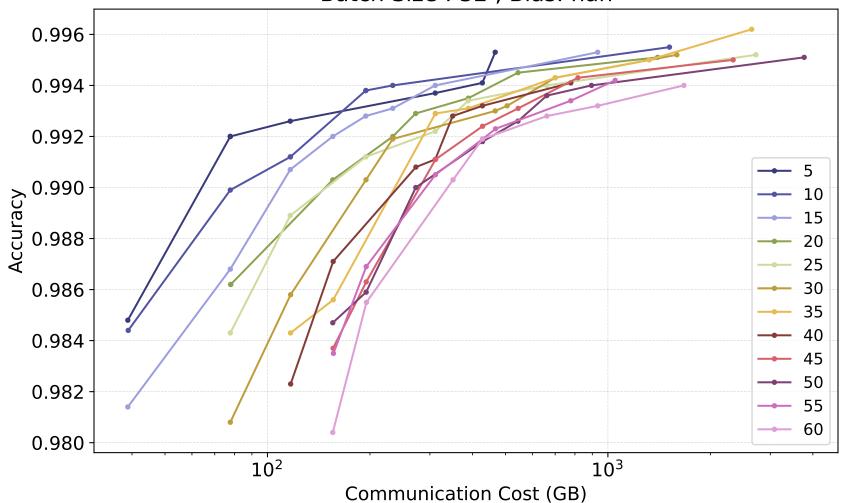
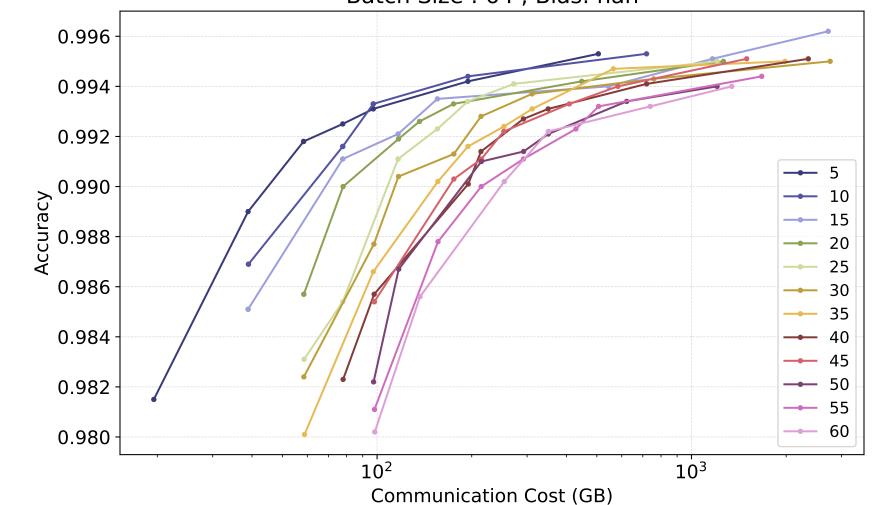
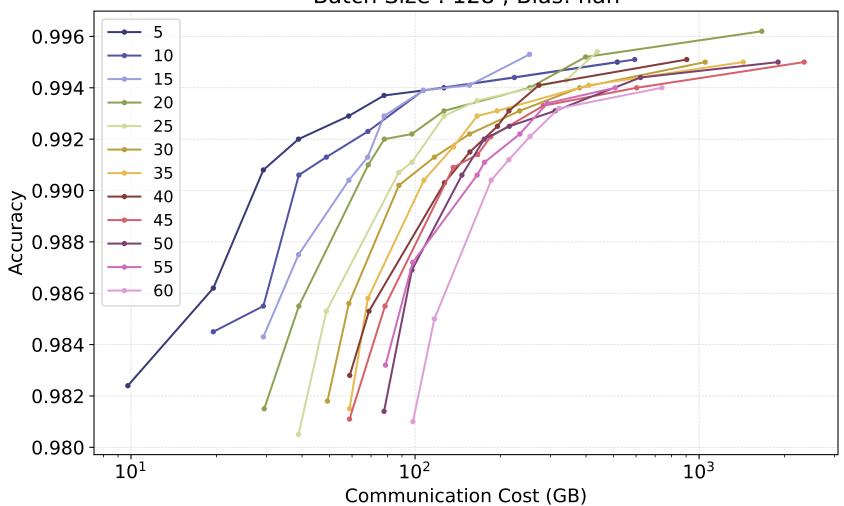
synchronous Batch Size: 32, Bias: nan



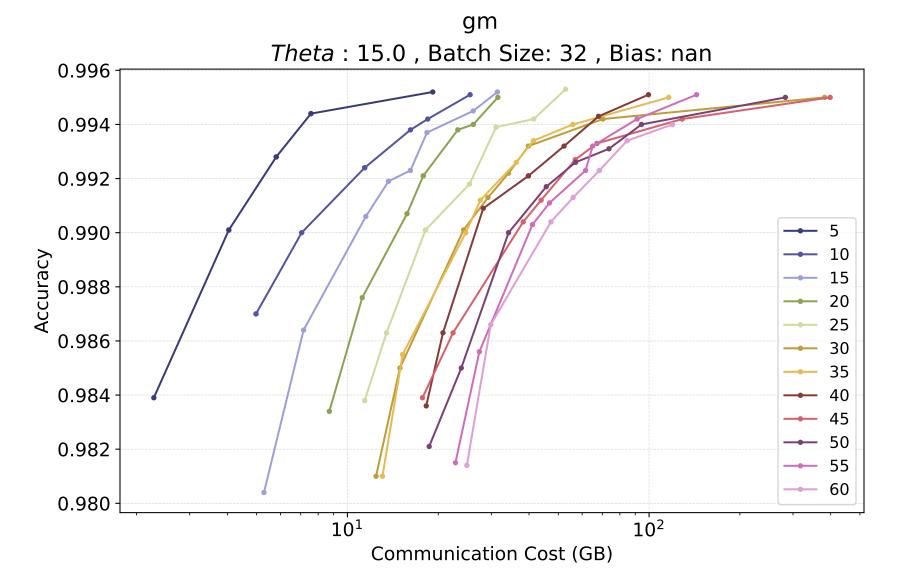
synchronous Batch Size : 64 , Bias: nan



synchronous Batch Size: 128, Bias: nan



synchronous Batch Size: 256, Bias: nan 0.996 0.994 0.992 Accuracy 0 0 0.0 20 25 0.986 30 **→** 35 0.984 40 **-** 45 **-** 50 0.982 55 ---- 60 0.980 10¹ 10² 10³ Communication Cost (GB)



naive Theta: 15.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 Accuracy 886.0 886.0 **—** 10 20 **—** 25 0.986 30 **→** 35 0.984 **-** 45 - 50 0.982 **–** 55 ---- 60 0.980 10^{1} 102 Communication Cost (GB)

linear Theta: 15.0, Batch Size: 32, Bias: nan **-** 10 20 **— 25 —** 30 **—** 35 **—** 50 **→** 55 **→** 60 10^{1} 10²

Communication Cost (GB)

0.996

0.994

0.992

Accuracy 886.0

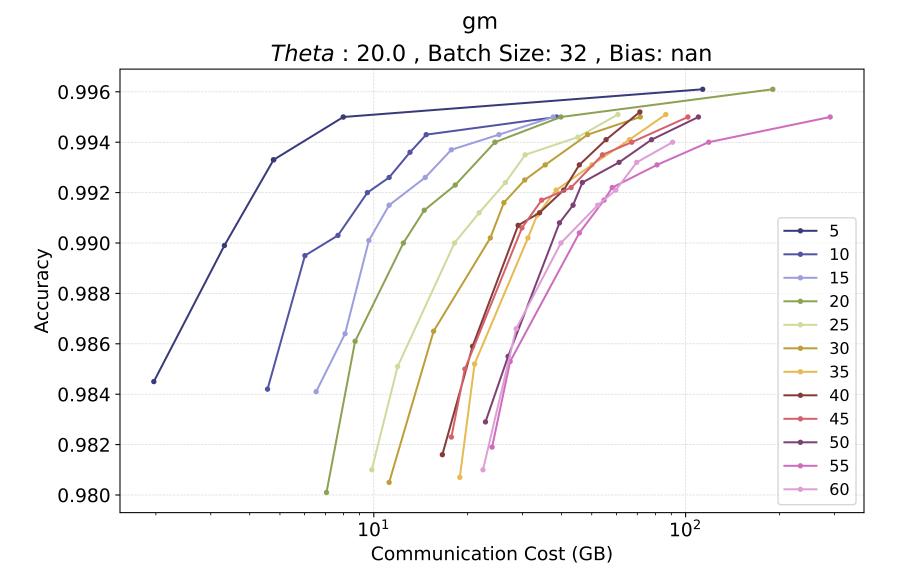
0.986

0.984

0.982

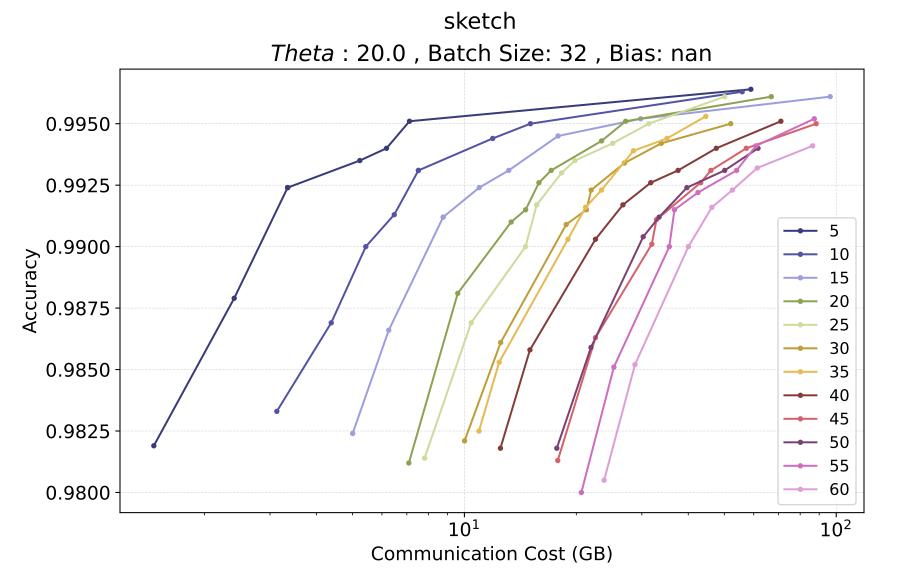
0.980

sketch Theta: 15.0, Batch Size: 32, Bias: nan 0.996 0.9940.992 Accuracy 886.0 - 10 20 25 30 0.986 **—** 35 40 0.984**-** 45 **-** 50 0.982 **-** 55 ---- 60 0.980 10^{1} 10² Communication Cost (GB)



naive Theta: 20.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 Accuracy .0 0:0 **-** 10 20 25 0.986 30 35 0.984 40 **-** 45 - 50 0.982 - 55 ---- 60 0.980 10^{1} 10² Communication Cost (GB)

linear Theta: 20.0, Batch Size: 32, Bias: nan 0.996 0.9940.992 Accuracy 886.0 - 10 20 25 0.986 30 **—** 35 0.984 **-** 45 - 50 0.982 **-** 55 **→** 60 0.980 10^{1} 10² Communication Cost (GB)



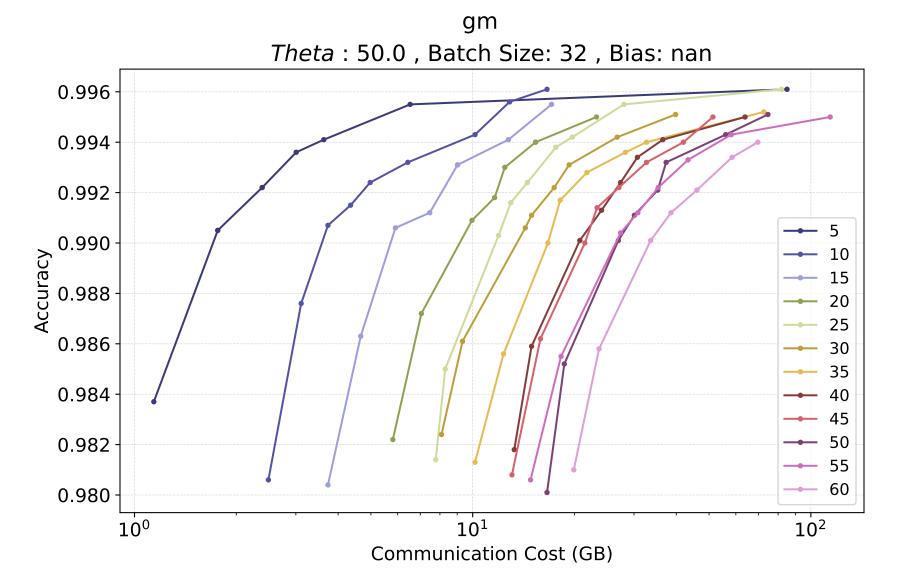
naive Theta: 25.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 Accuracy 0 0 0.0 20 25 0.986 30 **—** 35 0.984 **-** 45 0.982 - 50 **-** 55 ----- 60 0.980 10^{1} 10^2 Communication Cost (GB)

gm Theta: 30.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 Accuracy 886.0 **-** 10 20 25 0.986 30 **—** 35 40 0.984 - 45 **-** 50 0.982 55 0.980 10^{1} 10² Communication Cost (GB)

naive Theta: 30.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 Accuracy .0 0.0 20 25 0.986 30 **—** 35 0.984 0.982 50 **-** 55 ----- 60 0.980 10^{1} 10² Communication Cost (GB)

linear Theta: 30.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 0.990 - 10 Accuracy 0.988 20 25 0.986 30 **—** 35 0.984 **-** 45 0.982 - 50 **–** 55 -- 60 0.980 10^{1} Communication Cost (GB)

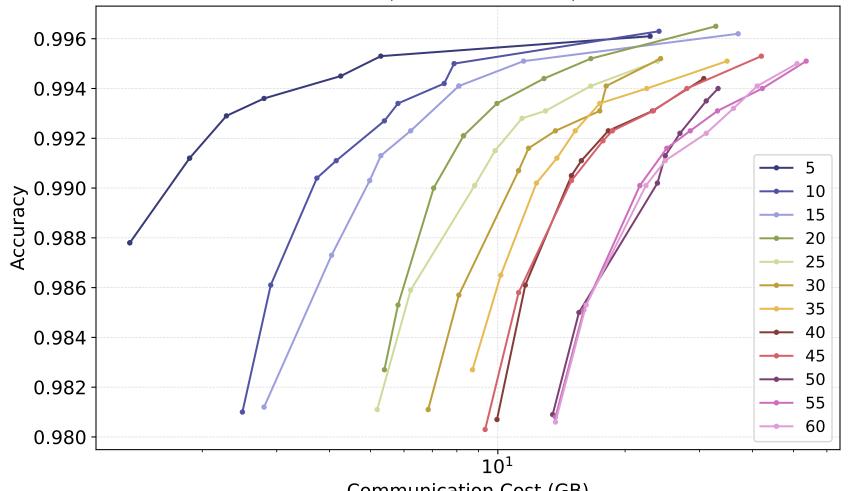
sketch Theta: 30.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 Accuracy 886.0 886.0 **-** 10 20 25 30 0.986 35 40 0.984 **-** 45 **—** 50 0.982 55 ----- 60 10^{1} Communication Cost (GB)



naive Theta: 50.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 Accuracy 886.0 886.0 **—** 10 20 25 0.986 30 **—** 35 40 0.984 - 45 **–** 50 0.982 55 **→** 60 0.980 10^{1} 100 Communication Cost (GB)

linear

Theta: 50.0, Batch Size: 32, Bias: nan



Communication Cost (GB)

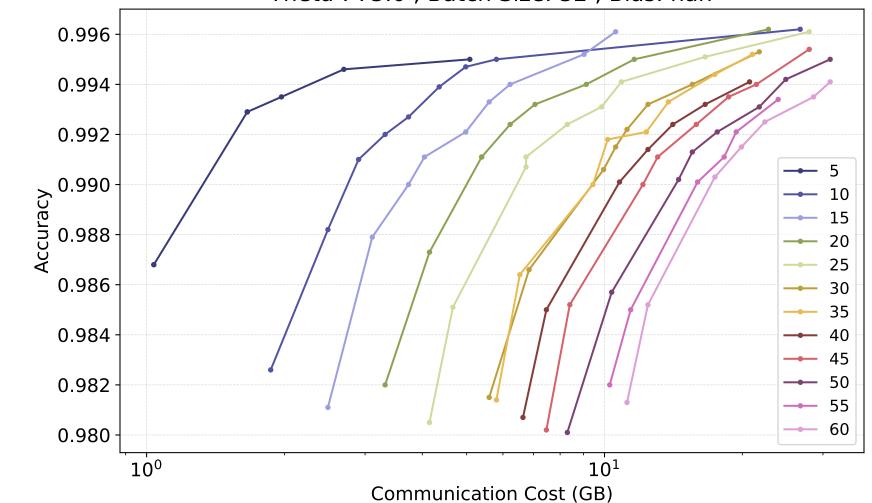
sketch Theta: 50.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 Accuracy 0.0 ر.0 **-** 10 20 25 0.986 30 **—** 35 0.984 40 **-** 45 - 50 0.982 55 ---- 60 0.980 100 10^{1} Communication Cost (GB)

gm Theta: 75.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 Accuracy .0 0:0 20 25 0.986 30 0.984 **-** 45 **-** 50 0.982 **–** 55 ---- 60 0.980 10^1 102 Communication Cost (GB)

naive Theta: 75.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 0.990 Accuracy **—** 10 0.988 20 25 0.986 30 **—** 35 0.984 40 **-** 45 0.982 - 50 **-** 55 -- 60 0.980 100 10^{1} Communication Cost (GB)

linear

Theta: 75.0, Batch Size: 32, Bias: nan



sketch Theta: 75.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 0.990 Accuracy **-** 10 0.98820 25 0.986 30 **—** 35 0.984 40 **-** 45 0.982 **-** 50 **–** 55 -- 60 0.980 100 10^{1} Communication Cost (GB)

gm Theta: 100.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 0.990 Accuracy **-** 10 0.98820 25 0.986 30 - 35 0.984 40 **-** 45 0.982 - 50 **–** 55 ---- 60 0.980 100 10^{1} Communication Cost (GB)

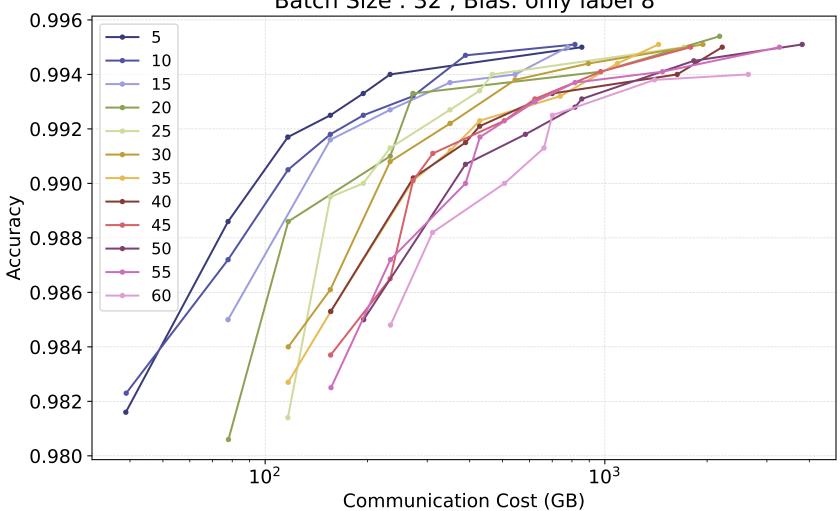
naive Theta: 100.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 0.990 Accuracy 10 0.98820 25 0.986 30 **–** 35 0.984 40 **-** 45 0.982 - 50 **-** 55 -- 60 0.980 100 10^{1} Communication Cost (GB)

linear Theta: 100.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 0.990 10 Accuracy 0.988 20 25 0.986 30 35 0.98440 - 45 0.982 50 55 -- 60 0.980 100 10^{1} Communication Cost (GB)

sketch Theta: 100.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 0.990 Accuracy 10 0.98820 25 0.986 30 **–** 35 0.984 40 - 45 0.982 50 **-** 55 ----- 60 0.980 100 10^{1} Communication Cost (GB)

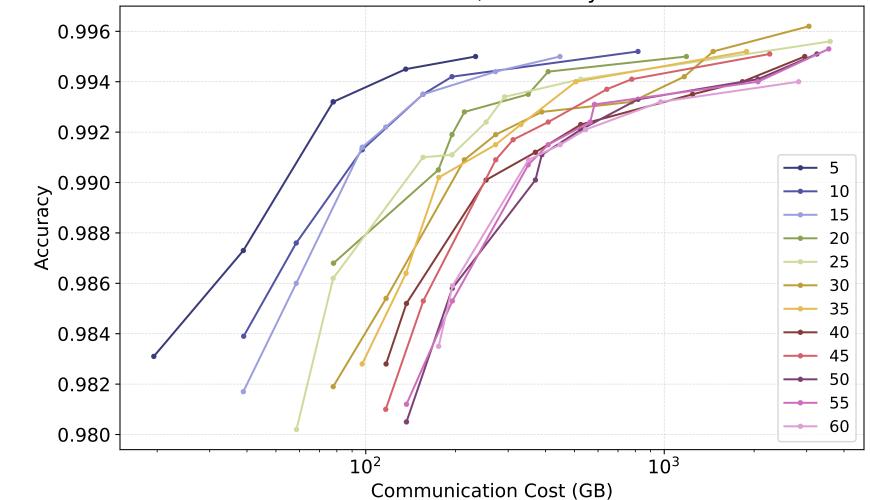
synchronous

Batch Size: 32, Bias: only label 8

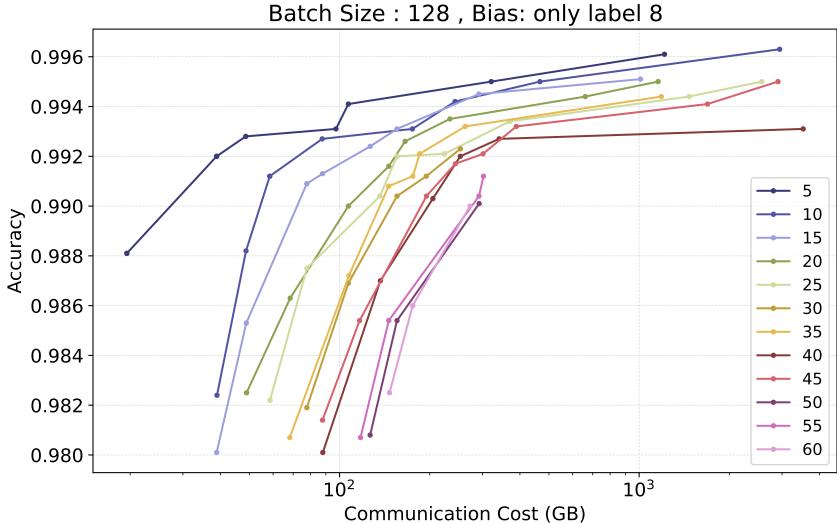


synchronous

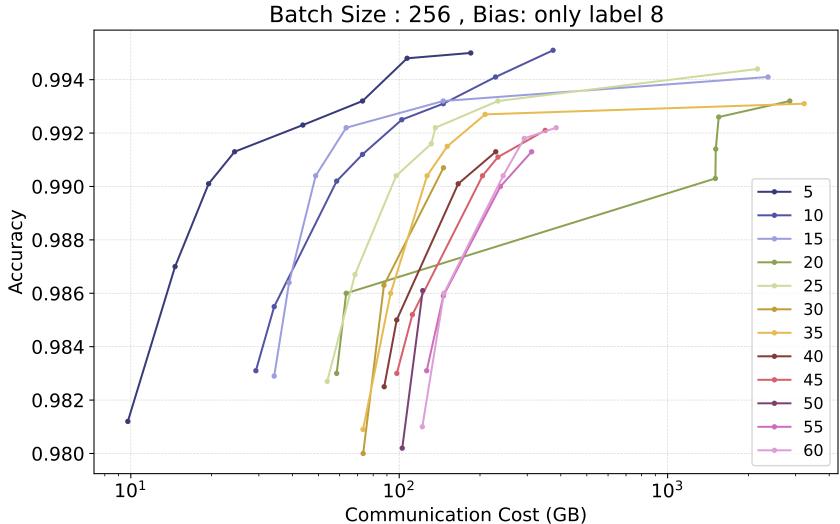
Batch Size: 64, Bias: only label 8



synchronous

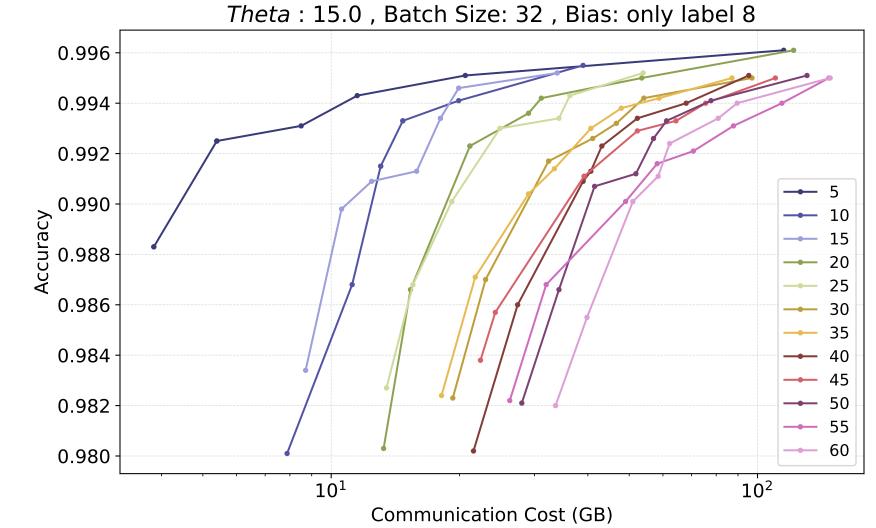


synchronous

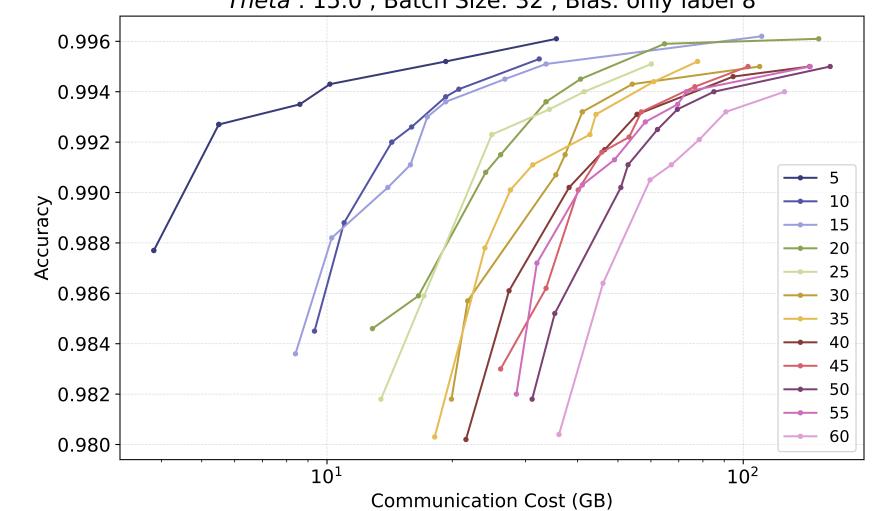


gm Theta: 15.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 Accuracy .0 0.0 20 25 0.986 30 **→** 35 0.984 40 **-** 45 - 50 0.982 55 0.980 10¹ 10² Communication Cost (GB)

naive



linear
Theta: 15.0, Batch Size: 32, Bias: only label 8

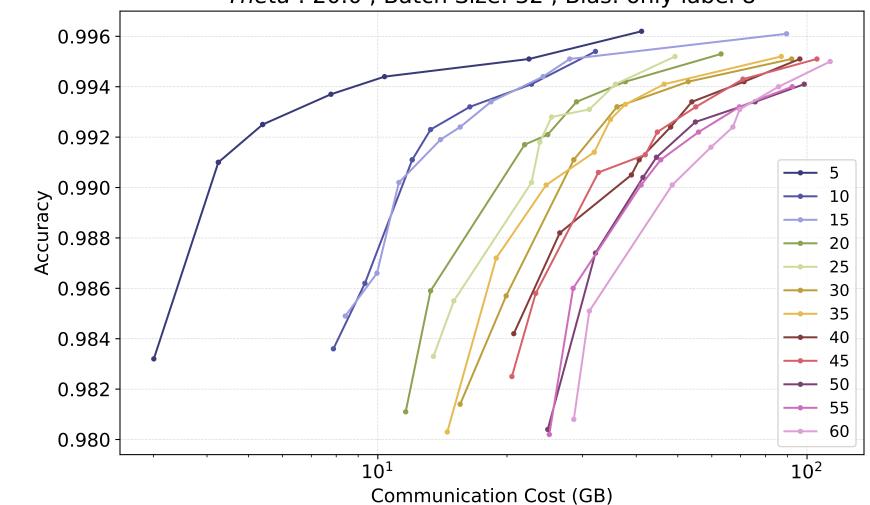


sketch Theta: 15.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 0.990 Accuracy 0.988 20 25 0.986 30 0.984 0.982 - 50 **–** 55 ---- 60 0.980 10^{1} 10^2 Communication Cost (GB)

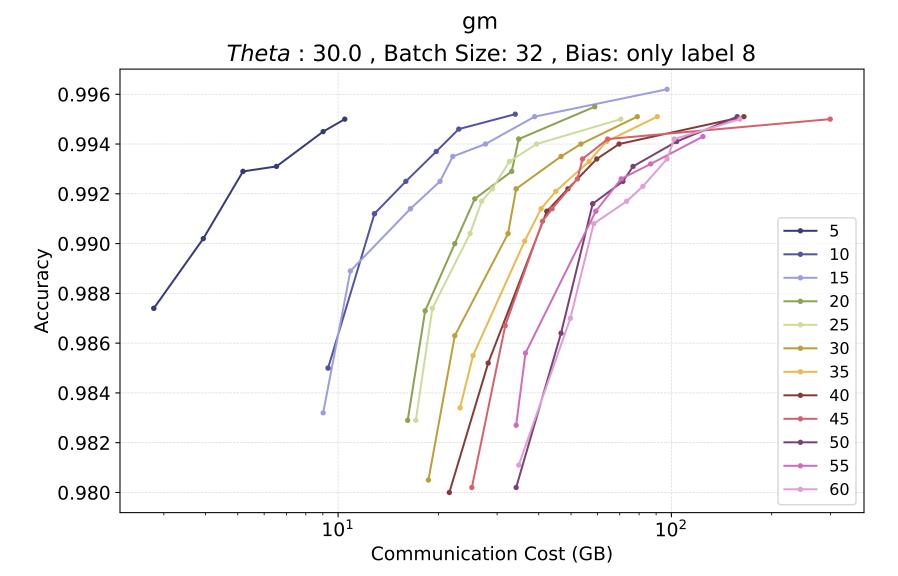
gm Theta: 20.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 Accuracy 886.0 **—** 10 20 25 30 0.986 **—** 35 40 0.984 **-** 45 **-** 50 0.982 55 ---- 60 0.980 10¹ 10² Communication Cost (GB)

naive Theta: 20.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 Accuracy 886.0 - 10 20 25 **–** 30 0.986 0.984 **—** 50 0.982 55 **60** 0.980 10^{1} 10² Communication Cost (GB)

linear *Theta*: 20.0, Batch Size: 32, Bias: only label 8



sketch Theta: 20.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 Accuracy 0 0 0.0 20 25 0.986 30 0.984 **-** 45 0.982 - 50 **–** 55 **60** 0.980 10^{1} 10² Communication Cost (GB)

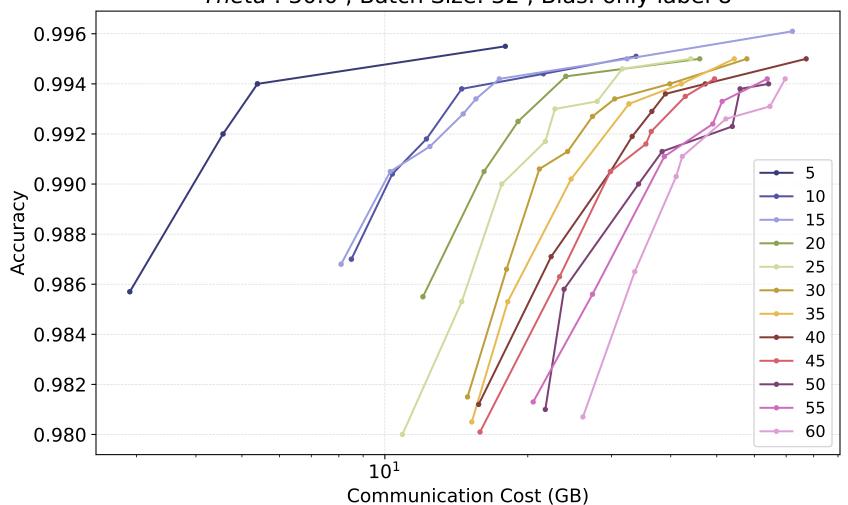


naive Theta: 30.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 Accuracy 886.0 - 10 20 25 30 0.986 **–** 35 0.984 **-** 45 **–** 50 0.982 55 -- 60 10² 10^{1}

Communication Cost (GB)

linear

Theta: 30.0, Batch Size: 32, Bias: only label 8



sketch Theta: 30.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 Accuracy 886.0 **—** 10 20 25 0.986 30 **—** 35 0.984**-** 45 **-** 50 0.982 **-** 55 **→** 60 0.980 10^{1} 10² Communication Cost (GB)

gm Theta: 50.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 Accuracy 0 0 0.0 20 25 0.986 30 **—** 35 0.984 **-** 45 0.982 - 50 **-** 55 -- 60 0.980 10^{1} 10² Communication Cost (GB)

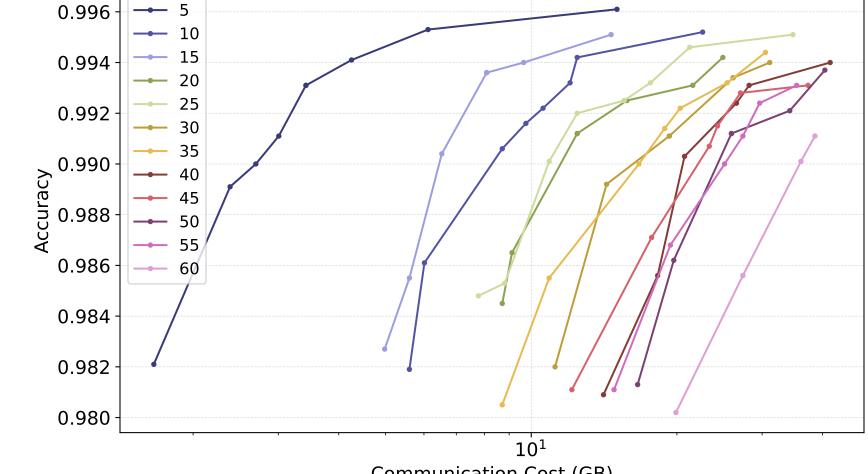
naive Theta: 50.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 0.990 Accuracy **-** 10 0.988 20 25 0.986 30 **—** 35 0.984 40 **-** 45 0.982 - 50 **-** 55 -- 60 0.980 10^{1} Communication Cost (GB)

linear Theta: 50.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 Accuracy .0 0:0 **—** 10 20 25 0.986 30 **—** 35 0.984 40 **-** 45 - 50 0.982 **-** 55 ---- 60 0.980 10^{1} Communication Cost (GB)

sketch Theta: 50.0, Batch Size: 32, Bias: only label 8 0.996 0.9940.992 Accuracy 886.0 886.0 **—** 10 20 25 30 0.986 **—** 35 40 0.984- 45 **–** 50 0.982 **-** 55 ---- 60 0.980 10^{1} Communication Cost (GB)

gm Theta: 75.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 Accuracy 886.0 886.0 **-** 10 20 25 30 0.986 0.984 **-** 45 **–** 50 0.982 55 **60** 10^{1} 10² Communication Cost (GB)

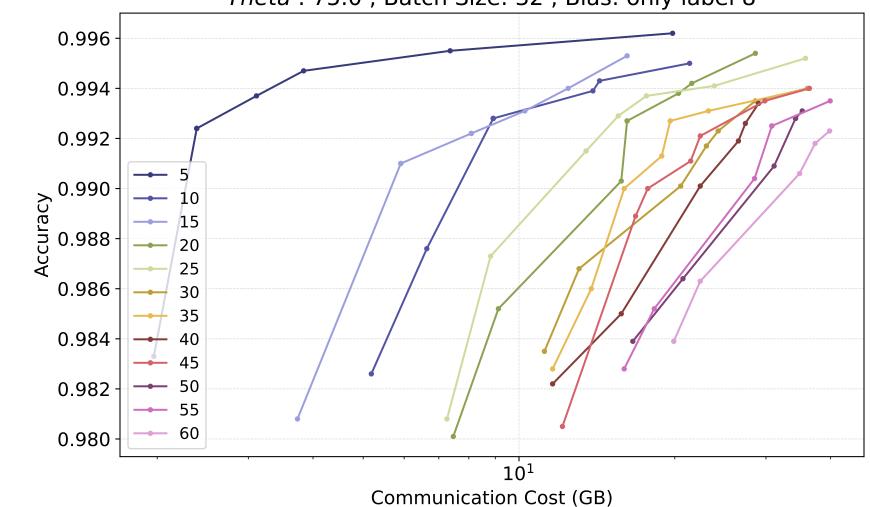
naive Theta: 75.0, Batch Size: 32, Bias: only label 8

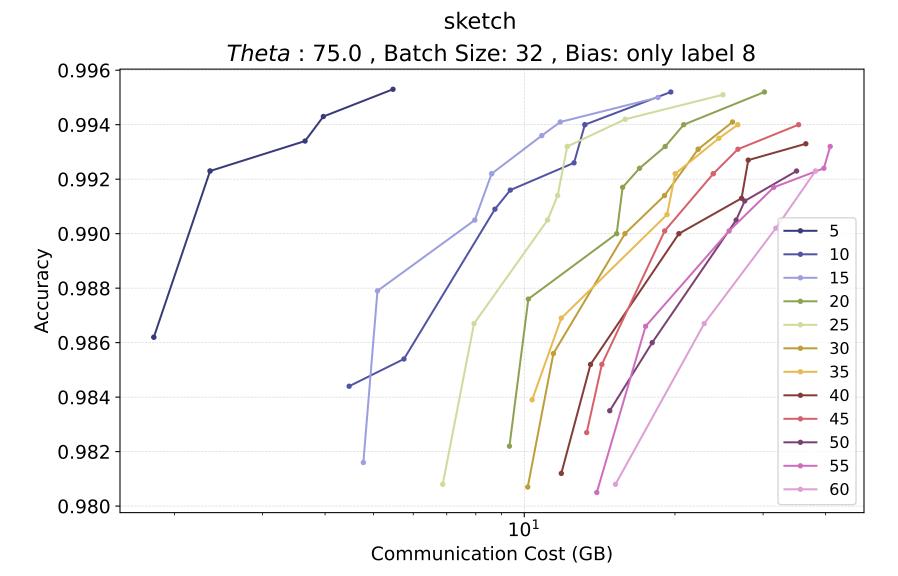


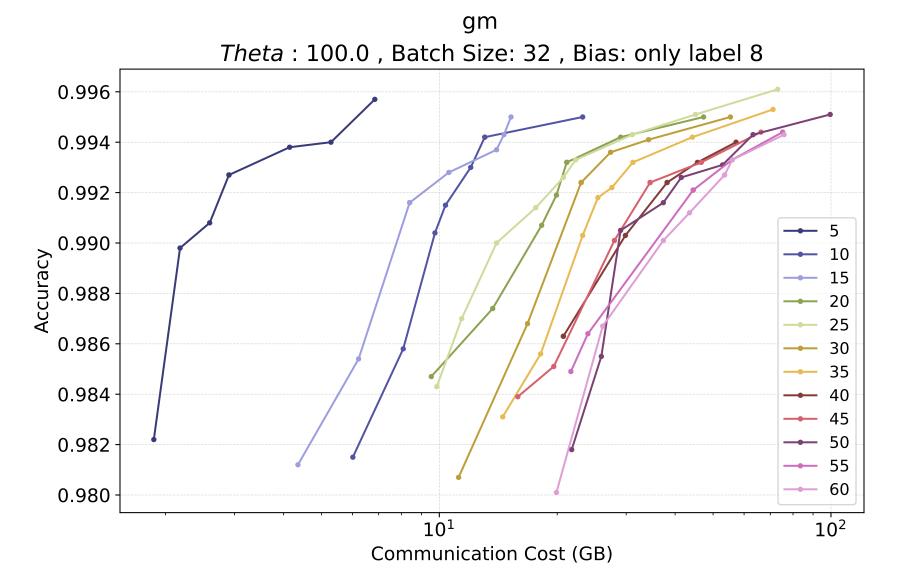
Communication Cost (GB)

linear

Theta: 75.0, Batch Size: 32, Bias: only label 8







naive Theta: 100.0, Batch Size: 32, Bias: only label 8 25 30 35 45 50 55 60

0.994

0.992

0.990

0.988

0.986

0.984

0.982

Accuracy

Communication Cost (GB)

 10^{1}

linear Theta: 100.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 25 30 0.990 35 Accuracy 0.988 50 55 0.986 60 0.984 0.982 0.980 10^{1} Communication Cost (GB)

sketch Theta: 100.0, Batch Size: 32, Bias: only label 8 10 20 25 30 35 40 45 55 60 10^{1}

0.996

0.994

0.992

0.986

0.984

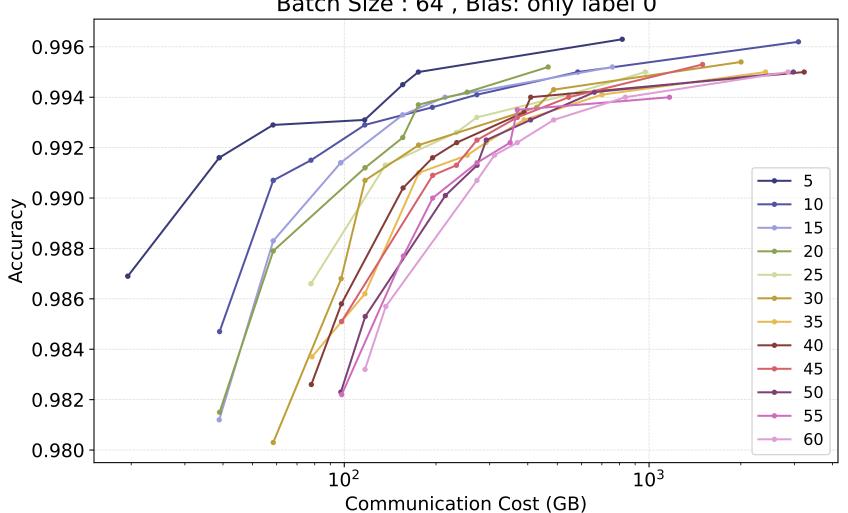
0.982

0.980

Communication Cost (GB)

synchronous Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 Accuracy 886.0 10 20 25 30 0.986 35 40 0.984 45 **–** 50 0.982 55 ----- 60 0.980 102 10³ Communication Cost (GB)

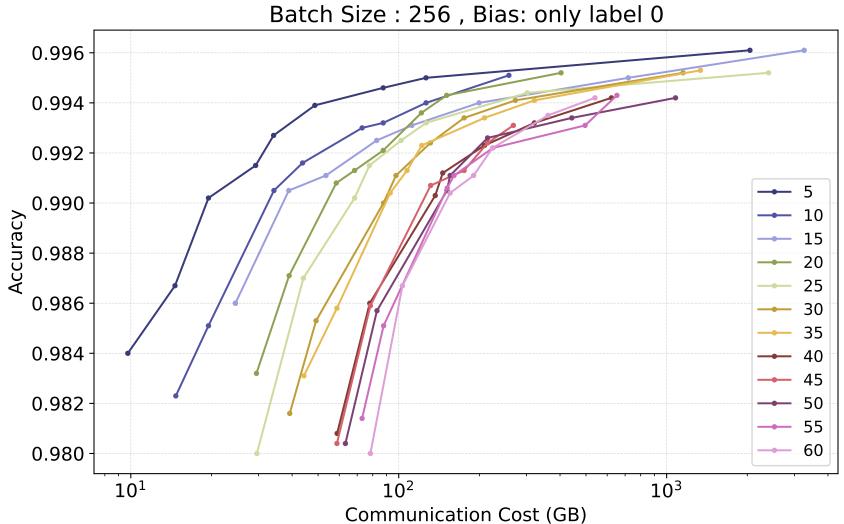
synchronous Batch Size : 64 , Bias: only label 0



synchronous Batch Size: 128, Bias: only label 0 0.996 0.994 0.992 Accuracy 886.0 - 10 20 25 30 0.986 35 0.984 **-** 45 **—** 50 0.982 55 **→** 60 102 10³

Communication Cost (GB)

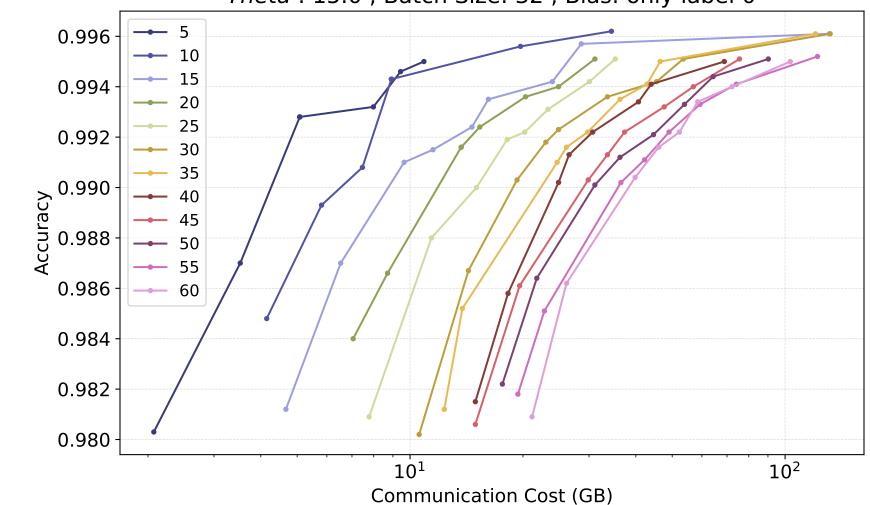
synchronous



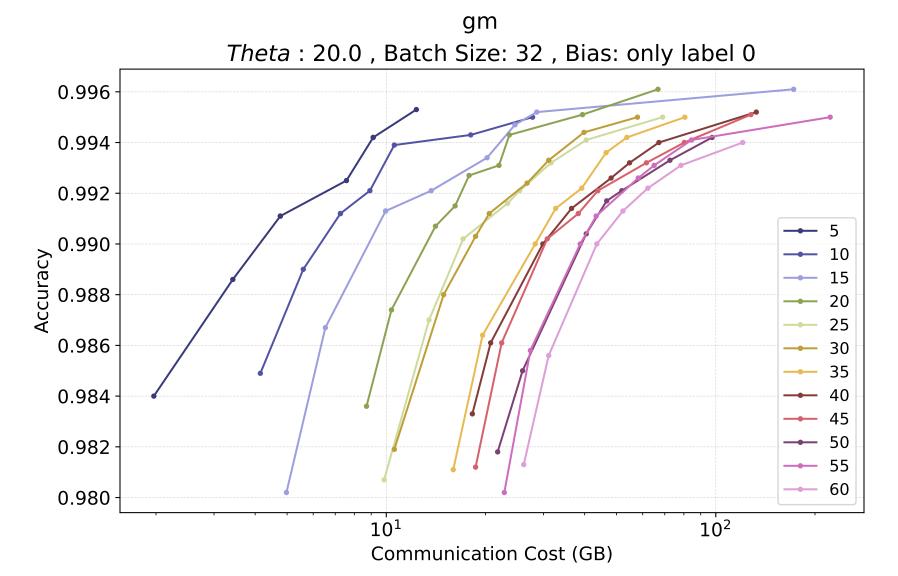
gm Theta: 15.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 Accuracy 886.0 886.0 **-** 10 20 25 0.986 30 **—** 35 40 0.984 **-** 45 **-** 50 0.982 55 ---- 60 0.980 10¹ 10² Communication Cost (GB)

naive Theta: 15.0, Batch Size: 32, Bias: only label 0 0.996 0.9940.992 Accuracy 0 0 0 66 20 25 0.986 30 **—** 35 0.984 **-** 45 - 50 0.982 **–** 55 **→** 60 0.980 10^{1} 10² Communication Cost (GB)

linear
Theta: 15.0, Batch Size: 32, Bias: only label 0



sketch Theta: 15.0, Batch Size: 32, Bias: only label 0 0.996 0.9940.992 Accuracy 0 0 0 0 20 25 0.986 30 **—** 35 0.984 **-** 45 - 50 0.982 **-** 55 ---- 60 0.980 101 102 Communication Cost (GB)



naive Theta: 20.0, Batch Size: 32, Bias: only label 0 20 25 30 **—** 35 40 **-** 45 50 **-** 55 ---- 60

10²

 10^1

Communication Cost (GB)

0.996

0.994

0.992

0.986

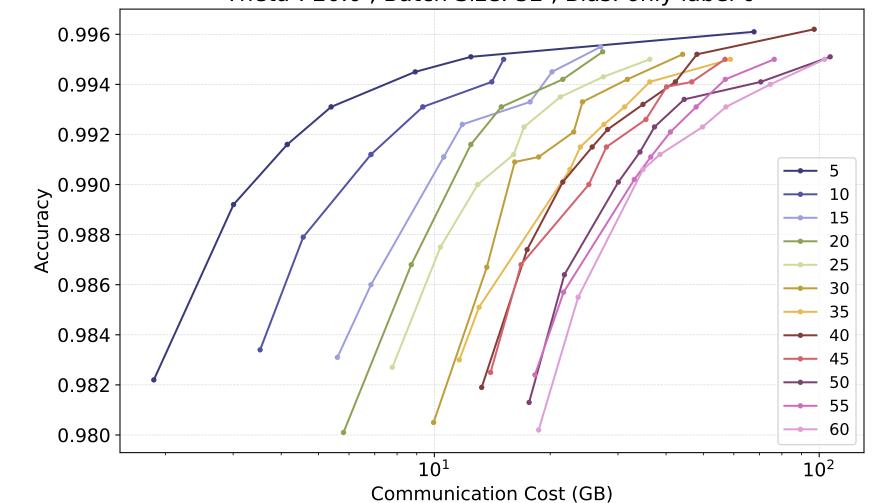
0.984

0.982

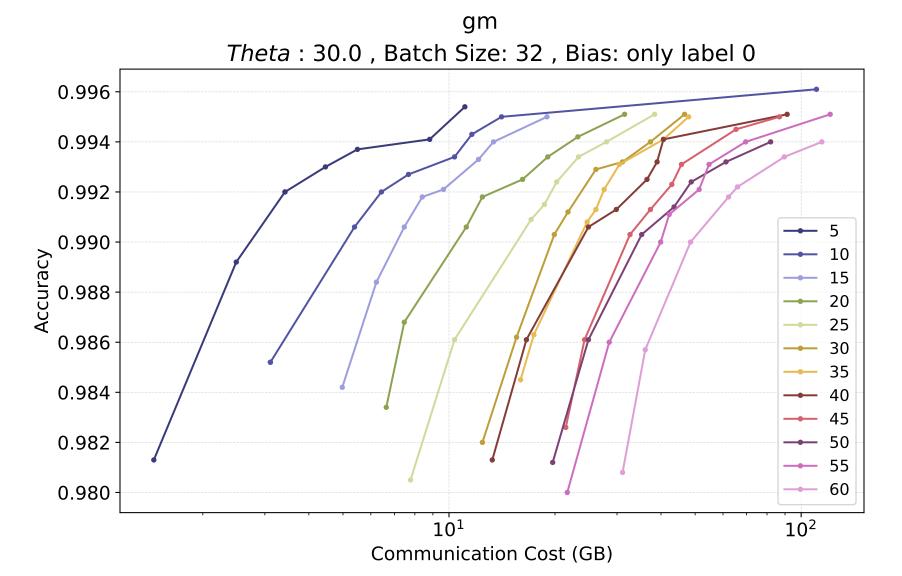
0.980

Accuracy 0 0 0.0

linear
Theta: 20.0, Batch Size: 32, Bias: only label 0

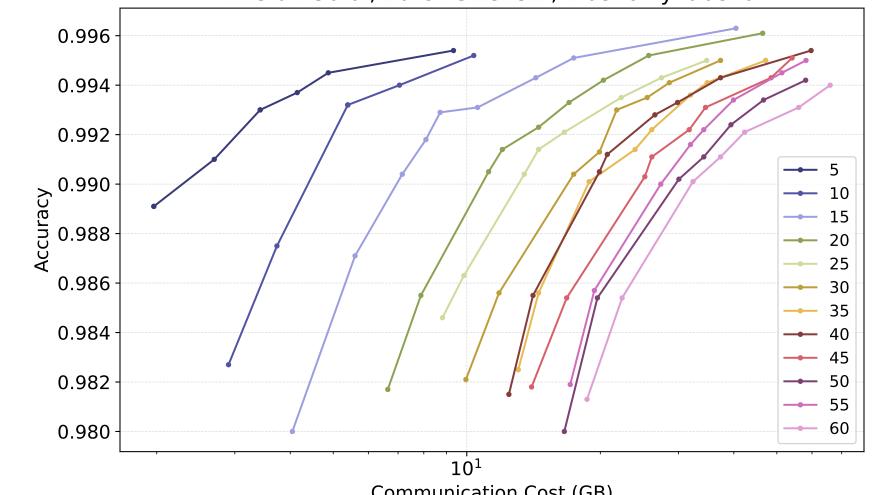


sketch Theta: 20.0, Batch Size: 32, Bias: only label 0 0.996 0.9940.992 Accuracy 0 0 0.0 20 25 0.986 30 - 35 0.984 **-** 45 - 50 0.982 **-** 55 ---- 60 0.980 10¹ 10² Communication Cost (GB)



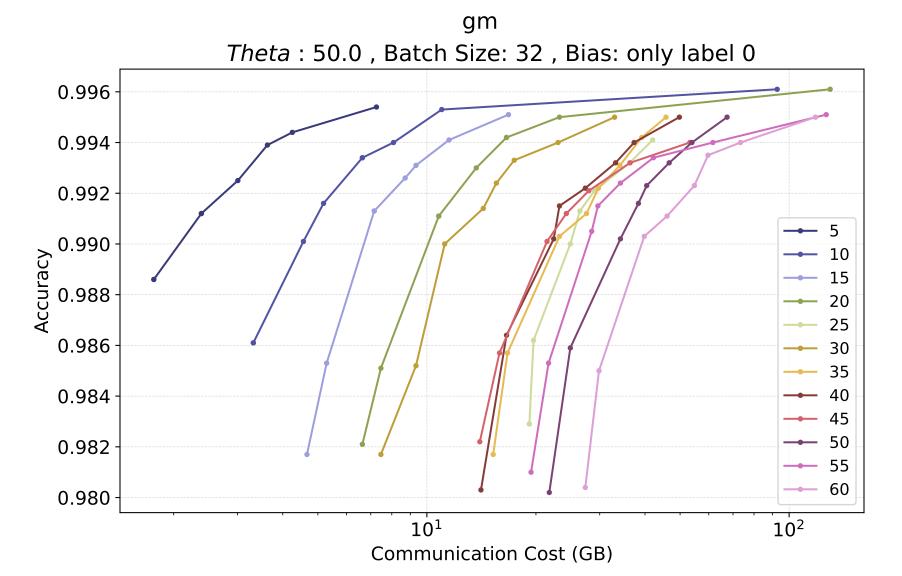
naive Theta: 30.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 Accuracy 886.0 **—** 10 20 **—** 25 0.986 30 **→** 35 0.984**-** 45 **—** 50 0.982 **–** 55 ---- 60 0.980 10^{1} Communication Cost (GB)

linear Theta: 30.0, Batch Size: 32, Bias: only label 0



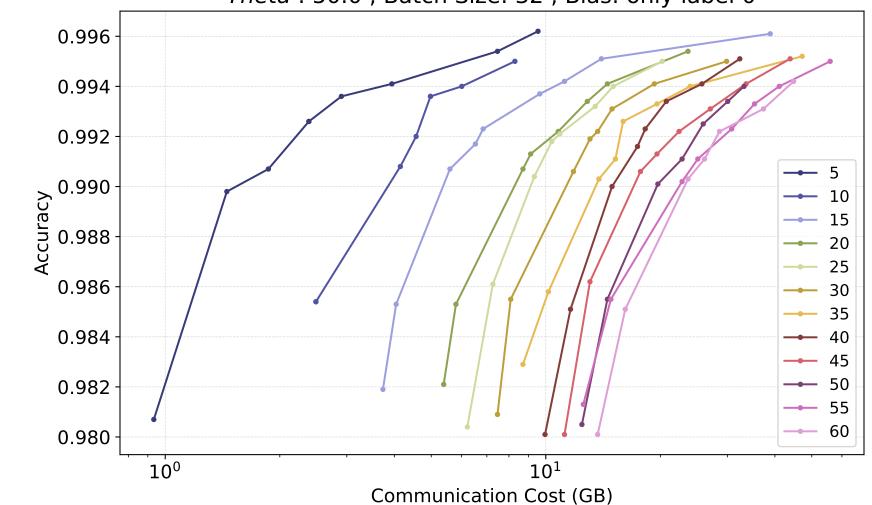
Communication Cost (GB)

sketch Theta: 30.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 0.990 Accuracy - 10 0.988 20 25 0.986 30 0.984- 45 0.982 50 55 -- 60 0.980 10^{1} 10² Communication Cost (GB)

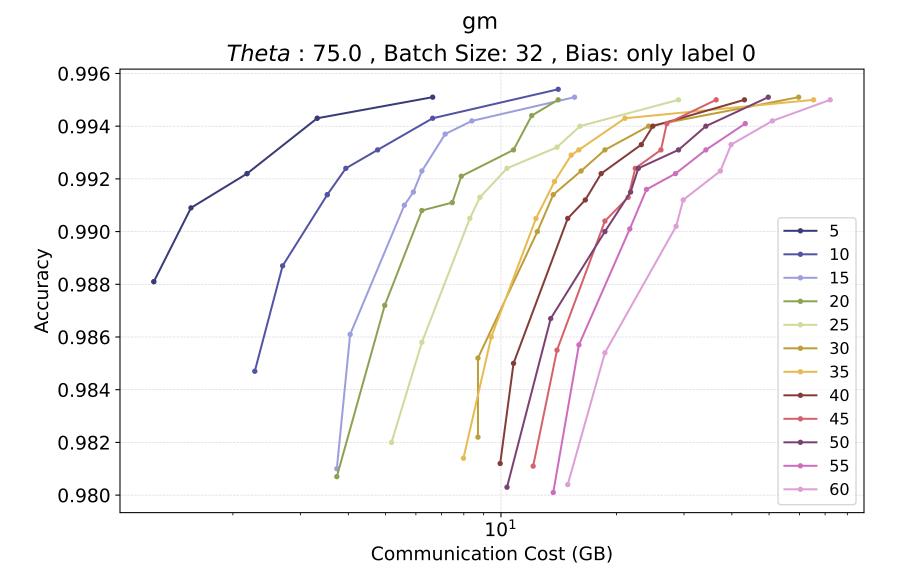


naive Theta: 50.0, Batch Size: 32, Bias: only label 0 0.9950 0.9925 0.9900 **—** 10 Accuracy 20 0.9875 25 **-** 30 0.9850 35 40 **-** 45 0.9825 **–** 50 55 -- 60 0.9800 10^{1} Communication Cost (GB)

linear *Theta*: 50.0, Batch Size: 32, Bias: only label 0



sketch Theta: 50.0, Batch Size: 32, Bias: only label 0 0.996 0.9940.992 Accuracy 886.0 - 10 20 25 **-** 30 0.986 40 0.984 45 **-** 50 55 0.982 **→** 60 100 10^{1} Communication Cost (GB)



naive Theta: 75.0, Batch Size: 32, Bias: only label 0 **—** 10 20 25 30 - 35 40 - 45 50 55 ---- 60 10^{1} 100

0.996

0.994

0.992

0.990

0.988

0.986

0.984

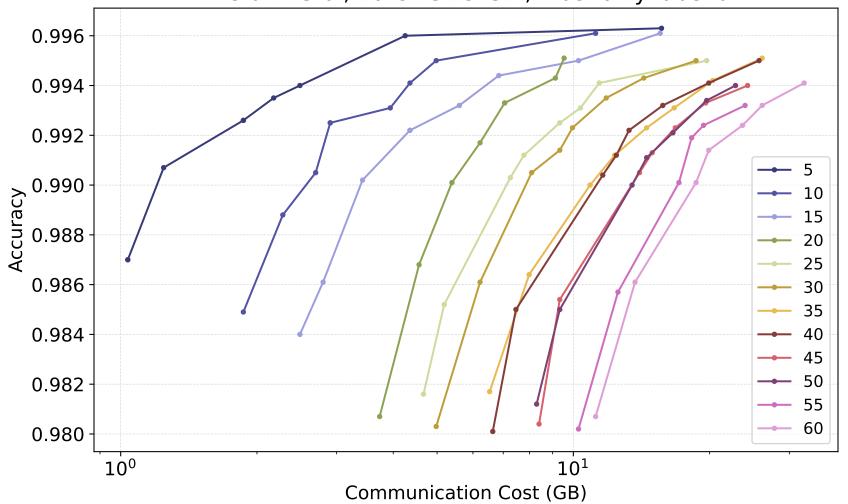
0.982

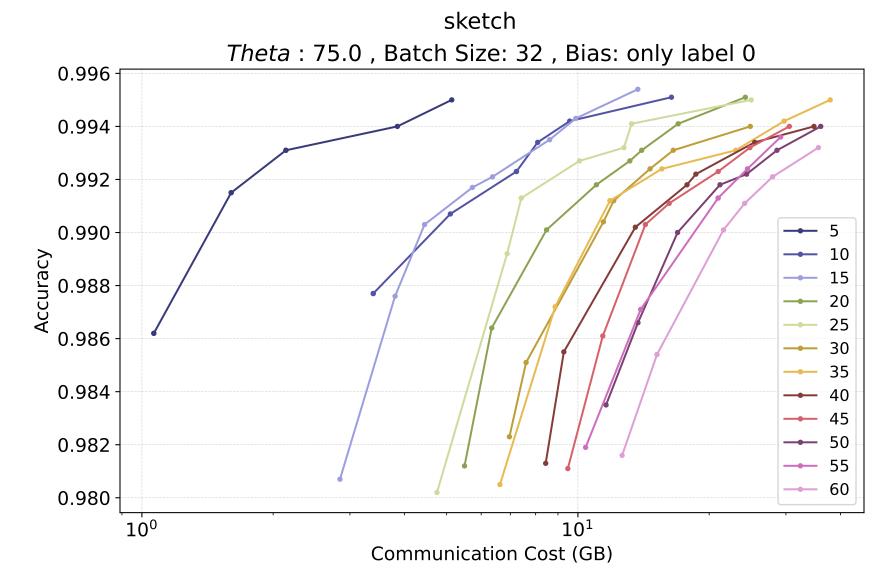
0.980

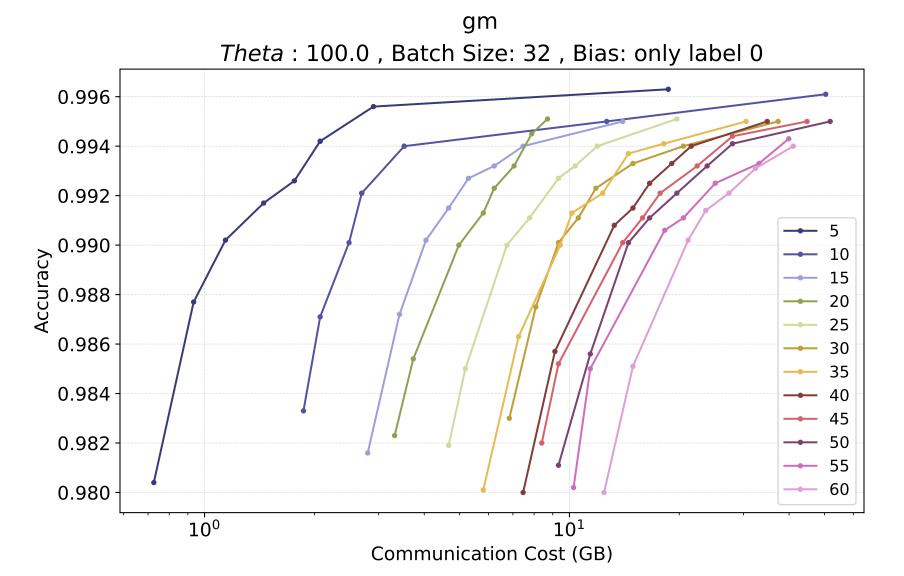
Accuracy

Communication Cost (GB)

linear *Theta*: 75.0, Batch Size: 32, Bias: only label 0



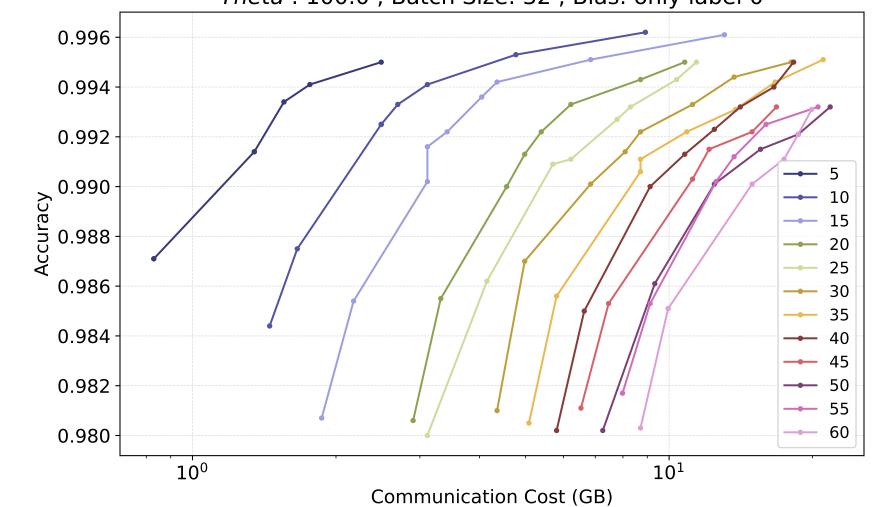


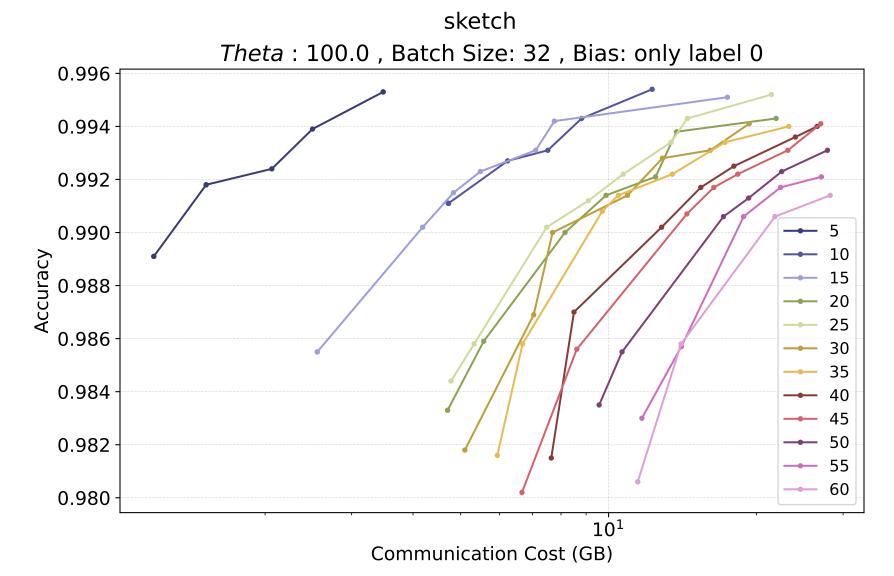


naive Theta: 100.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 0.990 Accuracy 10 0.988 20 25 0.986 30 - 35 0.984 40 - 45 0.982 50 55 -- 60 0.980 100 10^{1}

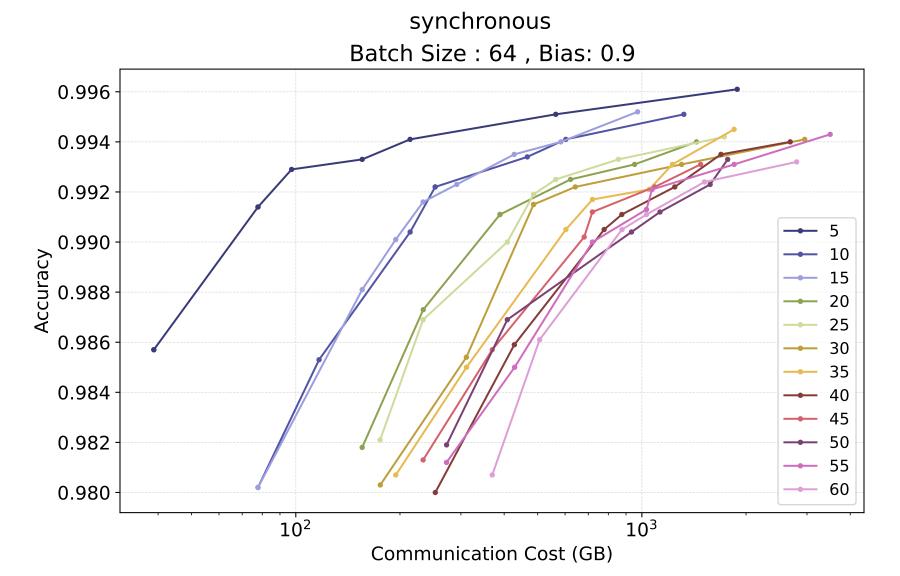
Communication Cost (GB)

linear
Theta: 100.0, Batch Size: 32, Bias: only label 0

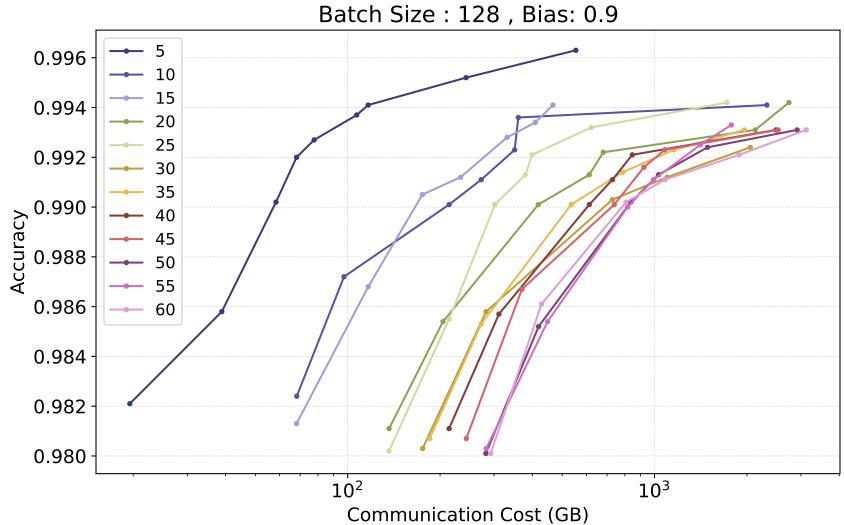


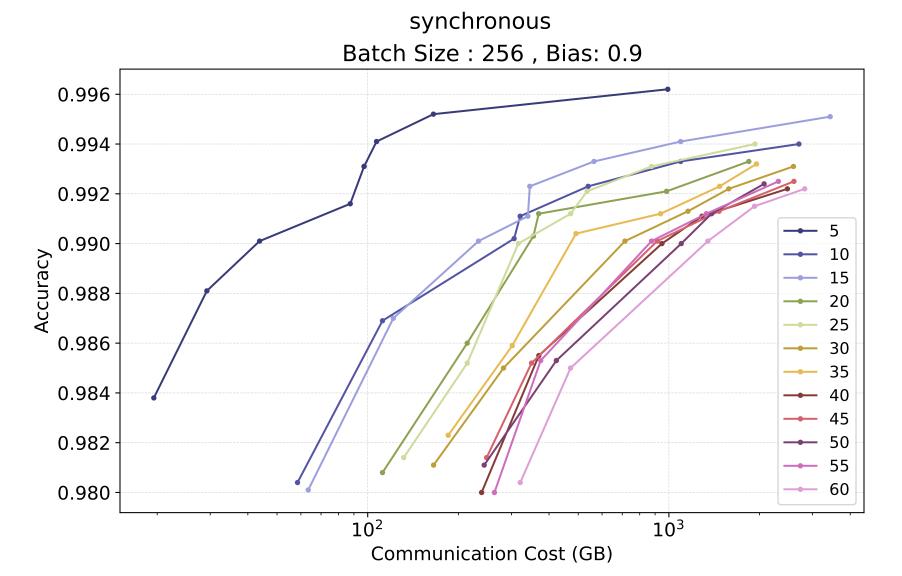


synchronous Batch Size: 32, Bias: 0.9 0.994 0.992 0.990 - 10 Accuracy 0.988 20 25 0.986 30 35 0.984 40 - 45 0.982 50 55 **-** 60 0.980 10² 10³ Communication Cost (GB)

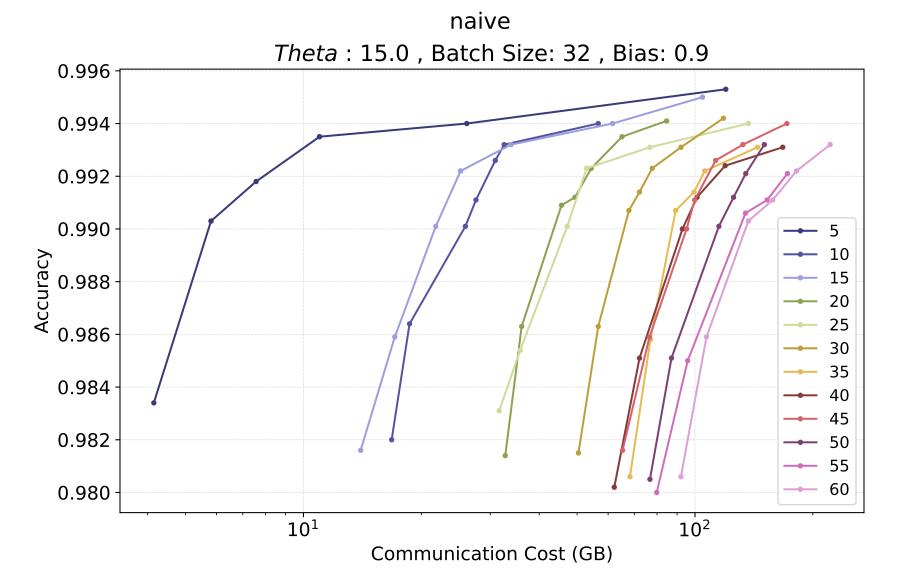


synchronous





gm Theta: 15.0, Batch Size: 32, Bias: 0.9 0.994 20 0.992 25 30 0.990 35 40 Accuracy 45 0.988 50 55 0.986 60 0.984 0.982 0.980 102 10^{1} Communication Cost (GB)

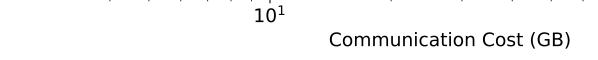


linear Theta: 15.0, Batch Size: 32, Bias: 0.9 10 15 20 25 30 **–** 35 40 **-** 45 50

55

60

10²



0.996

0.994

0.992

0.990

0.988

0.986

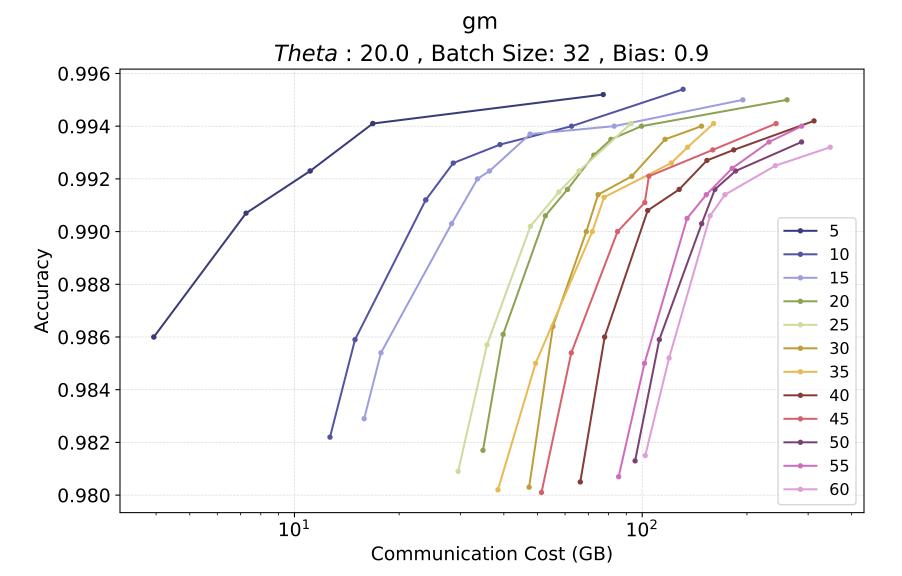
0.984

0.982

0.980

Accuracy

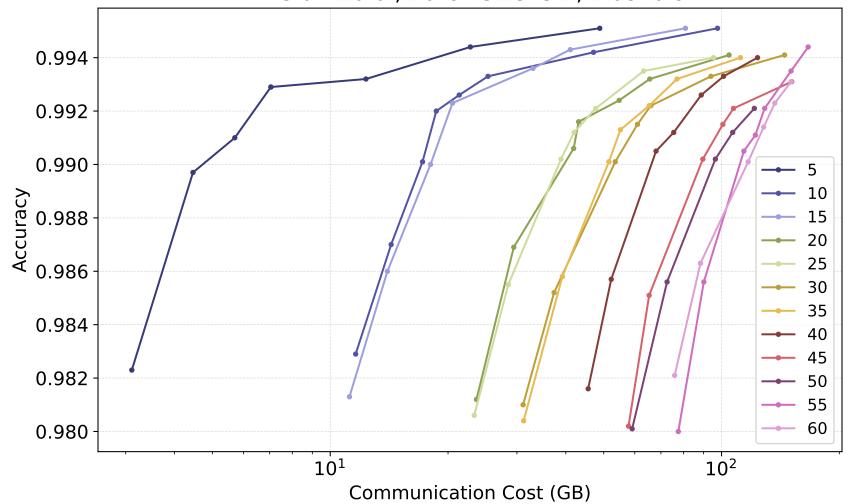
sketch Theta: 15.0, Batch Size: 32, Bias: 0.9 0.994 0.992 0.990 10 Accuracy 0.988 20 25 0.986 30 **-** 35 0.984 40 - 45 0.982 - 50 55 ---- 60 0.980 10^{1} 10² Communication Cost (GB)

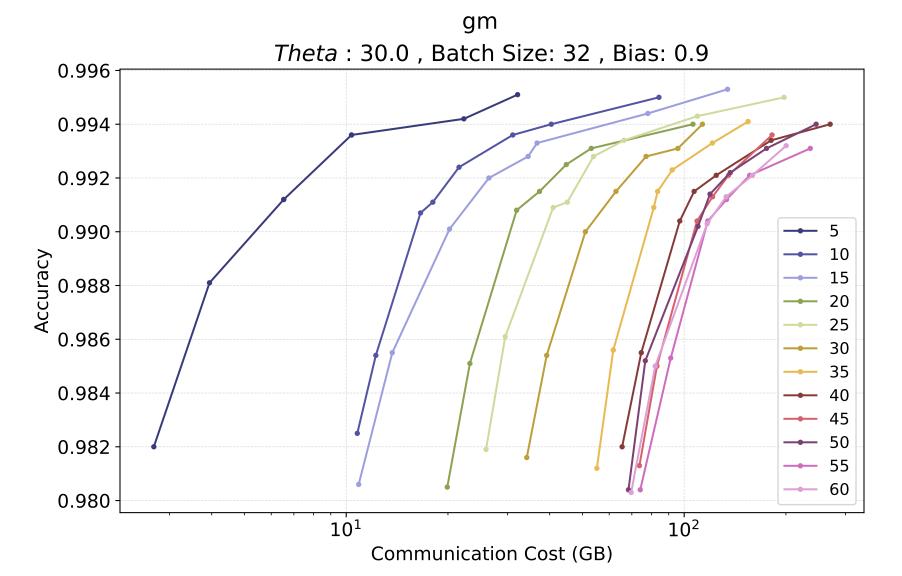


naive Theta: 20.0, Batch Size: 32, Bias: 0.9 0.994 0.992 0.990 10 Accuracy 0.988 20 25 0.986 30 **-** 35 0.984 40 **-** 45 0.982 **-** 50 **-** 55 -- 60 0.980 10^{1} 10² Communication Cost (GB)

linear

Theta: 20.0, Batch Size: 32, Bias: 0.9

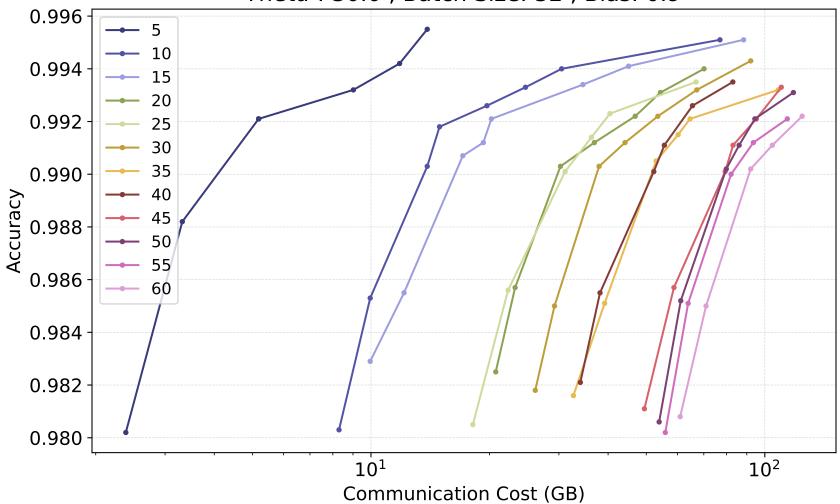




naive Theta: 30.0, Batch Size: 32, Bias: 0.9 0.994 0.992 0.990 10 Accuracy 0.988 20 25 0.986 30 **-** 35 0.984 40 **-** 45 0.982 **-** 50 **-** 55 **60** 0.980 10^{1} 10² Communication Cost (GB)

linear

Theta: 30.0, Batch Size: 32, Bias: 0.9

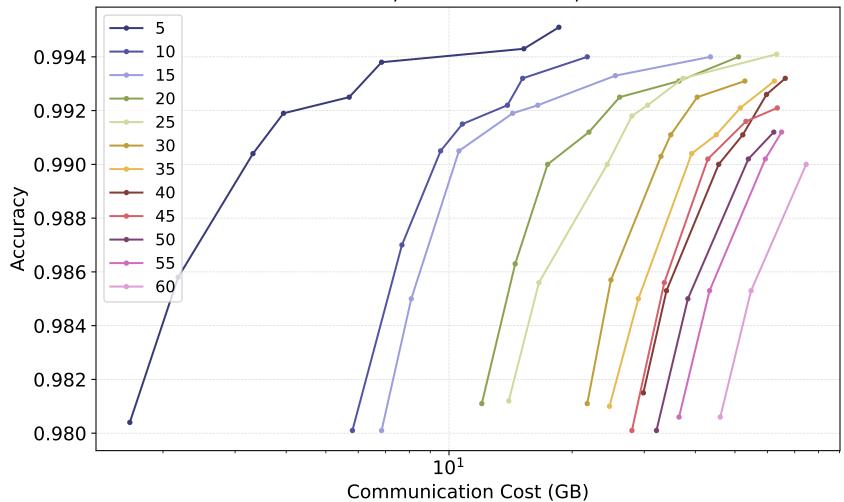


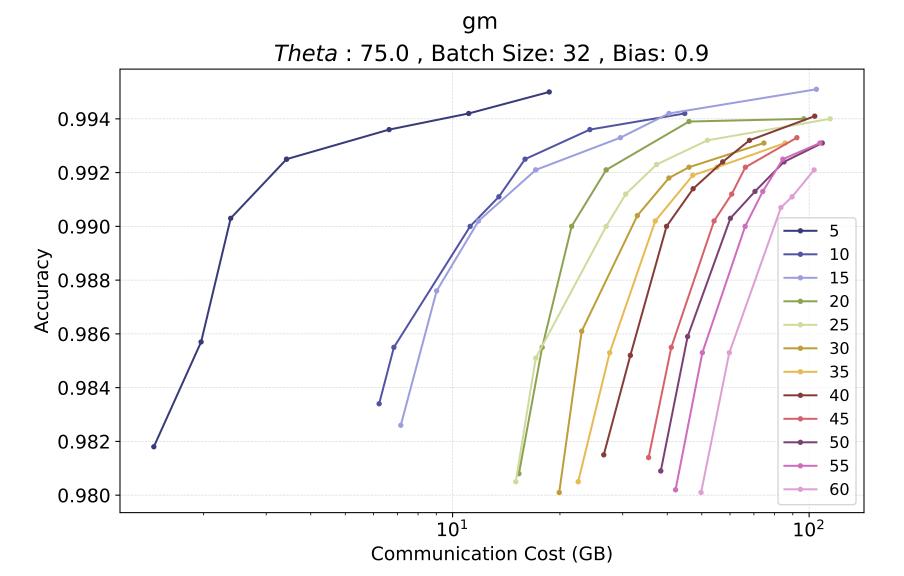
gm Theta: 50.0, Batch Size: 32, Bias: 0.9 0.994 0.992 25 30 0.990 35 Accuracy 986.0 50 55 60 0.984 0.982 0.980 10¹ 10² Communication Cost (GB)

naive Theta: 50.0, Batch Size: 32, Bias: 0.9 0.996 10 0.994 20 25 0.992 30 Accuracy .0 .0 .0 .0 35 45 55 0.986 60 0.984 0.982 0.980 10^{1} Communication Cost (GB)

linear

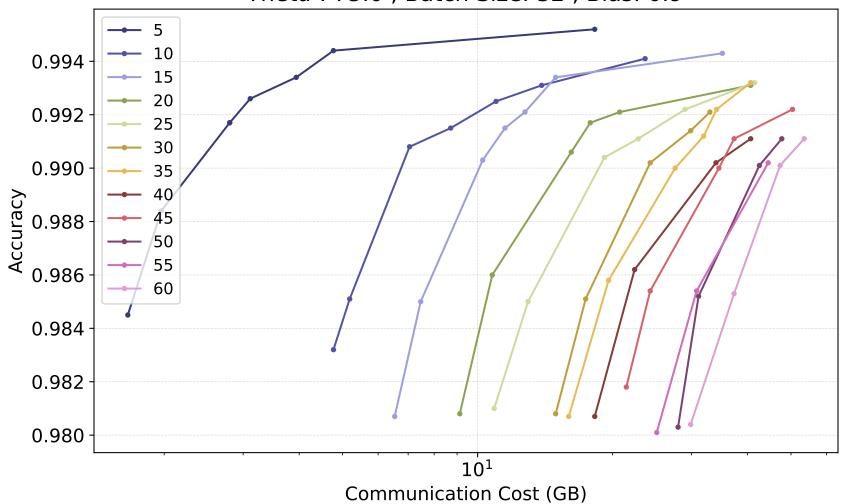
Theta: 50.0, Batch Size: 32, Bias: 0.9





naive Theta: 75.0, Batch Size: 32, Bias: 0.9 0.994 0.992 30 0.990 40 Accuracy 45 0.988 50 55 0.986 60 0.984 0.982 0.980 10^{1} Communication Cost (GB)

linear
Theta: 75.0, Batch Size: 32, Bias: 0.9



gm Theta: 100.0, Batch Size: 32, Bias: 0.9 0.994 0.992 0.990 Accuracy 986.0 **-** 10 20 25 30 35 0.984 40 **-** 45 0.982 **-** 50 55 -- 60 0.980 10^{1} 10² Communication Cost (GB)

naive Theta: 100.0, Batch Size: 32, Bias: 0.9 0.994 0.992 30 0.990 35 Accuracy 886.0 40 50 55 60 0.984 0.982 0.980 10^{1} Communication Cost (GB)

linear Theta: 100.0, Batch Size: 32, Bias: 0.9 10 0.994 0.992 25 30 0.990 35 40 Accuracy 45 0.988 55 0.986 60 0.984 0.982 0.980 10^{1} Communication Cost (GB)

sketch Theta: 100.0, Batch Size: 32, Bias: 0.9 0.994 0.992 25 30 0.990 35 Accuracy 986.0 50 55 60 0.984 0.982 0.980 100 10^{1} Communication Cost (GB)