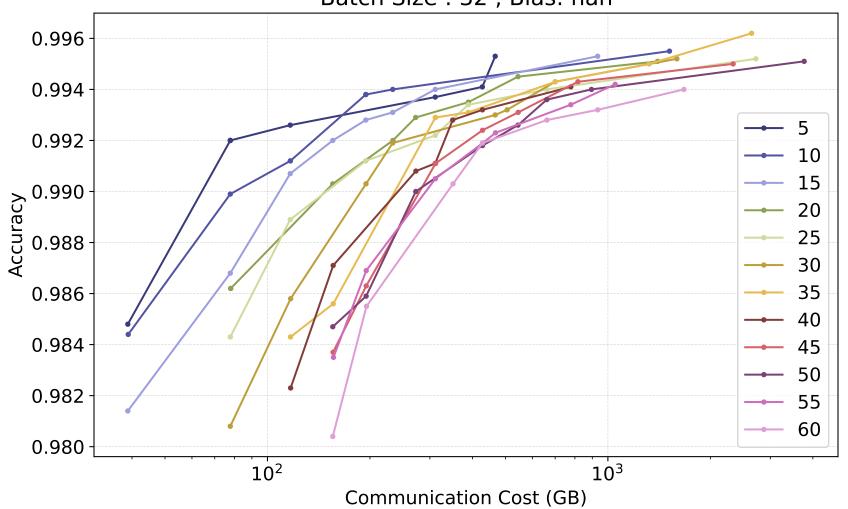
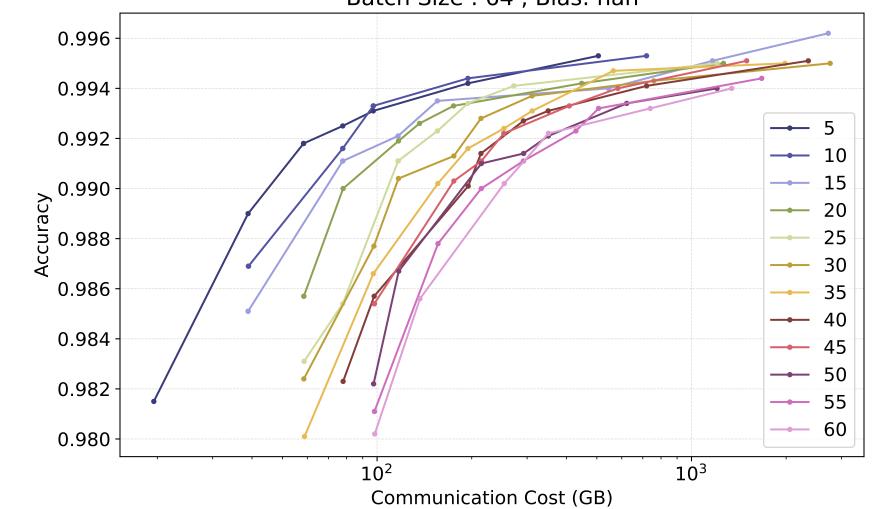
synchronous Batch Size : 32 , Bias: nan



synchronous Batch Size : 64 , Bias: nan



synchronous Batch Size: 128, Bias: nan 0.996 0.994 -0.992 10 15 Accuracy 886.0 20 25 30 35 0.986 40 0.984 45 50 0.982 55 60 0.980

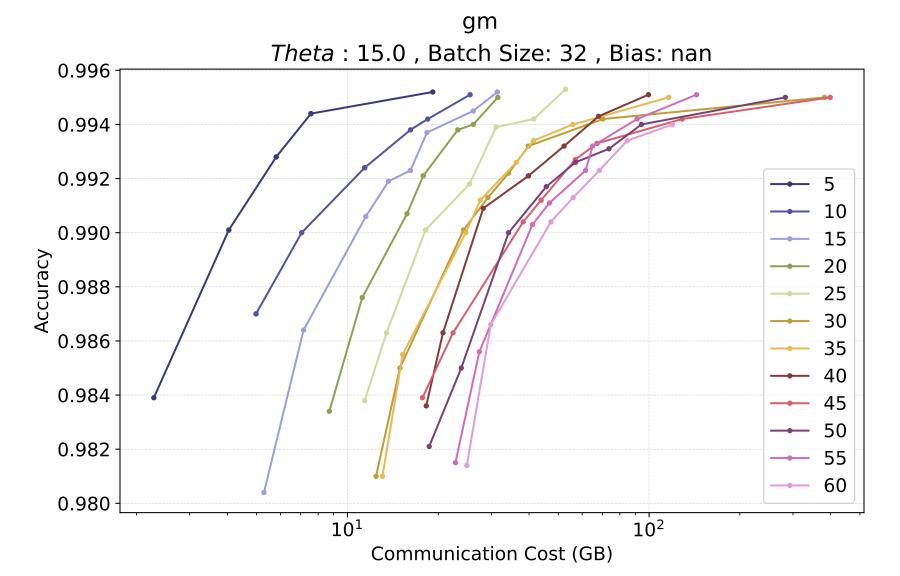
10²

Communication Cost (GB)

10³

10¹

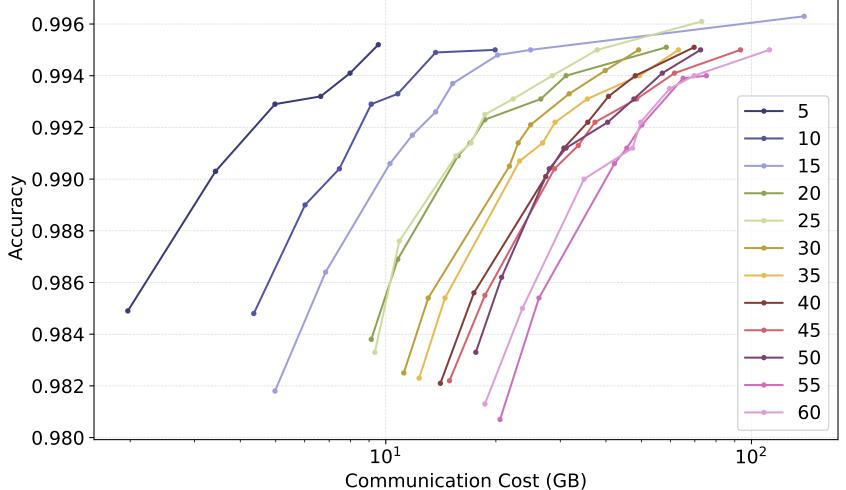
synchronous Batch Size: 256, Bias: nan 0.996 0.994 0.992 10 Accuracy 0 0 0 66 15 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 102 10³ 10¹ Communication Cost (GB)



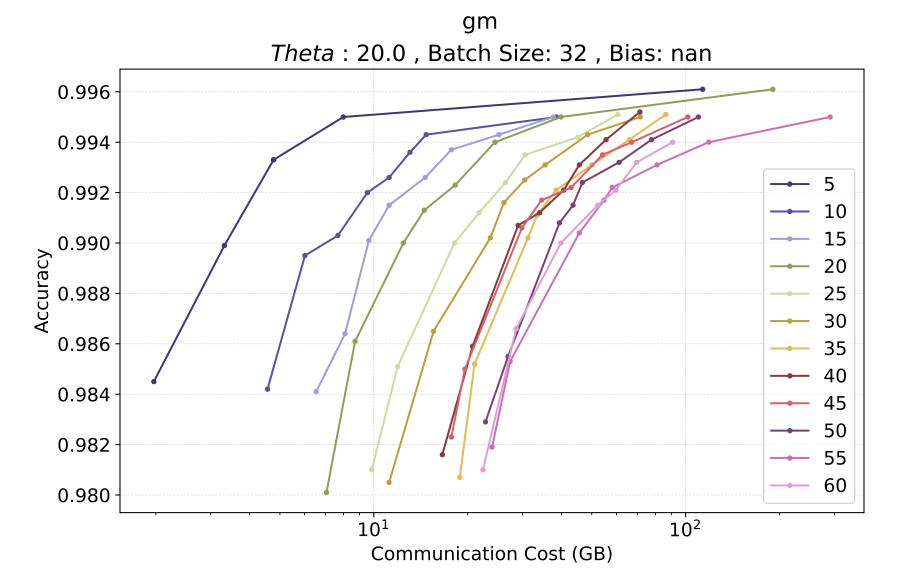
naive Theta: 15.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 10 15 Accuracy 886.0 886.0 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10^{1} 102 Communication Cost (GB)

linear

Theta: 15.0, Batch Size: 32, Bias: nan

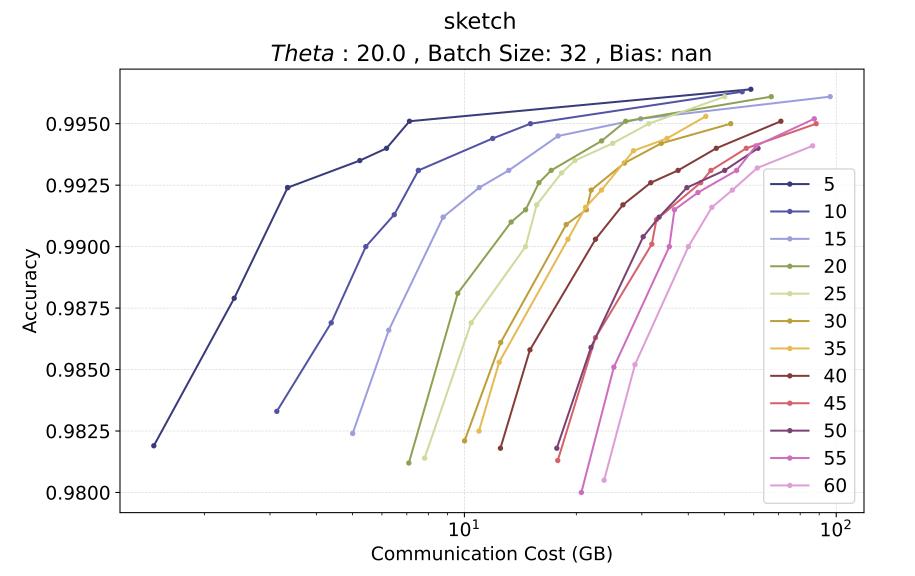


sketch Theta: 15.0, Batch Size: 32, Bias: nan 0.996 0.9940.992 10 15 Accuracy 886.0 20 25 30 35 0.986 40 0.984 45 50 0.982 55 60 0.980 10^{1} 10² Communication Cost (GB)



naive Theta: 20.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 10 Accuracy .0 0:0 15 20 25 30 0.986 35 40 0.984 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 10 Accuracy 886.0 15 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: 25.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 10 Accuracy 0 0 0 66 15 20 25 30 0.986 35 40 0.984 45 50 0.982 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 10 15 Accuracy 886.0 20 25 30 35 0.986 40 0.984 45 50 0.982 55 60 0.980 10^{1} 10² Communication Cost (GB)

naive Theta: 30.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 10 Accuracy 886.0 15 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10^{1} 10^2 Communication Cost (GB)

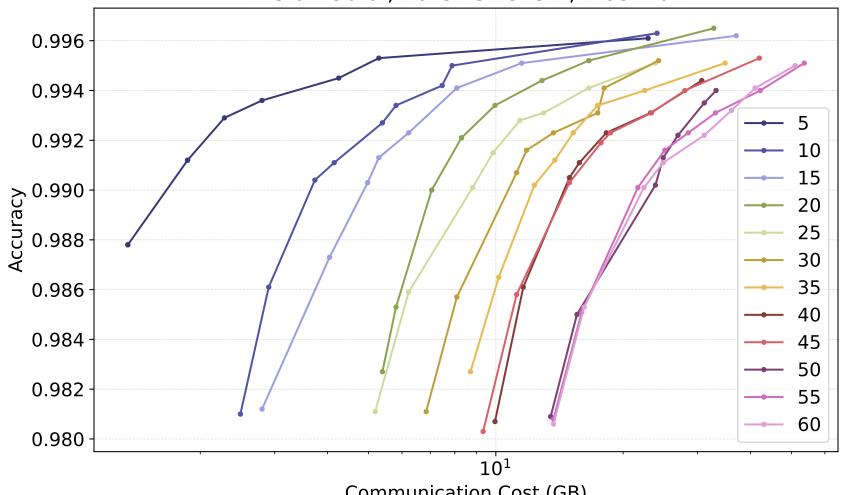
linear Theta: 30.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 10 0.990 15 Accuracy 20 0.988 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10^{1} Communication Cost (GB)

sketch Theta: 30.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 15 Accuracy 886.0 886.0 20 25 30 35 0.986 40 45 0.984 50 55 0.982 60 10^{1} Communication Cost (GB)

gm Theta: 50.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 10 15 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 100 10^{1} 10² Communication Cost (GB)

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

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



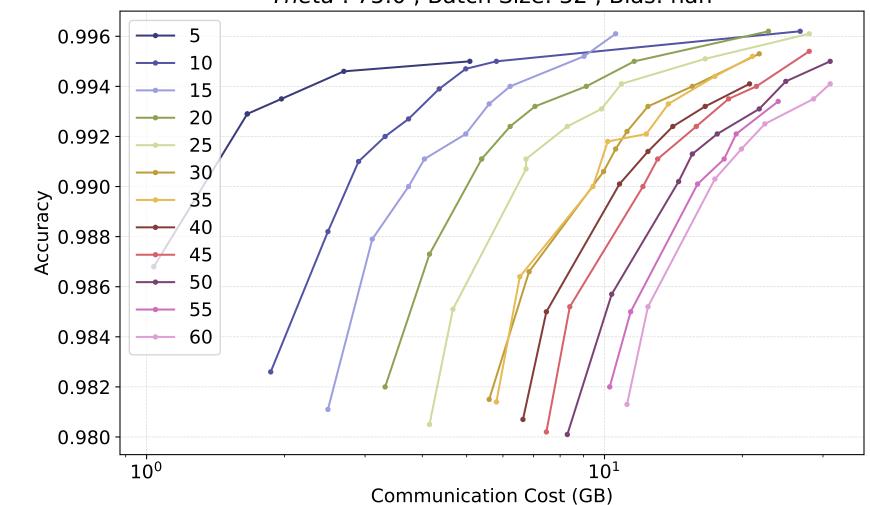
Communication Cost (GB)

sketch Theta: 50.0, Batch Size: 32, Bias: nan 0.996 0.994 0.992 Accuracy 0 0 0 66 15 20 25 30 0.986 35 40 0.984 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 10 15 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: 75.0, Batch Size: 32, Bias: nan 0.996 10 0.99420 0.992 25 Accuracy .0 .0 .0 66 30 35 40 45 50 0.986 55 0.984 60 0.982 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 10 15 0.990 Accuracy 20 0.988 25 30 0.986 35 40 0.984 45 50 0.982 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 10 15 0.990 Accuracy 20 0.988 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 100 10^{1} Communication Cost (GB)

naive Theta: 100.0, Batch Size: 32, Bias: nan 0.996 10 0.99415 20 0.992 25 Accuracy .0 .0 .0 .0 30 35 40 45 50 0.986 55 0.984 60 0.982 0.980 100 10^{1} Communication Cost (GB)

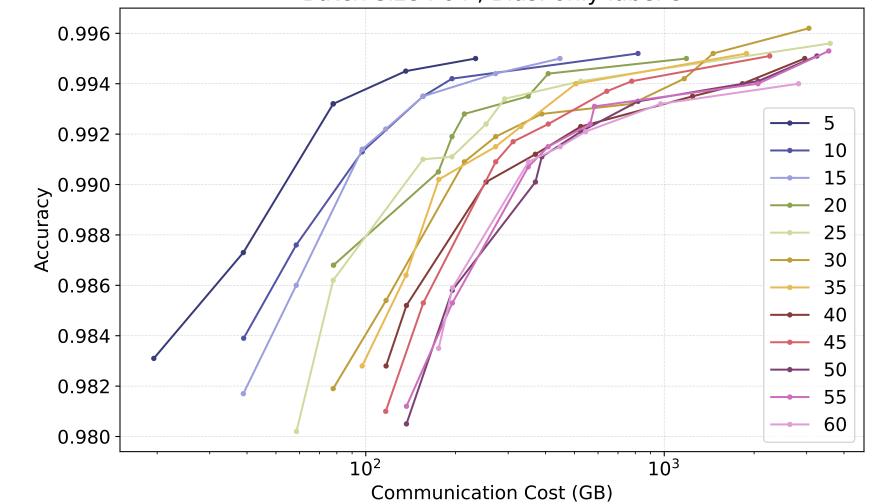
linear Theta: 100.0, Batch Size: 32, Bias: nan 0.996 0.994 15 20 0.992 25 0.990 30 Accuracy 35 0.988 40 45 0.986 50 55 0.98460 0.982 0.980 100 10^{1} Communication Cost (GB)

sketch Theta: 100.0, Batch Size: 32, Bias: nan 0.996 10 0.99415 20 0.992 25 Accuracy .0 .0 .0 66 30 35 40 45 50 0.986 55 0.984 60 0.982 0.980 100 10^{1} Communication Cost (GB)

synchronous Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 10 0.990 15 Accuracy 20 0.988 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10² 10³ Communication Cost (GB)

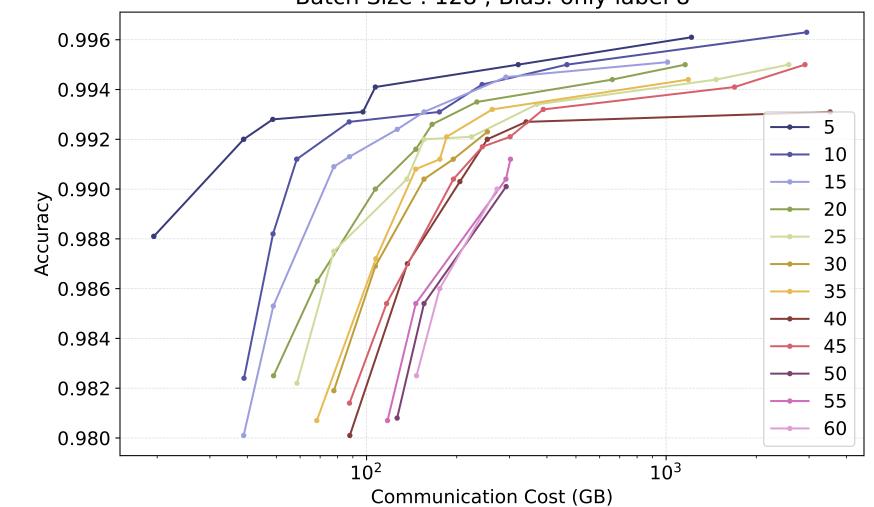
synchronous

Batch Size: 64, Bias: only label 8



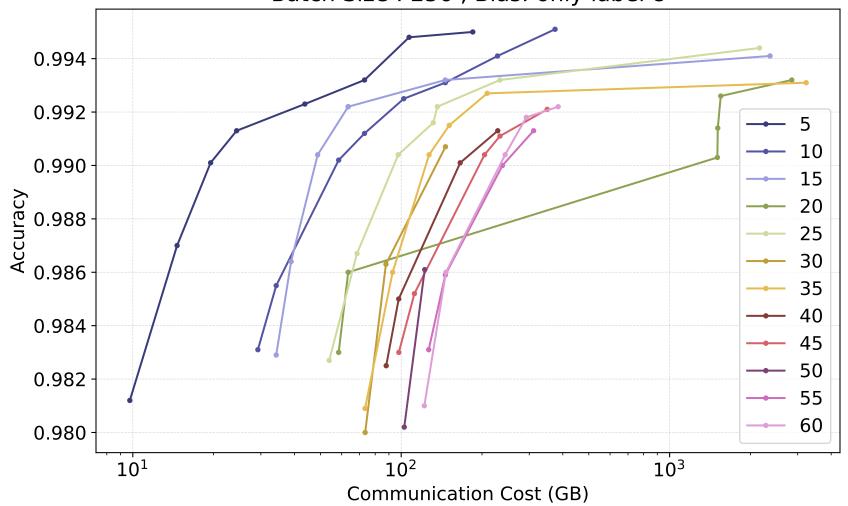
synchronous

Batch Size: 128, Bias: only label 8



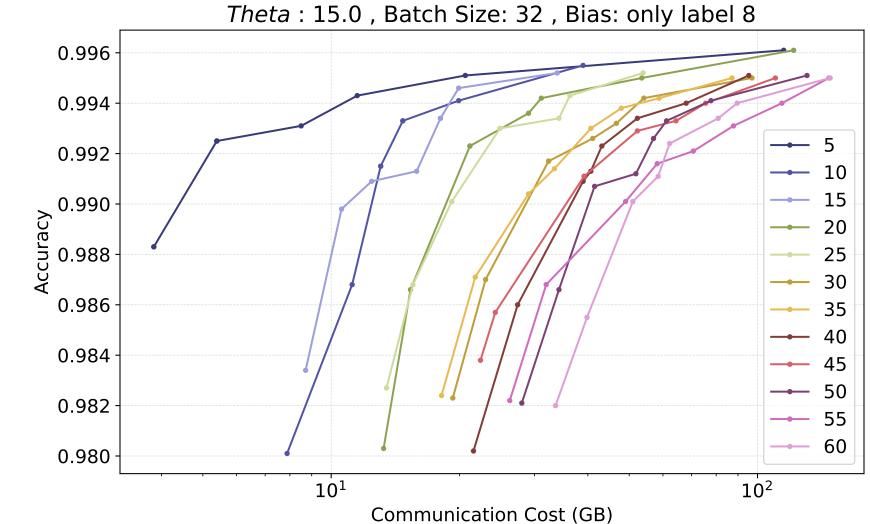
synchronous

Batch Size: 256, Bias: only label 8

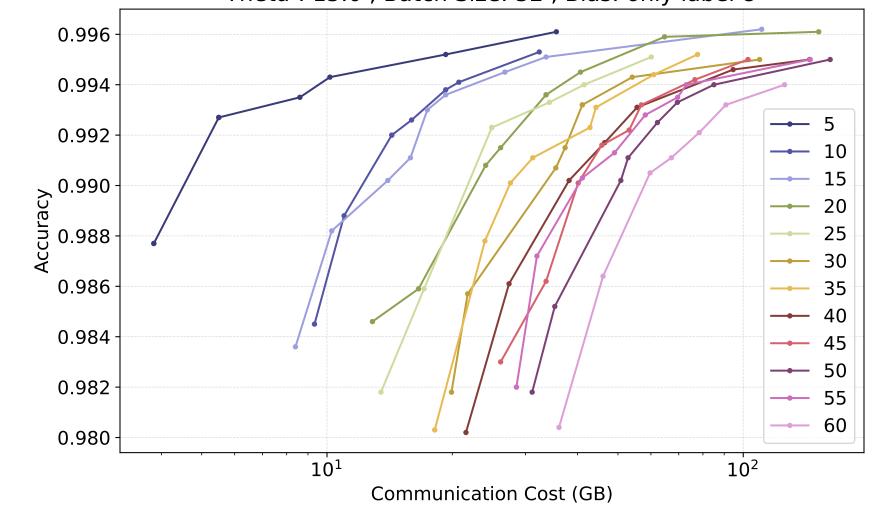


gm Theta: 15.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 10 Accuracy .0 0.0 15 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10¹ 10² Communication Cost (GB)

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 10 15 0.990 Accuracy 20 0.988 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10^{1} 10^2

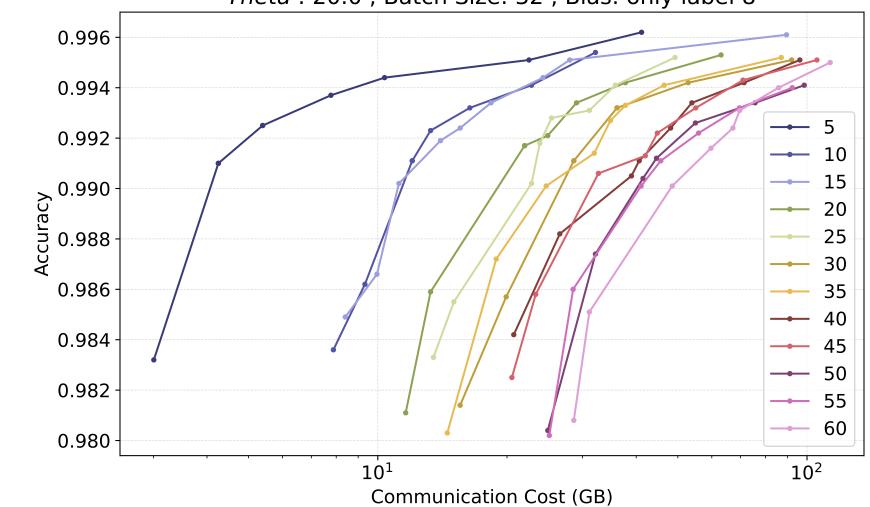
Communication Cost (GB)

gm Theta: 20.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 10 15 Accuracy 886.0 20 25 30 35 0.986 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 10 15 Accuracy 886.0 20 25 30 35 0.986 40 45 0.984 50 55 0.982 60 0.980 10^{1} 10² Communication Cost (GB)

linear

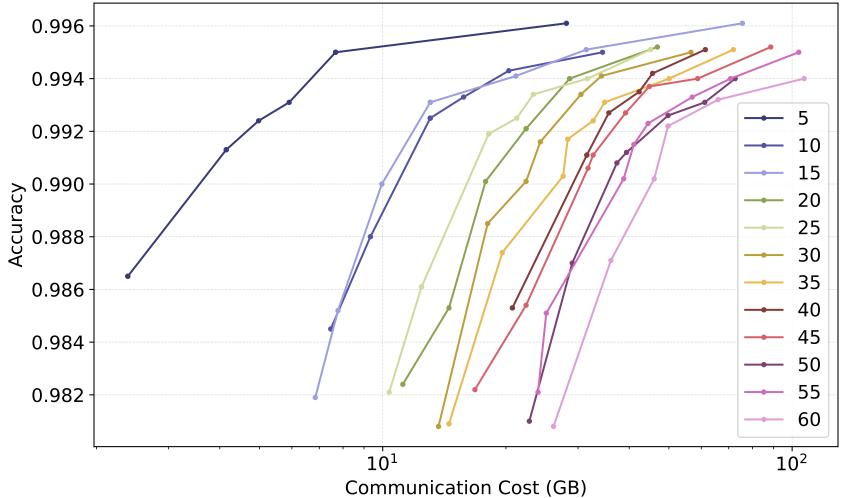
Theta: 20.0, Batch Size: 32, Bias: only label 8



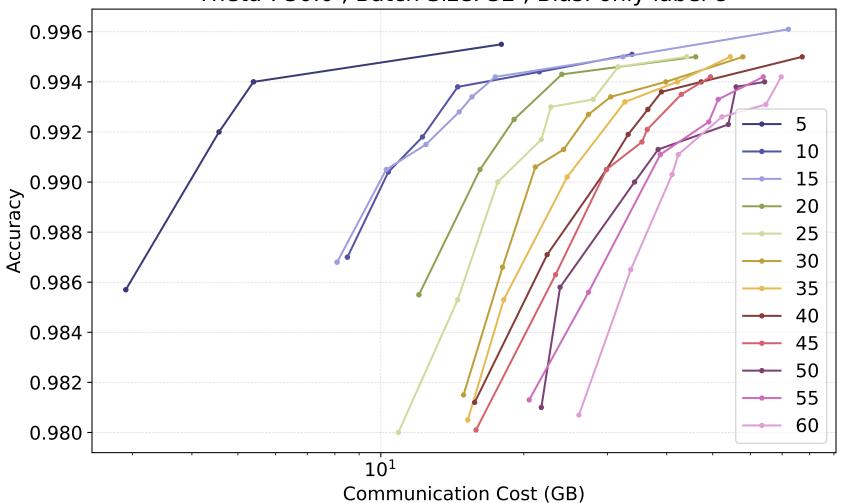
sketch Theta: 20.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 10 Accuracy 0 0 0 66 15 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10^{1} 10² Communication Cost (GB)

gm Theta: 30.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 10 Accuracy 0 0 0.0 15 20 25 30 0.986 35 40 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 8

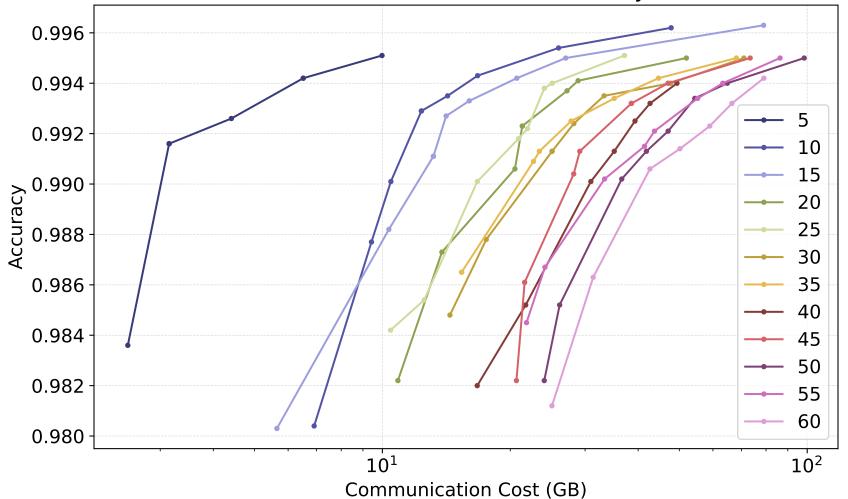


linear *Theta*: 30.0, Batch Size: 32, Bias: only label 8



sketch

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



gm Theta: 50.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 10 Accuracy 0 0 0 66 15 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: 50.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 15 0.990 Accuracy 20 0.988 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10^{1} Communication Cost (GB)

linear Theta: 50.0, Batch Size: 32, Bias: only label 8 10 15 20 25 30 35 40 45 50 55 60 10^{1}

0.996

0.994

0.992

0.986

0.984

0.982

0.980

Accuracy .0 0:0

Communication Cost (GB)

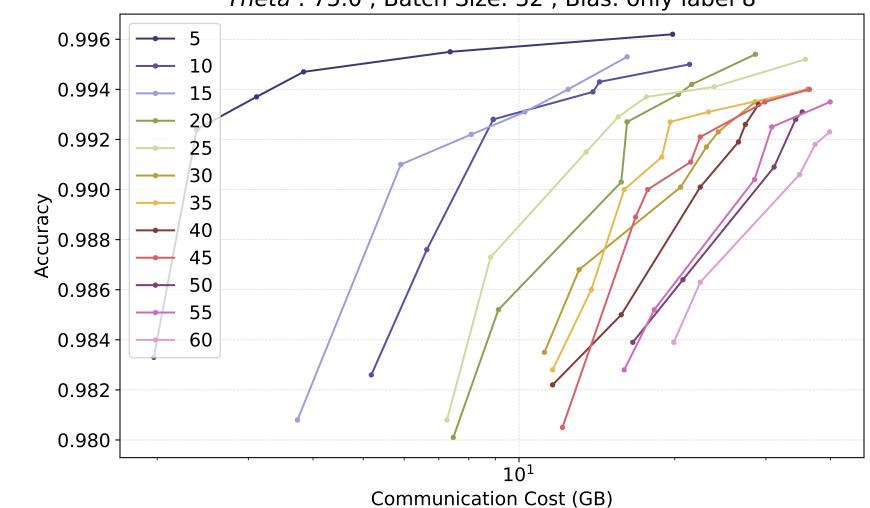
sketch Theta: 50.0, Batch Size: 32, Bias: only label 8 0.996 0.9940.992 15 Accuracy 886.0 20 25 30 35 0.986 40 0.98445 50 0.982 55 60 0.980 10^{1} Communication Cost (GB)

gm Theta: 75.0, Batch Size: 32, Bias: only label 8 0.996 0.994 0.992 10 15 Accuracy 886.0 20 25 30 35 0.986 40 45 0.984 50 55 0.982 60 10^{1} 10² Communication Cost (GB)

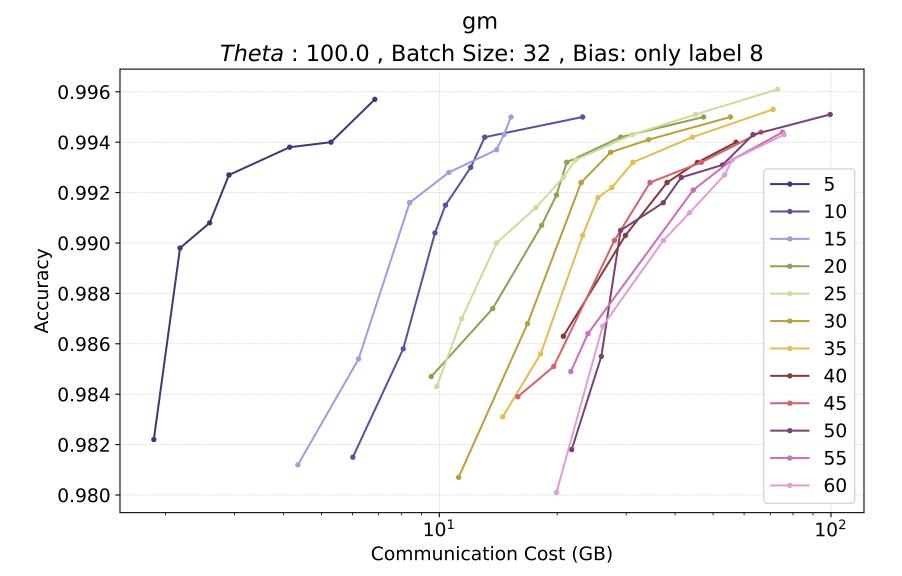
naive Theta: 75.0, Batch Size: 32, Bias: only label 8 0.996 10 0.994 20 0.992 25 Accuracy 0 0 0 0 30 35 40 45 50 0.986 55 60 0.984 0.982 0.980 10^{1} Communication Cost (GB)

linear

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







naive Theta: 100.0, Batch Size: 32, Bias: only label 8 0.994 0.992 25 30 0.990 Accuracy 40 0.988 45 0.986 55 60 0.984 0.982 10^1 Communication Cost (GB)

linear Theta: 100.0, Batch Size: 32, Bias: only label 8 0.996 0.994 15 0.992 20 25 0.990 30 Accuracy 35 0.988 40 45 0.986 50 55 0.984 60 0.982 0.980 10^1 Communication Cost (GB)

sketch Theta: 100.0, Batch Size: 32, Bias: only label 8 0.996 10 0.99420 0.992 25 Accuracy 0 0 0 0 30 35 40 45 50 0.986 55 60 0.984 0.982 0.980 10^{1} Communication Cost (GB)

synchronous Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 10 15 Accuracy 886.0 20 25 30 35 0.986 40 45 0.984 50 0.982 55 60

102

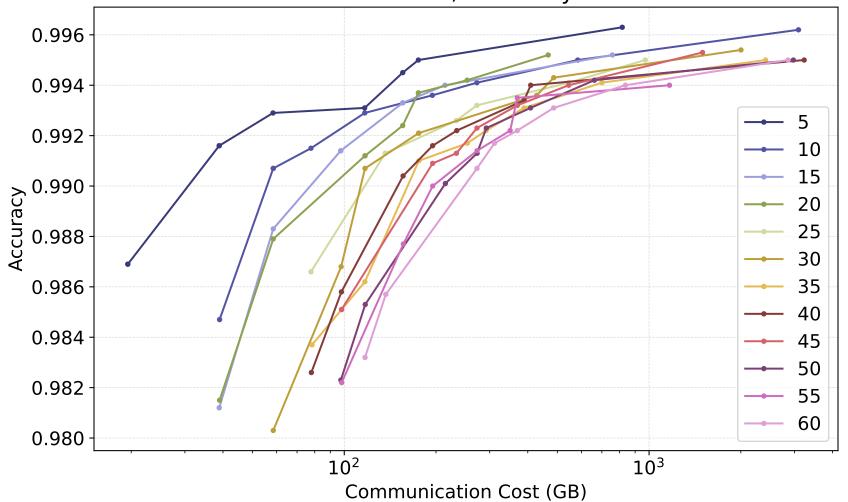
0.980

Communication Cost (GB)

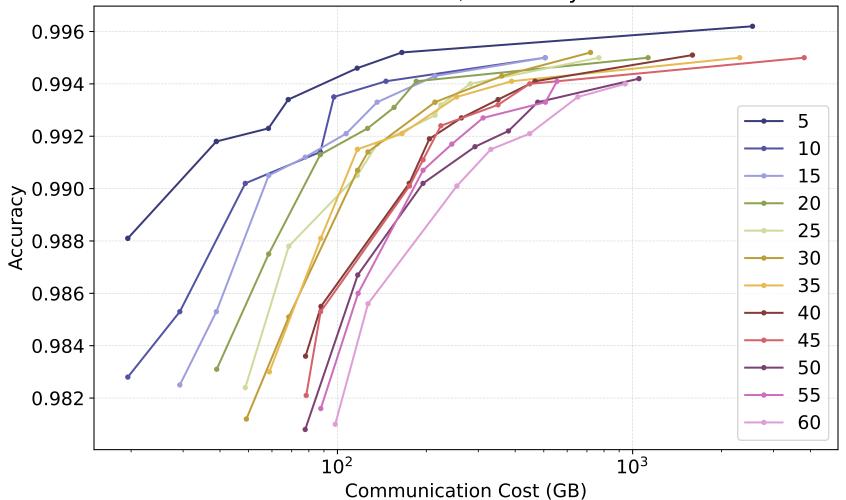
10³

synchronous

Batch Size: 64, Bias: only label 0



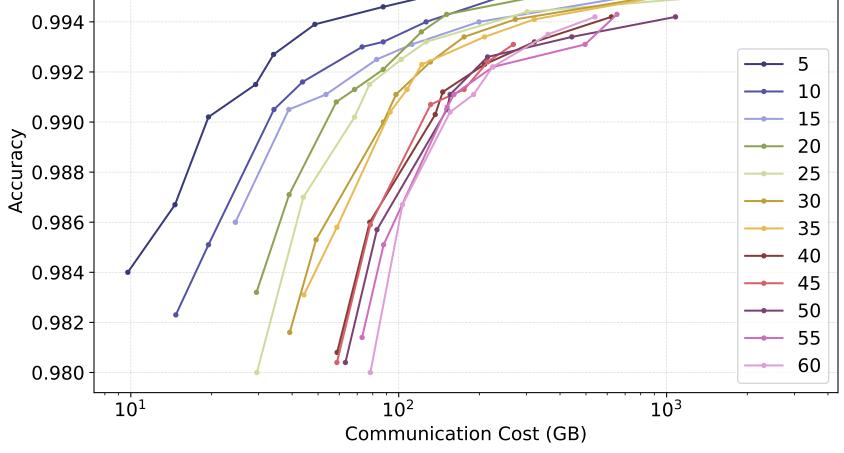
synchronous Batch Size: 128, Bias: only label 0



synchronous

0.996

Batch Size: 256, Bias: only label 0



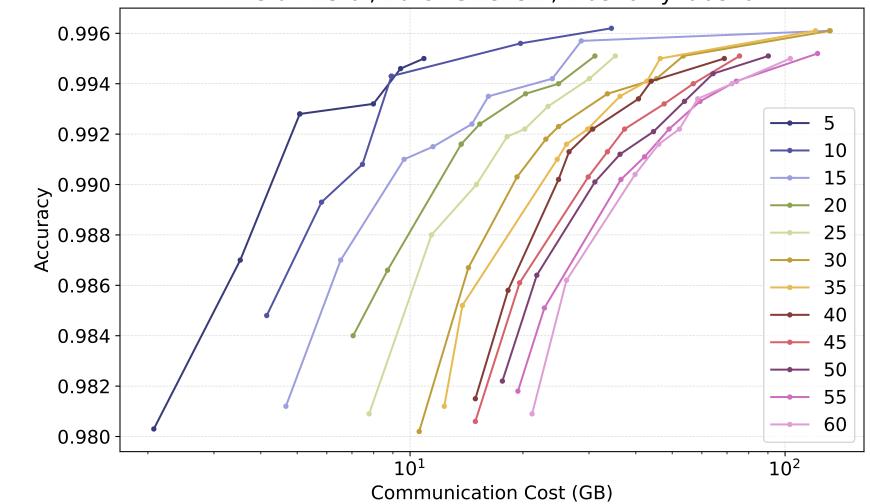
gm Theta: 15.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 10 15 Accuracy 886.0 886.0 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 102 10¹ Communication Cost (GB)

naive Theta: 15.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 10 Accuracy 0 0 0.0 15 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10^{1} 10²

Communication Cost (GB)

linear

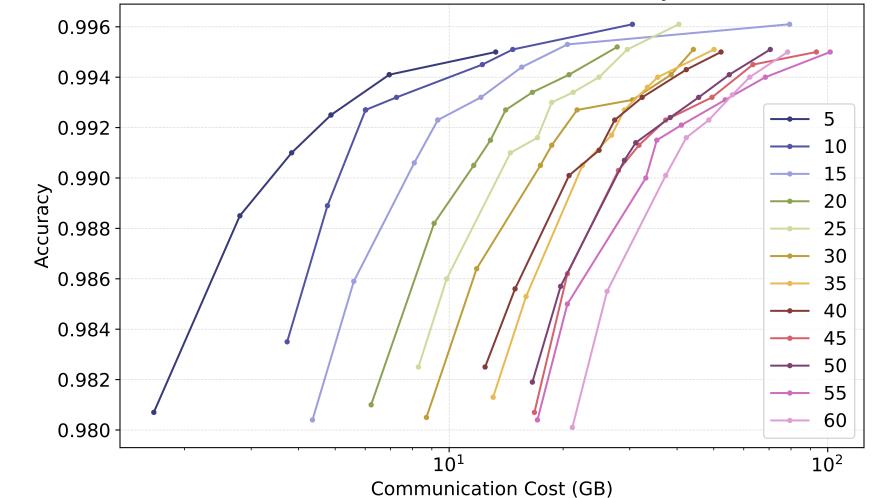
Theta: 15.0, Batch Size: 32, Bias: only label 0



sketch Theta: 15.0, Batch Size: 32, Bias: only label 0 0.996 0.9940.992 10 Accuracy 0 0 0 0 15 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 101 10² Communication Cost (GB)

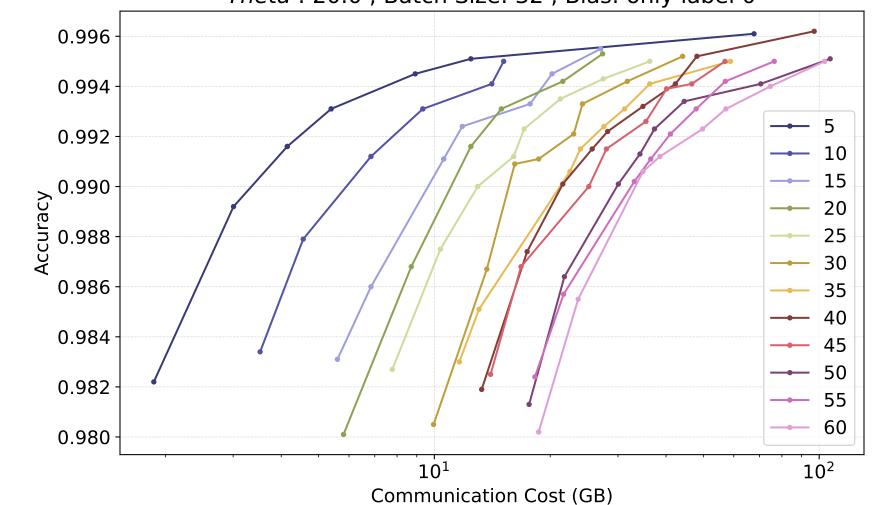
gm Theta: 20.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 10 Accuracy 0.0 ر.0 15 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10¹ 10² Communication Cost (GB)

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



linear

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

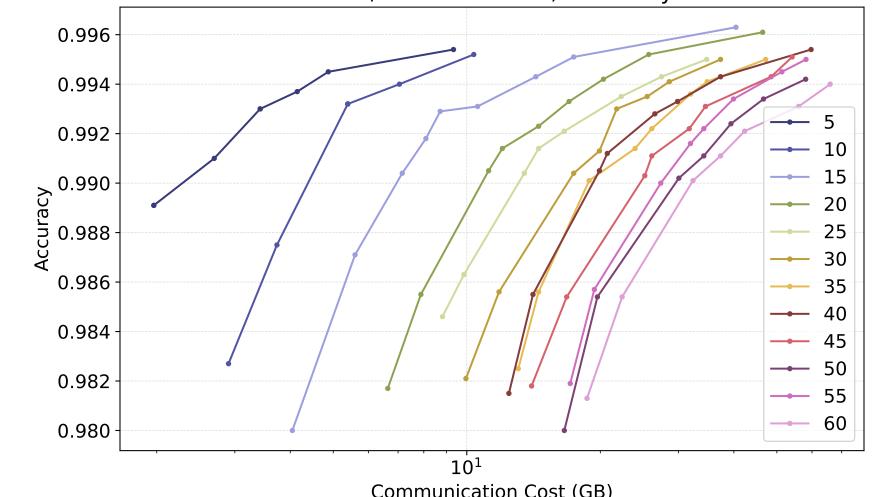


sketch Theta: 20.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 10 Accuracy 0 0 0.0 15 20 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 10¹ 10² Communication Cost (GB)

gm Theta: 30.0, Batch Size: 32, Bias: only label 0 0.996 10 0.994 20 0.992 25 30 35 40 45/ 50 0.986 55 0.984 60 0.982 0.980 10¹ 10² Communication Cost (GB)

naive Theta: 30.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 10 15 Accuracy 886.0 20 25 30 0.986 35 40 0.98445 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)

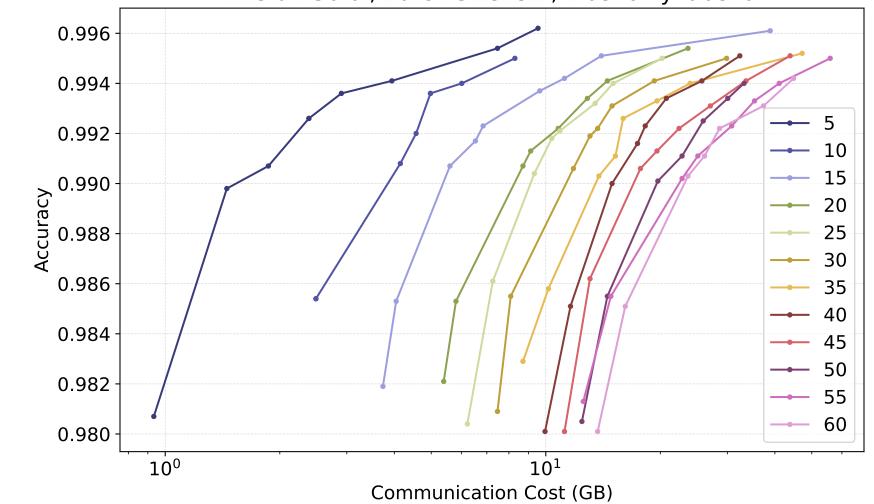
sketch Theta: 30.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 10 15 0.990 Accuracy 20 0.988 25 30 0.986 35 40 0.98445 50 0.982 55 60 0.980 10^{1} 10² Communication Cost (GB)

gm Theta: 50.0, Batch Size: 32, Bias: only label 0 0.996 0.994 0.992 10 Accuracy 0.0 ر.0 15 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: 50.0, Batch Size: 32, Bias: only label 0 0.9950 0.9925 10 15 0.9900 Accuracy 20 25 0.9875 30 35 0.9850 40 45 0.9825 50 55 60 0.9800 10^{1} Communication Cost (GB)

linear

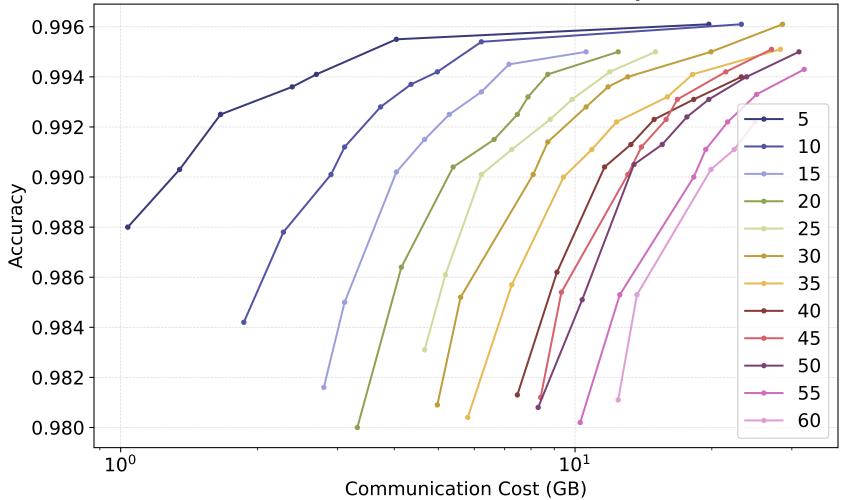
Theta: 50.0, Batch Size: 32, Bias: only label 0



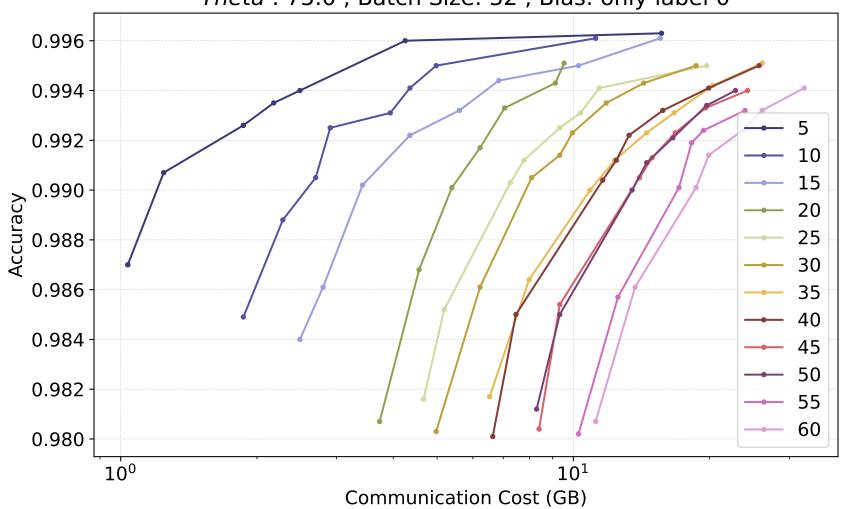
sketch Theta: 50.0, Batch Size: 32, Bias: only label 0 0.996 0.9940.992 15 Accuracy 886.0 20 25 30 35 0.986 40 45 0.984 50 55 0.982 60 100 10^{1} Communication Cost (GB)

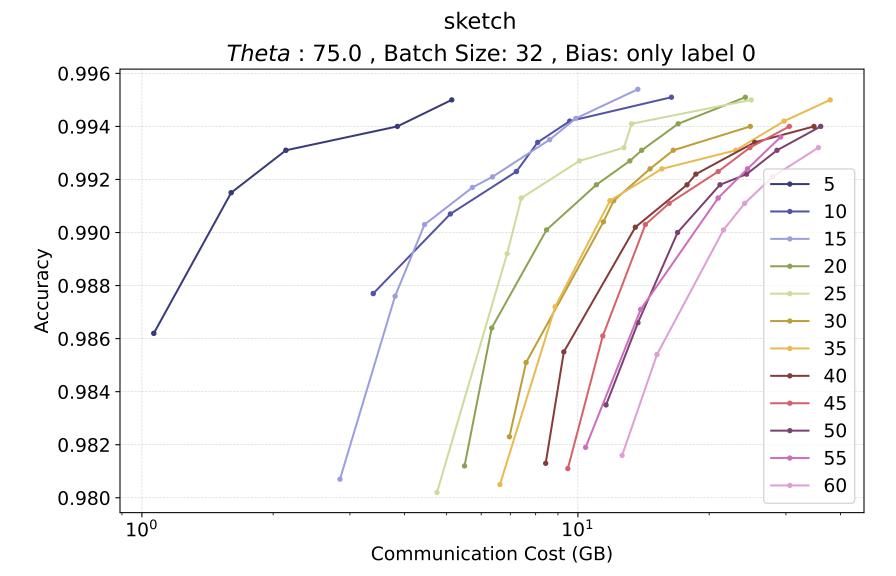


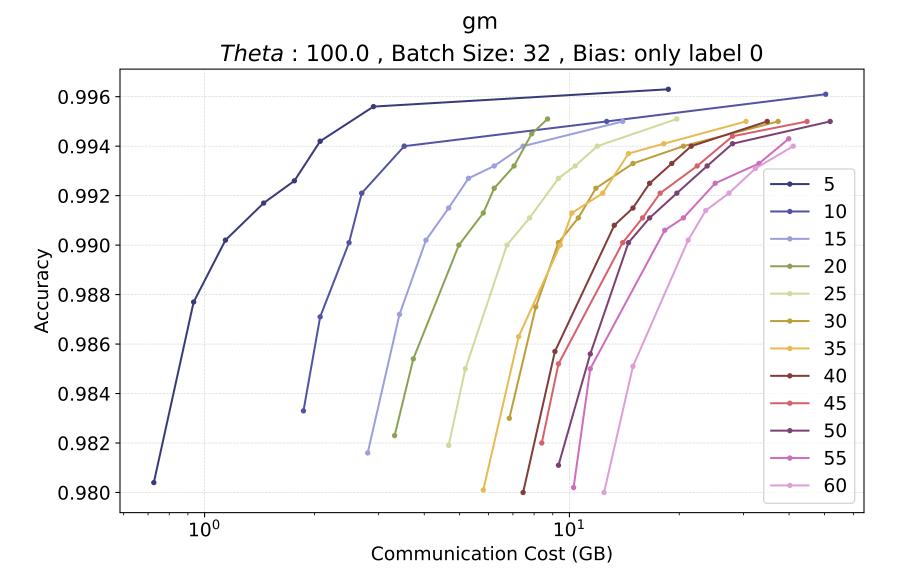
naive Theta: 75.0, Batch Size: 32, Bias: only label 0



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







naive Theta: 100.0, Batch Size: 32, Bias: only label 0 10 20 25 30 35 40 45 50 55 60

0.996

0.994

0.992

0.986

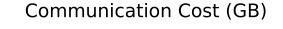
0.984

0.982

0.980

100

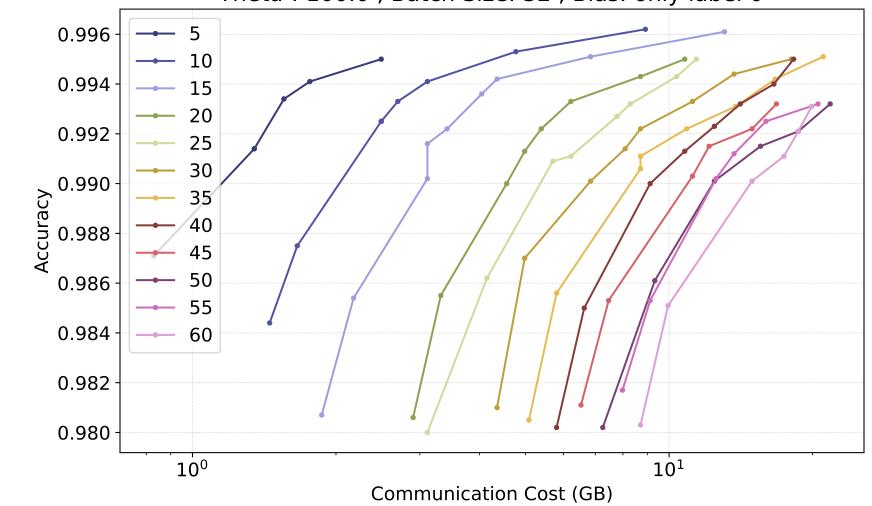
Accuracy .0 .0 .0 .0

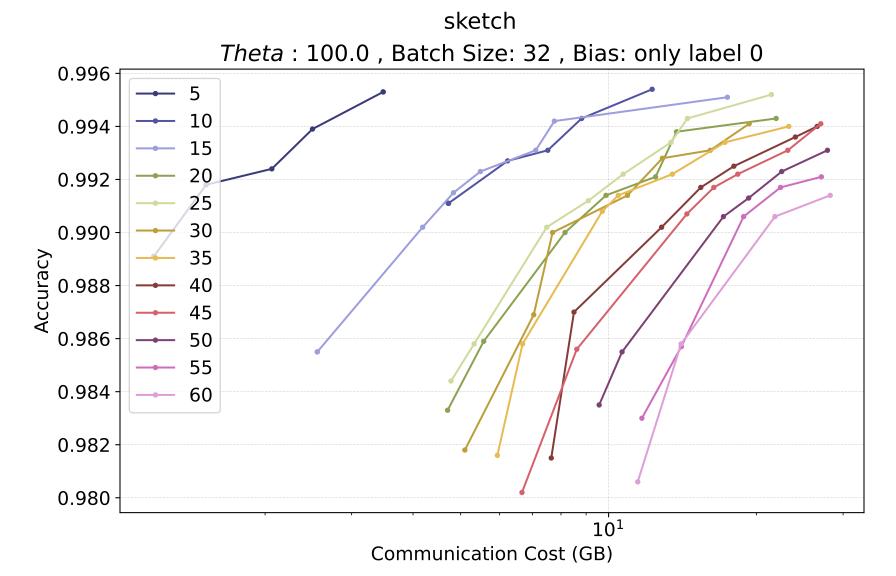


 10^{1}

linear

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





synchronous Batch Size: 32, Bias: 0.9 10 15 20 25 30 35 40 45 50

55 60

0.994

0.992

0.990

0.988

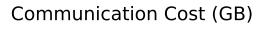
0.986

0.984

0.982

0.980

Accuracy

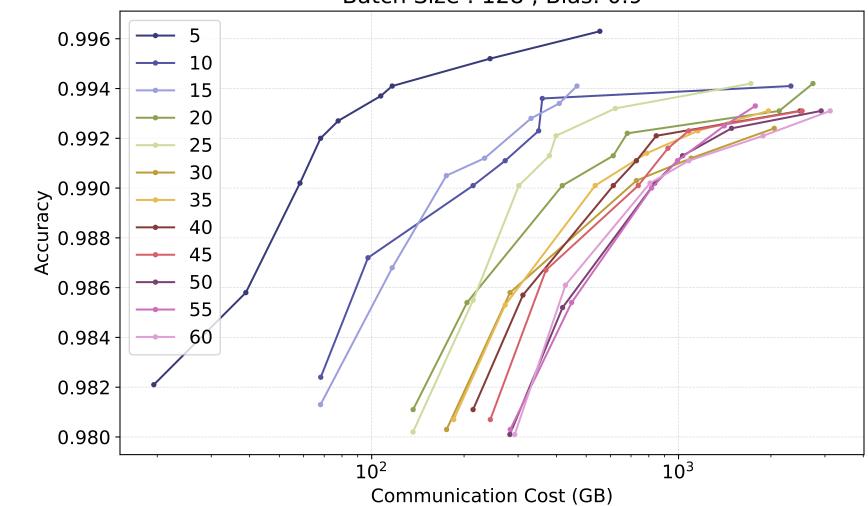


10³

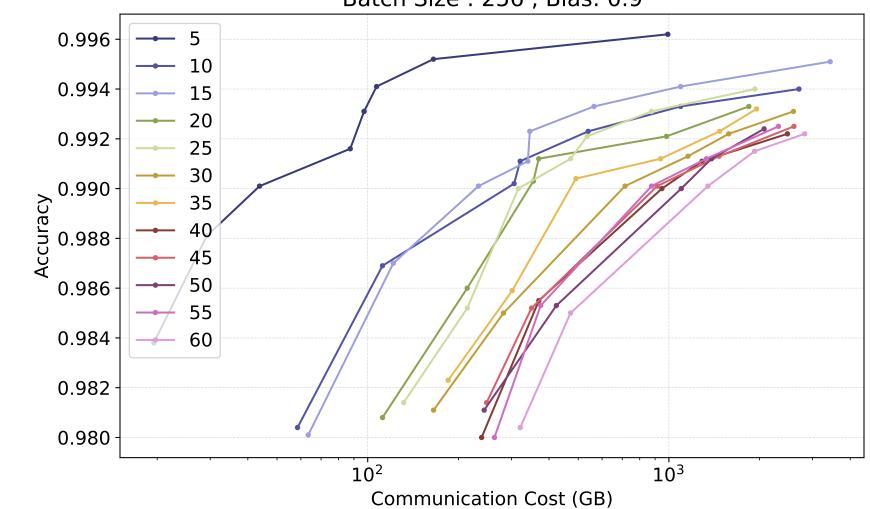
102

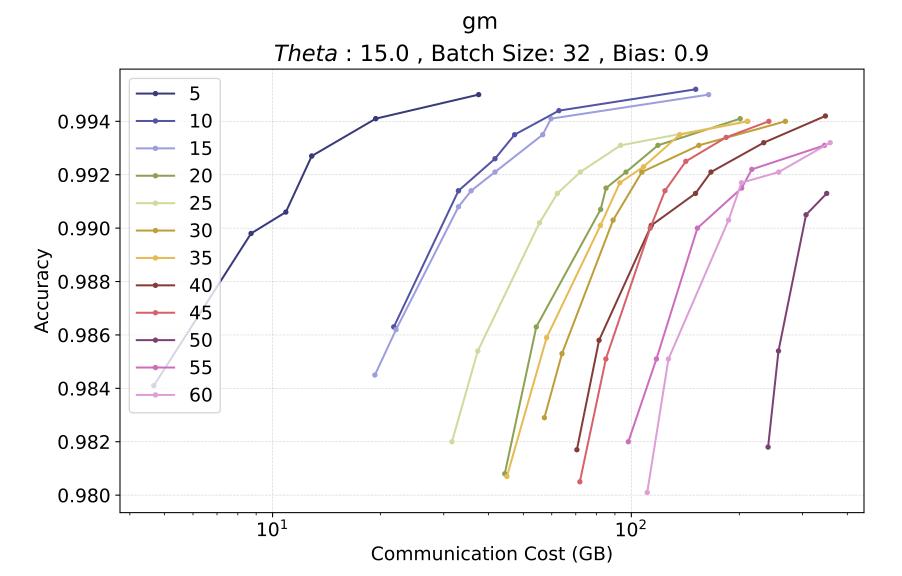
synchronous Batch Size: 64, Bias: 0.9 0.996 0.994 0.992 10 15 0.990 Accuracy 20 0.988 25 30 0.986 35 40 0.984 45 50 0.982 55 60 0.980 -103 10² Communication Cost (GB)

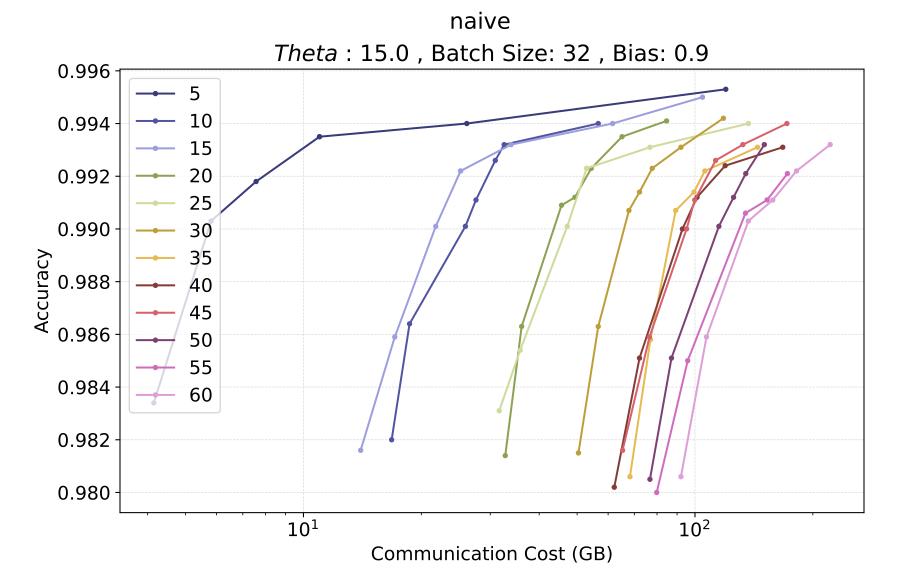
synchronous Batch Size: 128, Bias: 0.9



synchronous Batch Size : 256 , Bias: 0.9

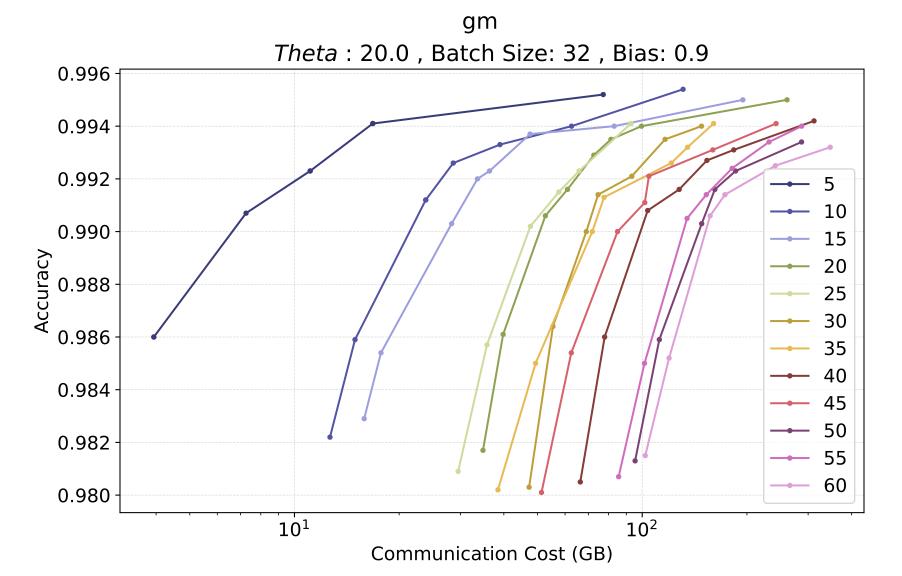






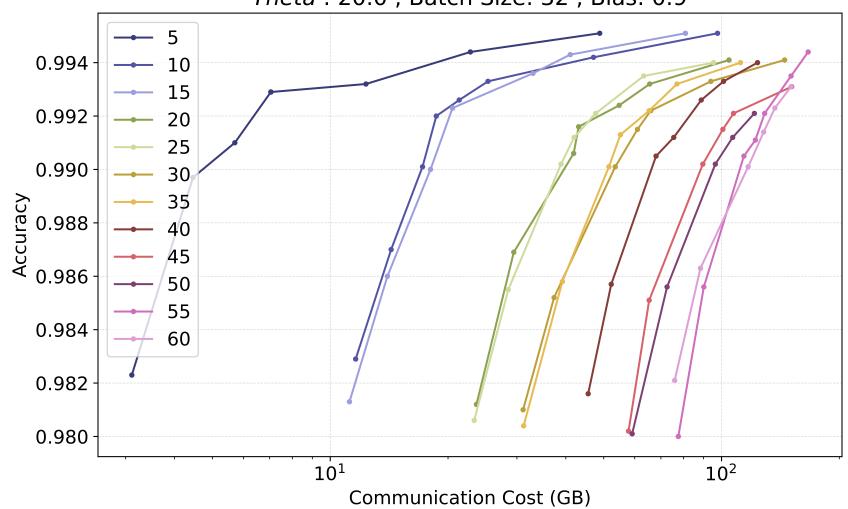
linear Theta: 15.0, Batch Size: 32, Bias: 0.9 0.996 -0.994 15 20 0.992 25 30 0.990 Accuracy 35 0.988 40 45 0.986 50 55 0.984 60 0.982 0.980 10^{1} 10² Communication Cost (GB)

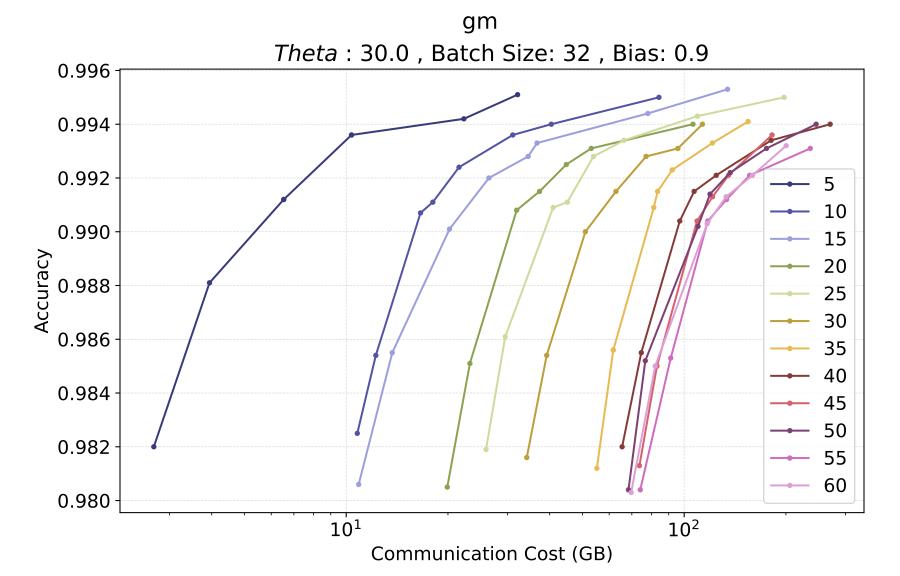
sketch Theta: 15.0, Batch Size: 32, Bias: 0.9 0.994 0.992 25 0.990 30 Accuracy 35 0.988 40 45 0.986 50 55 0.984 60 0.982 0.980 10^{1} 10² Communication Cost (GB)



naive Theta: 20.0, Batch Size: 32, Bias: 0.9 0.994 0.992 20 25 0.990 30 Accuracy 986.0 40 45 50 55 0.984 60 0.982 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 20 25 0.990 30 Accuracy 986.0 35 40 50 55 0.984 60 0.982 0.980 10^{1} 10² Communication Cost (GB)

linear Theta: 30.0, Batch Size: 32, Bias: 0.9 0.99610 0.994 15 20 0.992 25 30 0.990 Accuracy 35 0.988 45 0.986 50 55 0.984 60 0.982 0.980 10^{1} 10² Communication Cost (GB)

gm Theta: 50.0, Batch Size: 32, Bias: 0.9 0.994 0.992 20 25 0.990 30 Accuracy 0.988 40 45 0.986 50 55 0.984 60 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 0.992 25 Accuracy .0 .0 .0 .0 30 35 40 45 50 0.986 55 0.984 60 0.982 0.980 10^1 Communication Cost (GB)

linear Theta: 50.0, Batch Size: 32, Bias: 0.9 0.994 0.992 25 0.990 30 Accuracy 35 0.988 40 0.986 50 55 0.984 60 0.982 0.980 10^{1} Communication Cost (GB)

gm Theta: 75.0, Batch Size: 32, Bias: 0.9 0.994 0.992 25 0.990 30 Accuracy 0.988 40 0.986 50 55 0.984 60 0.982 0.980 10^{1} 10² Communication Cost (GB)

naive Theta: 75.0, Batch Size: 32, Bias: 0.9 0.994 0.992 25 0.990 30 Accuracy 35 0.988 40 0.986 50 55 0.984 60 0.982 0.980 10^{1} Communication Cost (GB)

linear Theta: 75.0, Batch Size: 32, Bias: 0.9 15 20 25 30 35 40 45 50 55 60 10^{1} Communication Cost (GB)

0.994

0.992

0.990

0.988

0.986

0.984

0.982

0.980

Accuracy

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

naive Theta: 100.0, Batch Size: 32, Bias: 0.9 0.994 0.992 20 25 0.990 30 Accuracy 35 0.988 40 45 0.986 50 55 0.984 60 0.982 0.980 10^{1} Communication Cost (GB)

linear Theta: 100.0, Batch Size: 32, Bias: 0.9 0.994 0.992 25 0.990 30 Accuracy 35 0.988 40 0.986 50 55 0.984 60 0.982 0.980 10^{1} Communication Cost (GB)

sketch Theta: 100.0, Batch Size: 32, Bias: 0.9 0.994 0.992 20 25 0.990 30 Accuracy 35 0.988 40 45 0.986 50 55 0.984 60 0.982 0.980 100 10^{1} Communication Cost (GB)