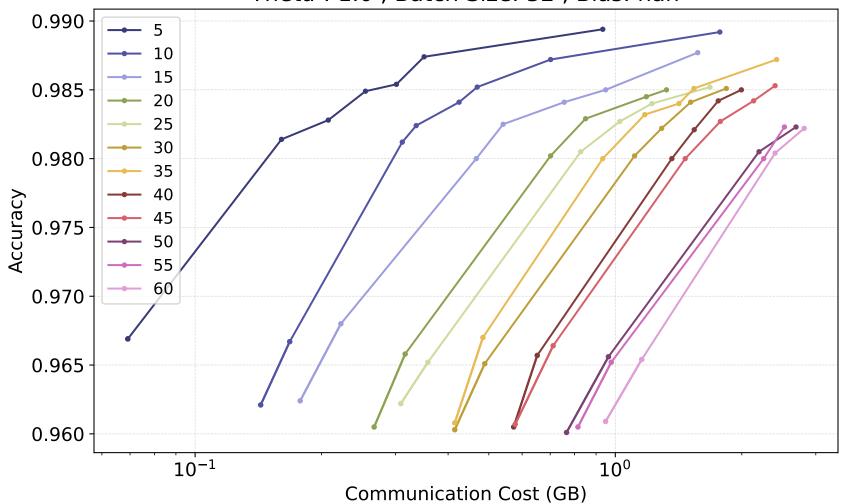
linear Theta: 0.5, Batch Size: 32, Bias: nan 0.990 0.985 0.980 Accuracy 0.975 20 25 30 0.970 35 40 45 0.965 **-** 50 55 ---- 60 0.960 100 Communication Cost (GB)

sketch Theta: 0.5, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 100 Communication Cost (GB)

naive Theta: 1.0, Batch Size: 32, Bias: nan 0.990 0.985 0.980 10 Accuracy 0.975 20 25 30 0.970 35 40 **–** 45 0.965 **-** 50 55 **→** 60 0.960 10^{-1} 100 Communication Cost (GB)

linear

Theta: 1.0, Batch Size: 32, Bias: nan



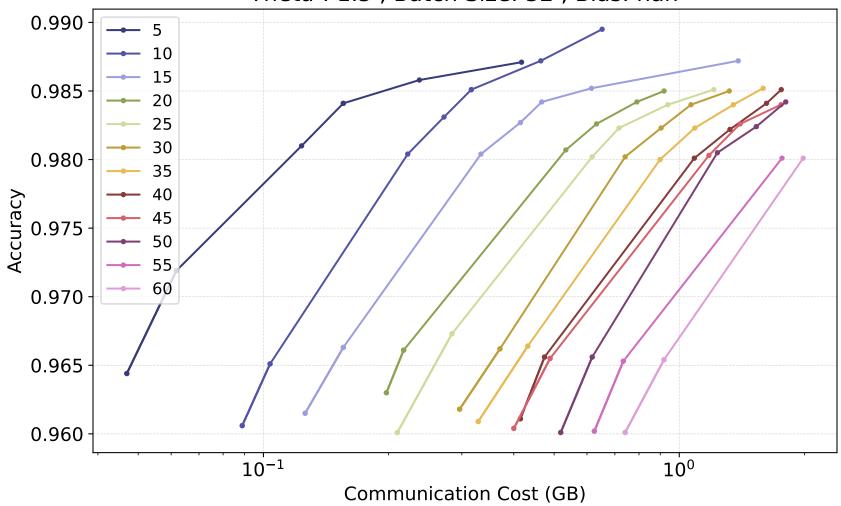
sketch Theta: 1.0, Batch Size: 32, Bias: nan 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10^{-1} 10⁰ Communication Cost (GB)

gm Theta: 1.5, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 Accuracy 0.975 50 60 0.970 0.965 0.960 10⁻¹ 10° Communication Cost (GB)

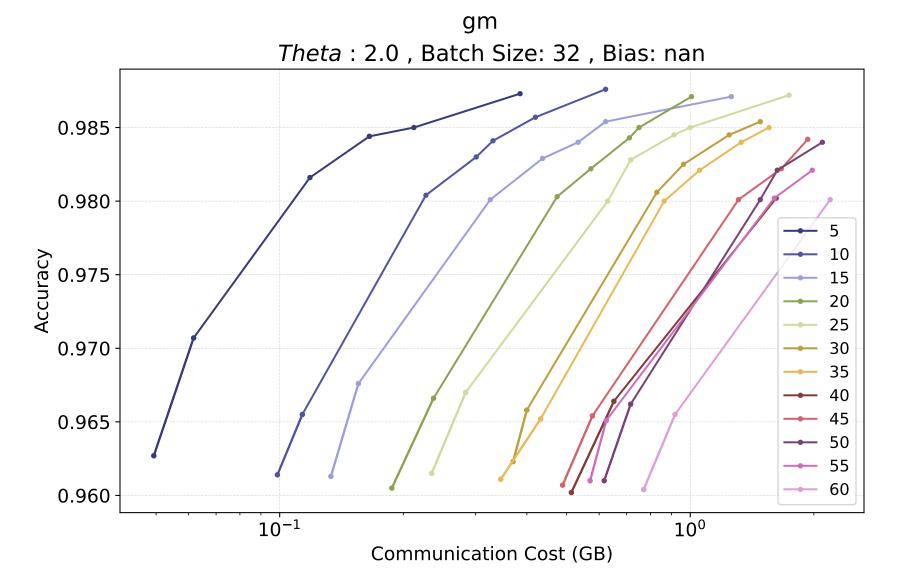
naive Theta: 1.5, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 Accuracy 0.975 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)

linear

Theta: 1.5, Batch Size: 32, Bias: nan



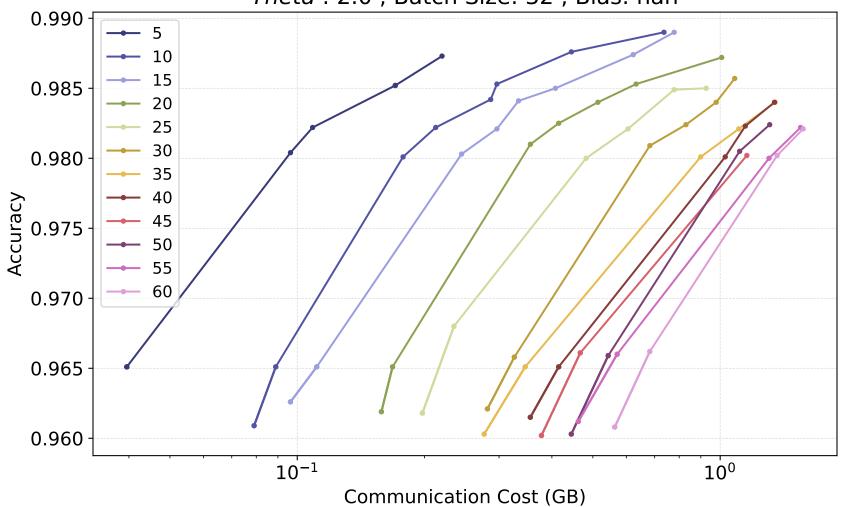
sketch Theta: 1.5, Batch Size: 32, Bias: nan 0.985 0.980 Accuracy 0.975 20 25 30 0.970 **–** 35 40 **-** 45 0.965 **—** 50 55 ---- 60 0.960 10^{-1} 10° Communication Cost (GB)



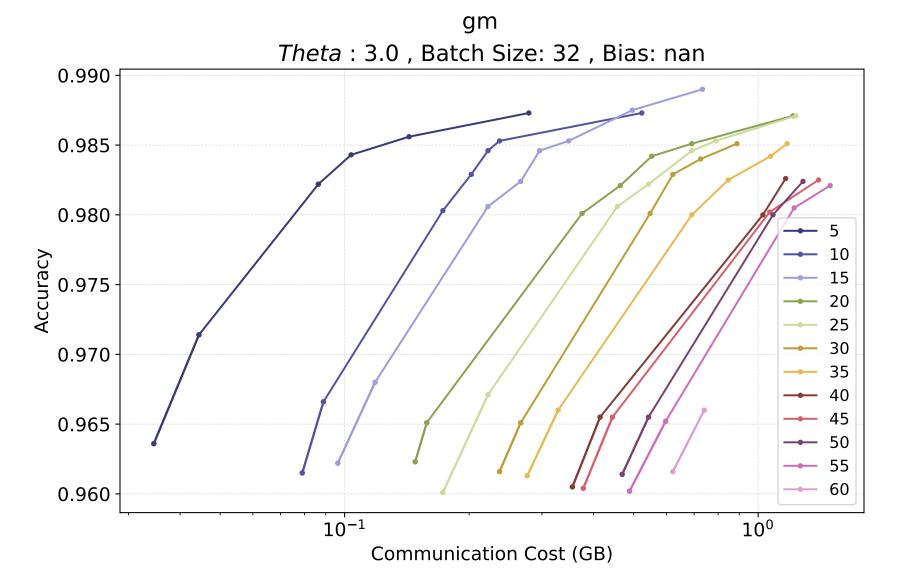
naive Theta: 2.0, Batch Size: 32, Bias: nan 0.990 0.985 0.980 10 Accuracy 0.975 20 25 30 0.970 35 40 **-** 45 0.965 **-** 50 55 **→** 60 0.960 10^{-1} 100 Communication Cost (GB)

linear

Theta: 2.0, Batch Size: 32, Bias: nan



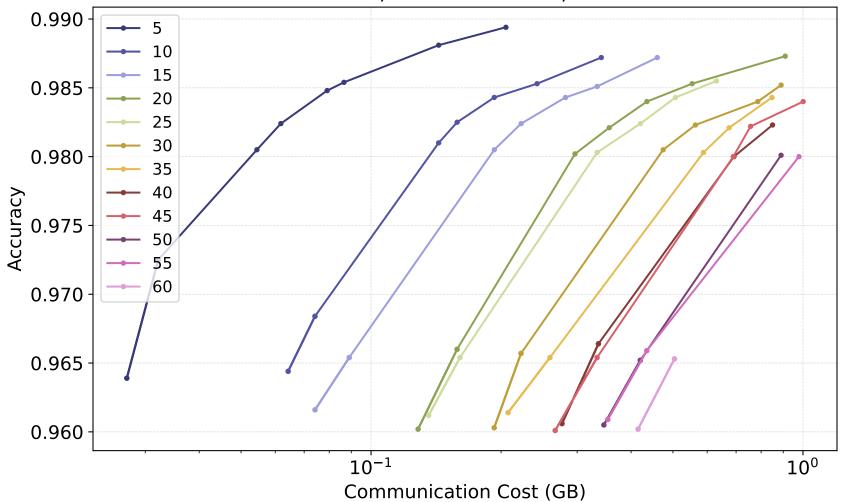
sketch Theta: 2.0, Batch Size: 32, Bias: nan 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10^{-1} 10⁰ Communication Cost (GB)



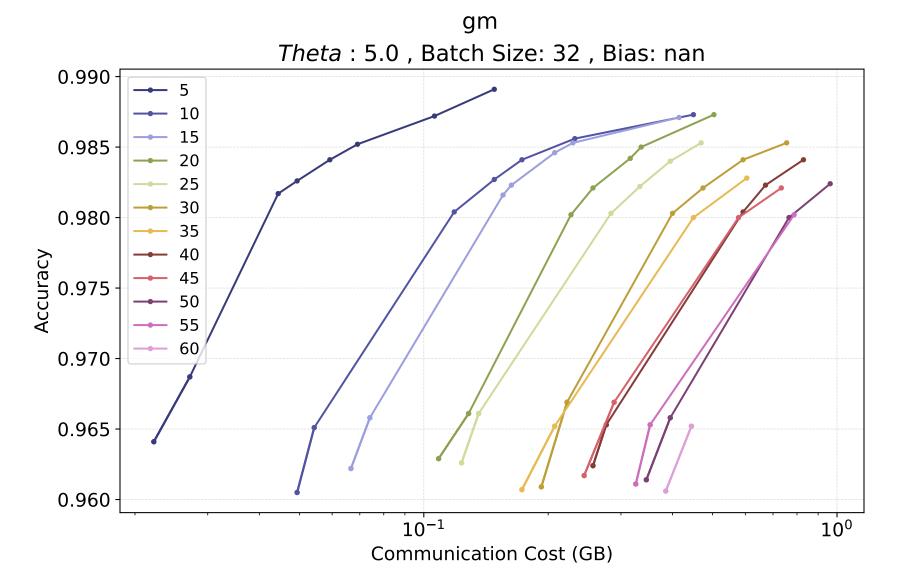
naive Theta: 3.0, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 Accuracy 0.975 50 60 0.970 0.965 0.960 10-1 100 Communication Cost (GB)

linear

Theta: 3.0, Batch Size: 32, Bias: nan



sketch Theta: 3.0, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)



naive Theta: 5.0, Batch Size: 32, Bias: nan 0.985 0.980 Accuracy 2.60 2.60 50 55 0.970 60 0.965 0.960 10^{-1} Communication Cost (GB)

linear Theta: 5.0, Batch Size: 32, Bias: nan 10-1

0.990

0.985

0.980

0.975

0.970

0.965

0.960

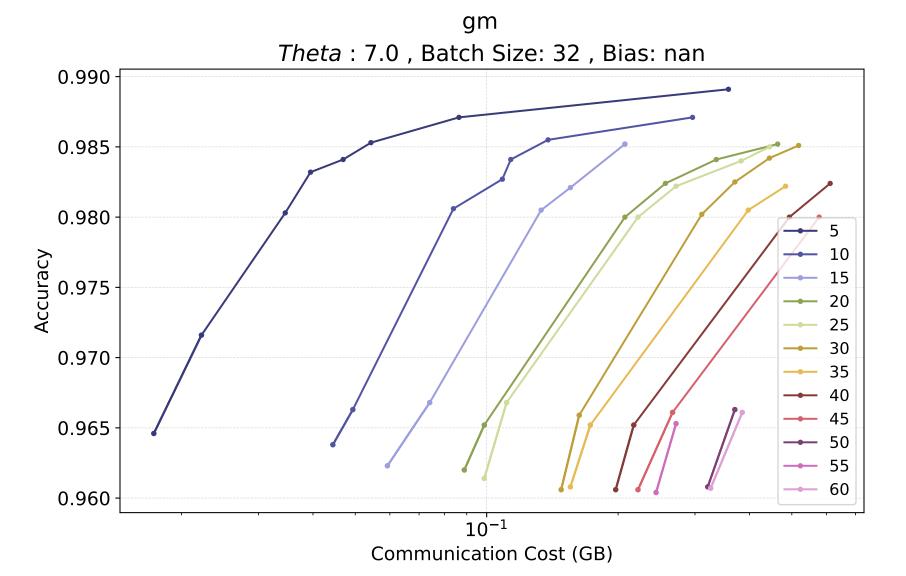
Accuracy

10

50 55 60

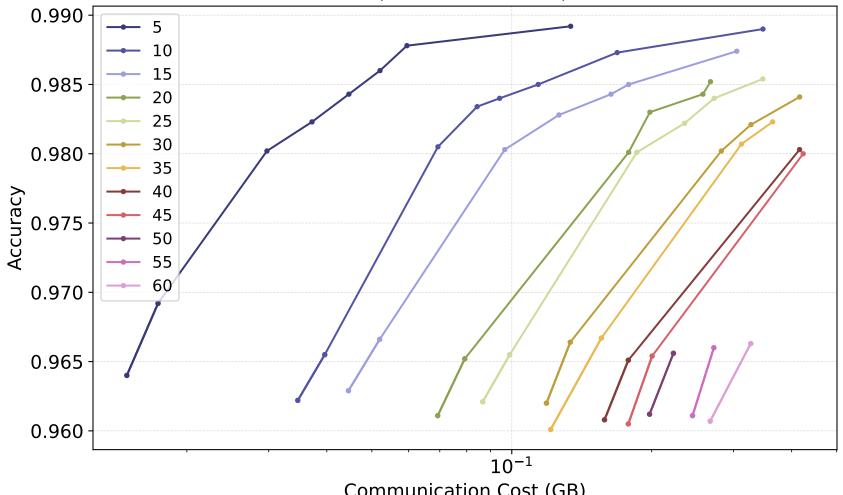
Communication Cost (GB)

sketch Theta: 5.0, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 Accuracy 0.975 55 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)



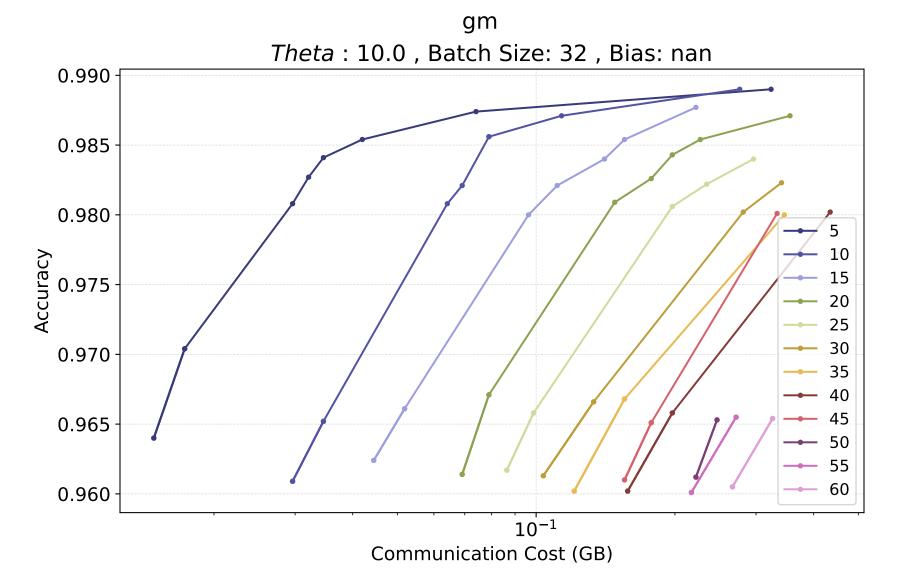
naive Theta: 7.0, Batch Size: 32, Bias: nan 0.985 0.980 Accuracy 0.052 50 0.970 0.965 0.960 10-1 Communication Cost (GB)

linear Theta: 7.0, Batch Size: 32, Bias: nan



Communication Cost (GB)

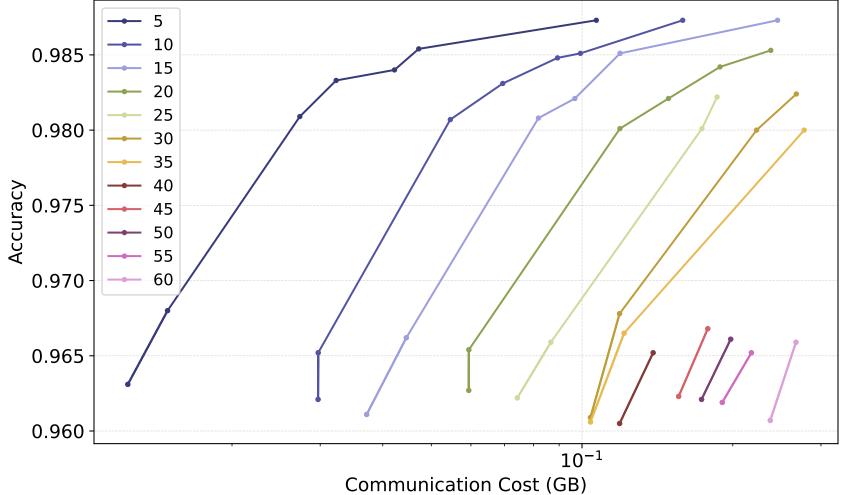
sketch Theta: 7.0, Batch Size: 32, Bias: nan 0.985 0.980 Accuracy 9.00 9.00 50 55 0.970 60 0.965 0.960 10^{-1} 100 Communication Cost (GB)



naive Theta: 10.0, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 Accuracy 0.975 50 55 60 0.970 0.965 0.960 10^{-1} Communication Cost (GB)

linear

Theta: 10.0, Batch Size: 32, Bias: nan



sketch Theta: 10.0, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 Accuracy 0.975 55 60 0.970 0.965 0.960 10^{-1} 100 Communication Cost (GB)

gm Theta: 12.0, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 40 Accuracy 45 0.975 50 60 0.970 0.965 0.960 10^{-1} Communication Cost (GB)

naive Theta: 12.0, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 Accuracy 0.975 50 60 0.970 0.965 0.960 10^{-1} Communication Cost (GB)

linear Theta: 12.0, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 35 40 Accuracy 45 0.975 55 60 0.970 0.965 0.960 10^{-1} Communication Cost (GB)

sketch Theta: 12.0, Batch Size: 32, Bias: nan

0.990

0.985

0.980

0.975

0.970

0.965

0.960

Accuracy

10

55 60

 10^{-1}

Communication Cost (GB)

100

gm Theta: 15.0, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 Accuracy 40 45 0.975 55 60 0.970 0.965 0.960 10^{-1} Communication Cost (GB)

naive Theta: 15.0, Batch Size: 32, Bias: nan 0.990 10 0.985 0.980 35 Accuracy 0.975 60 0.970 0.965 0.960 10^{-1} Communication Cost (GB)

linear Theta: 15.0, Batch Size: 32, Bias: nan 0.990 10 15 0.985 20 25 30 0.980 35 40 Accuracy 45 0.975 50 55 **—** 60 0.970 0.965 0.960 10-1 10^{-2} Communication Cost (GB)

sketch Theta: 15.0, Batch Size: 32, Bias: nan 0.990 0.985 0.980 **-** 10 Accuracy 0.975 20 25 30 0.970 - 35 40 45 0.965 50 55 **6**0 0.960 10^{-1} 100 Communication Cost (GB)