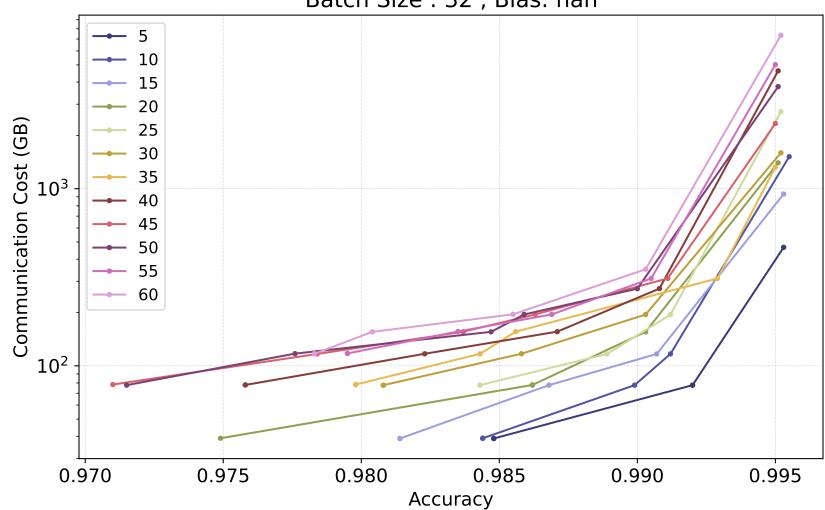
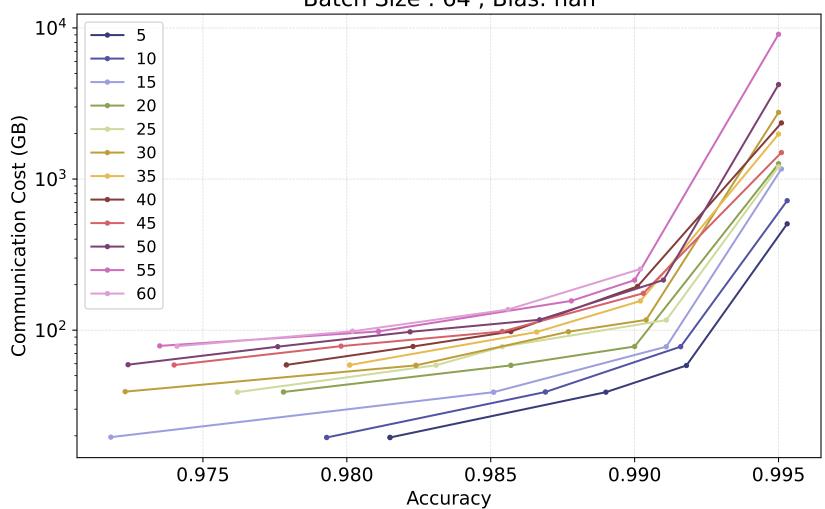
synchronous Batch Size : 32 , Bias: nan

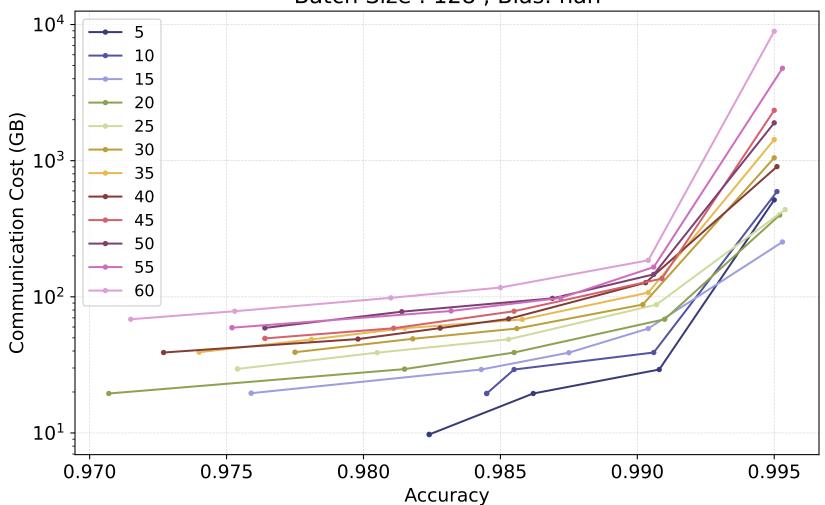


synchronous Batch Size: 64, Bias: nan

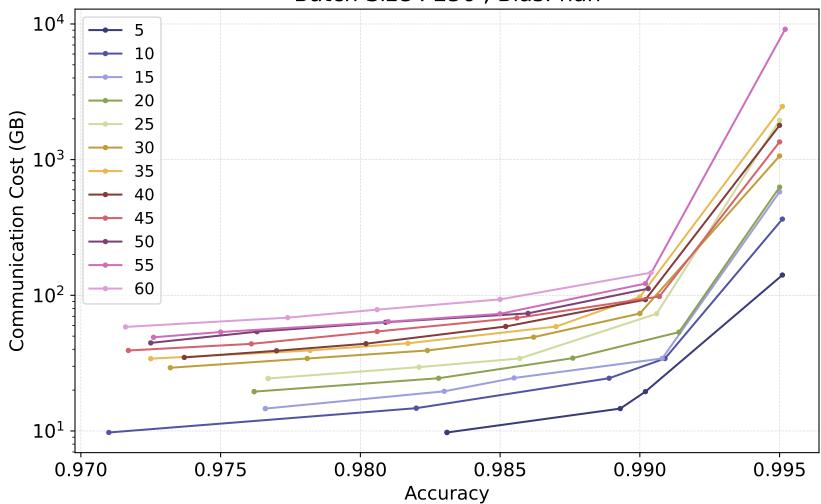


synchronous

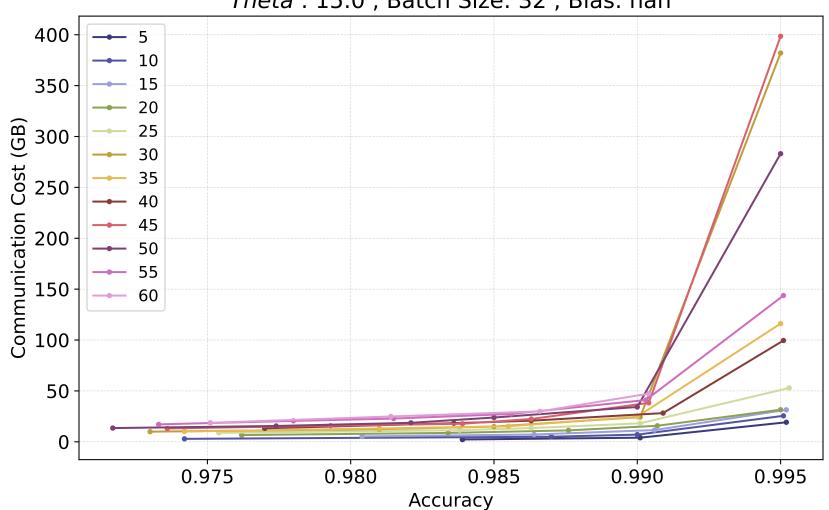
Batch Size: 128, Bias: nan



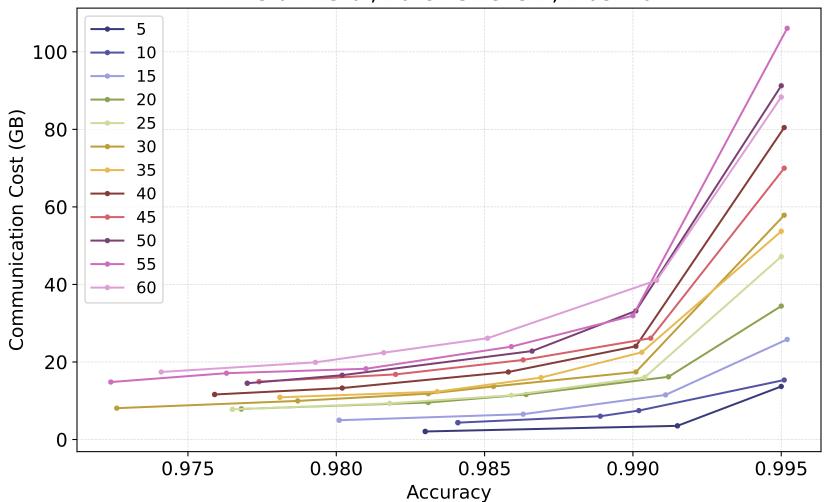
synchronous Batch Size : 256 , Bias: nan



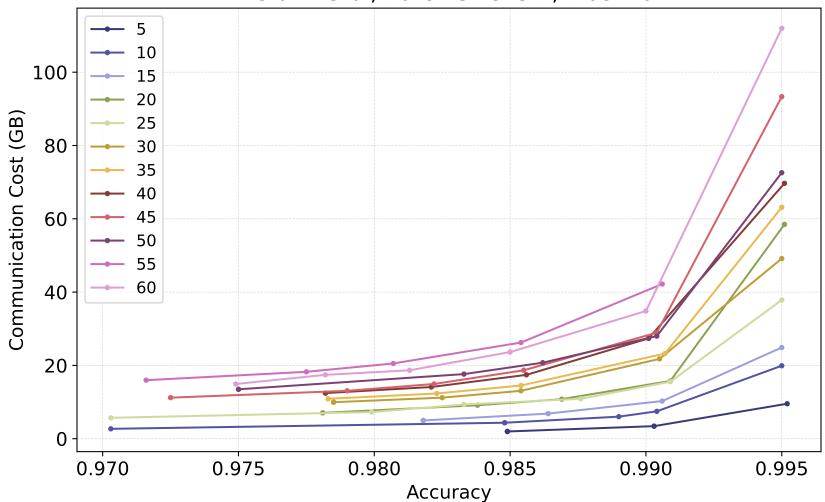
gm *Theta*: 15.0 , Batch Size: 32 , Bias: nan



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

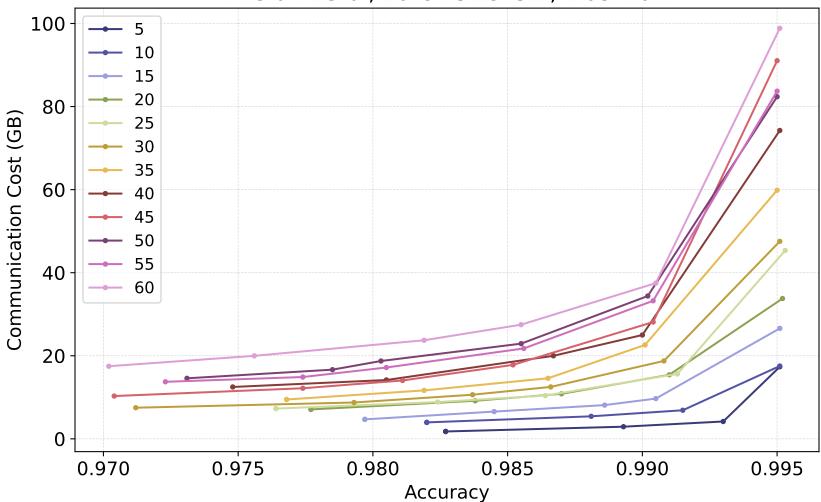


linear
Theta: 15.0, Batch Size: 32, Bias: nan

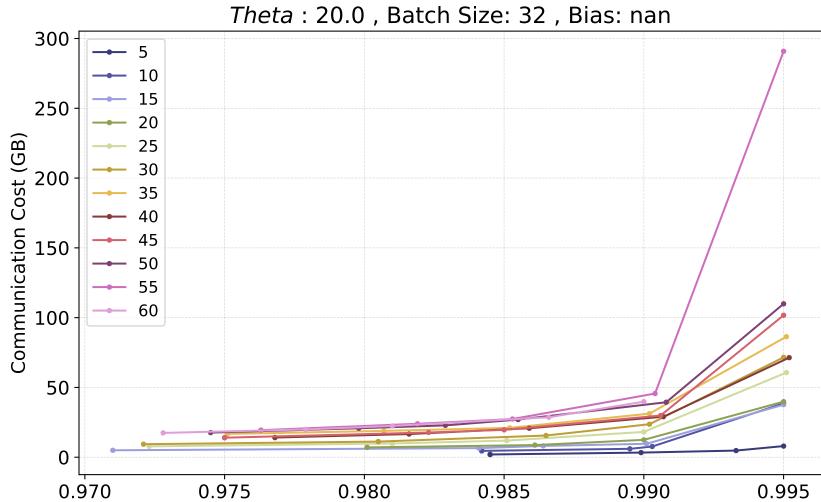


sketch

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

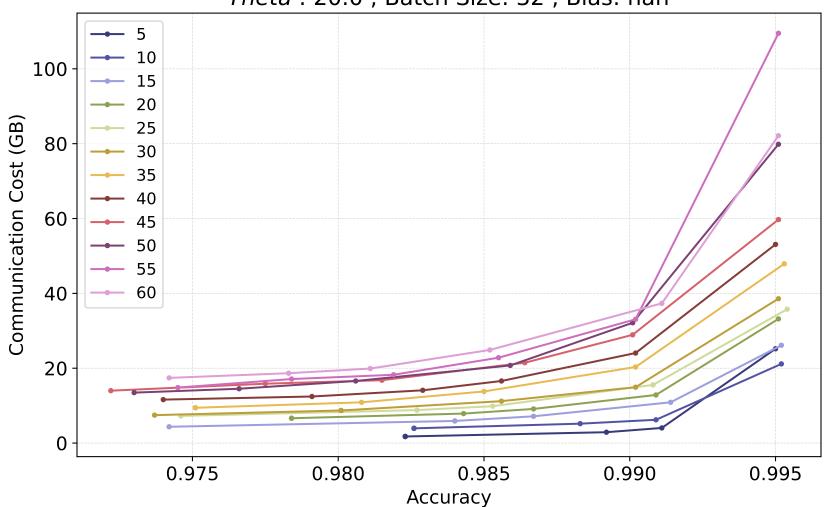


gm



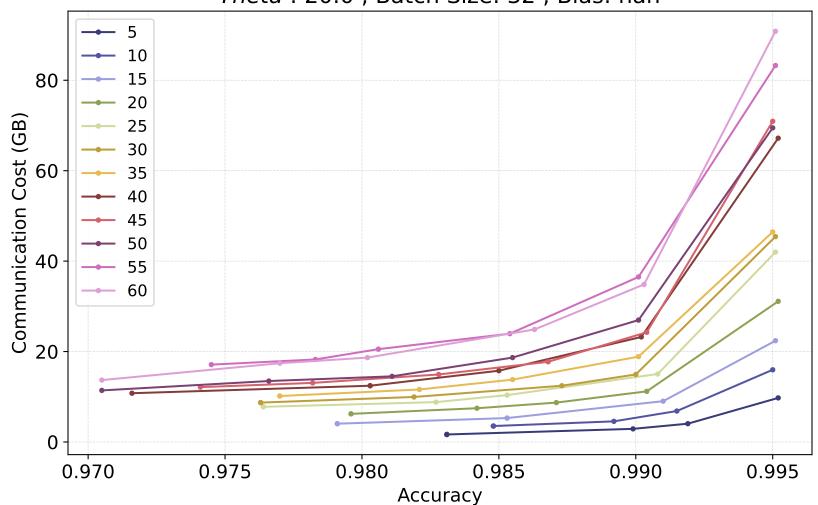
Accuracy

naive Theta: 20.0, Batch Size: 32, Bias: nan



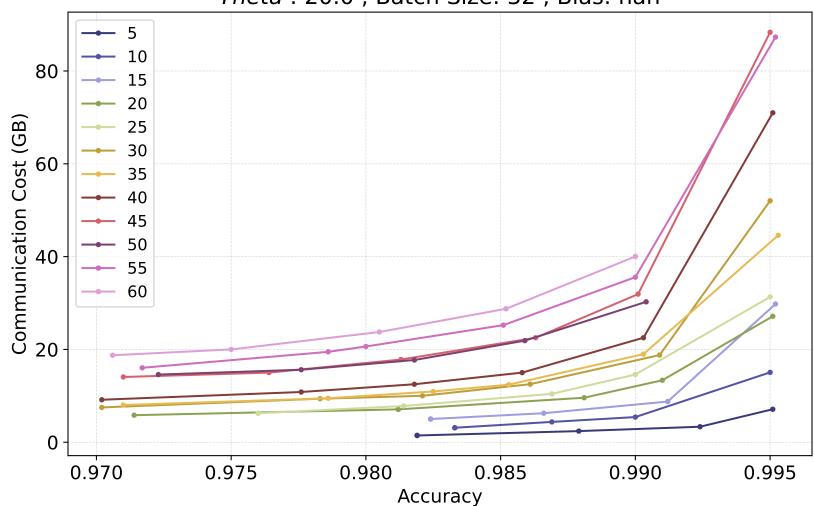
linear

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

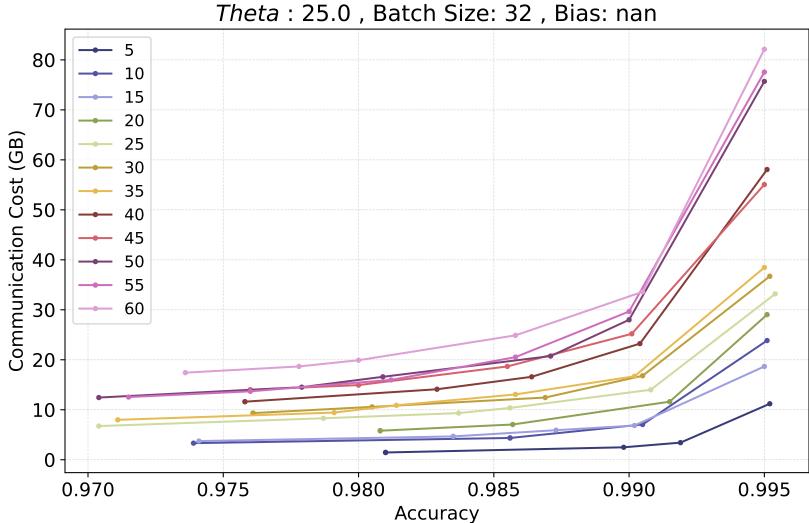


sketch

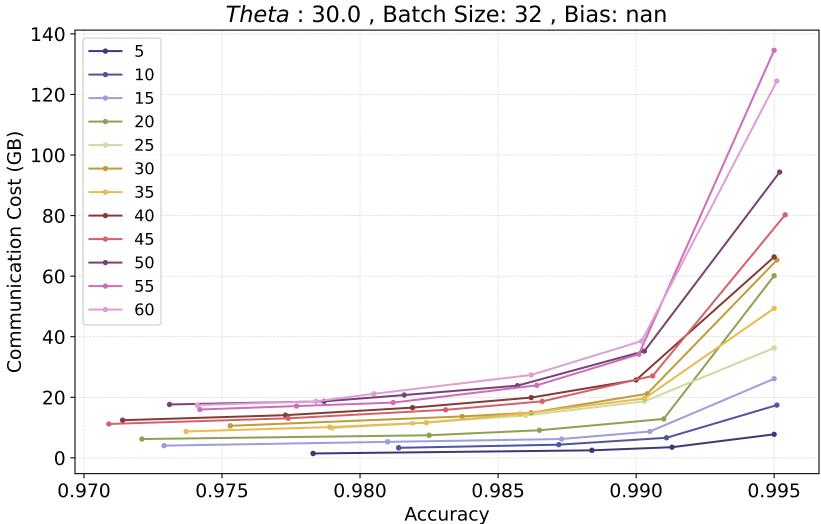
Theta: 20.0, Batch Size: 32, Bias: nan



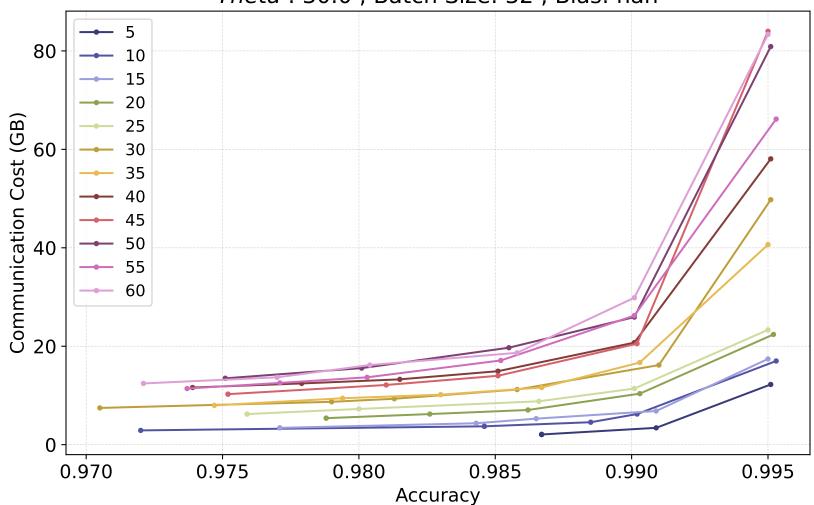
naive



gm *Theta* : 30.0 . Batch :

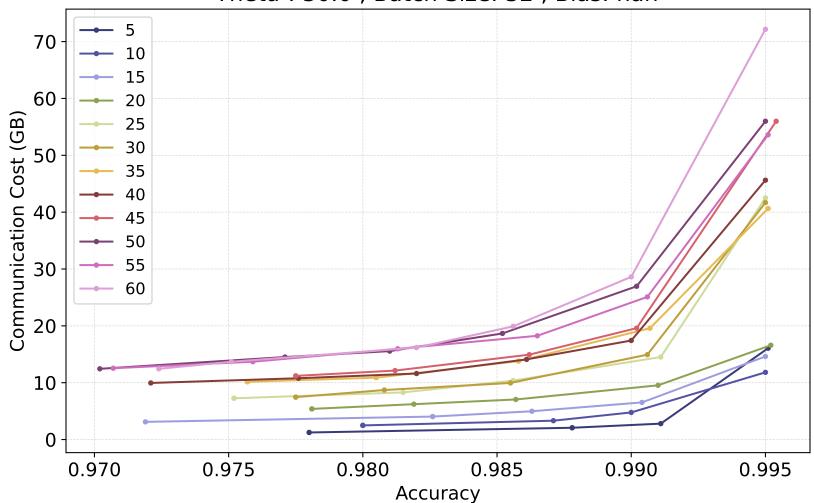


naive Theta: 30.0, Batch Size: 32, Bias: nan



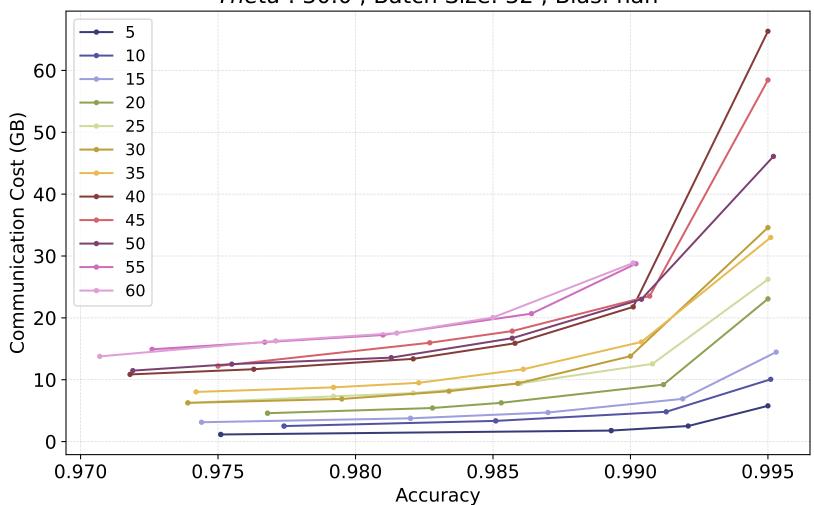
linear

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

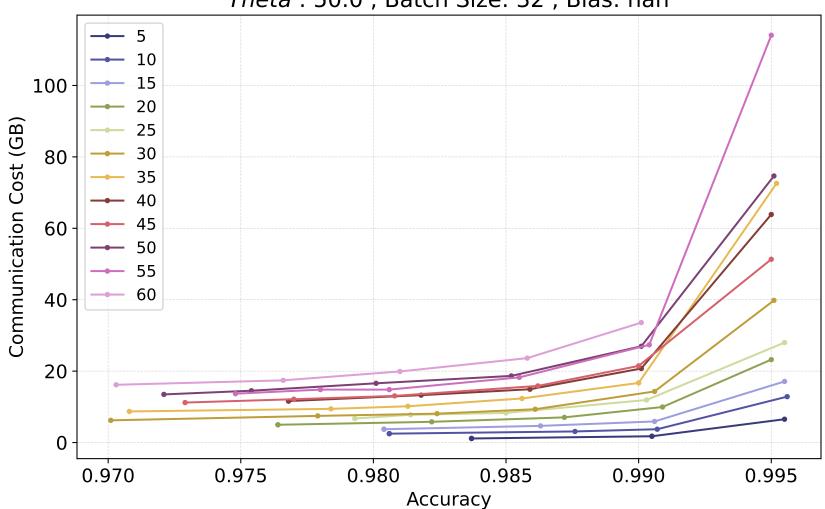


sketch

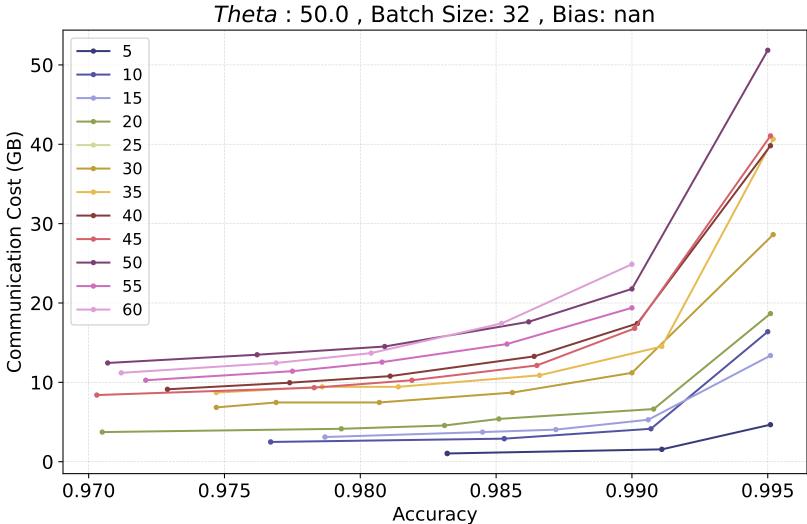
Theta: 30.0, Batch Size: 32, Bias: nan



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

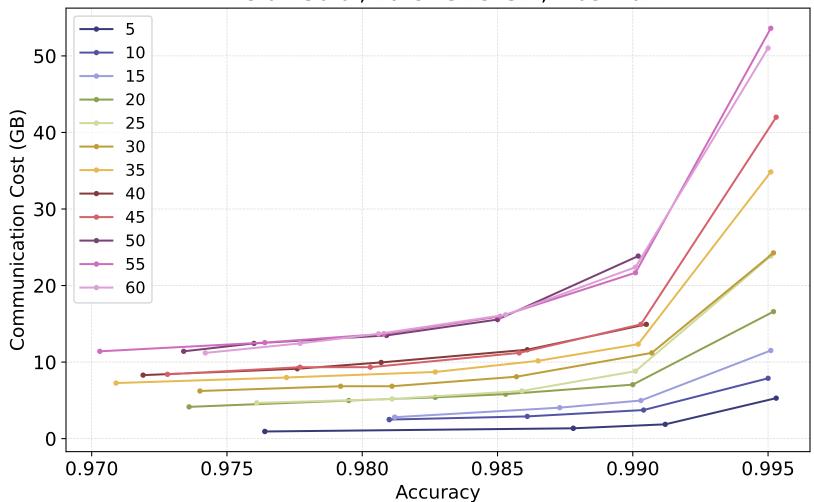


naive

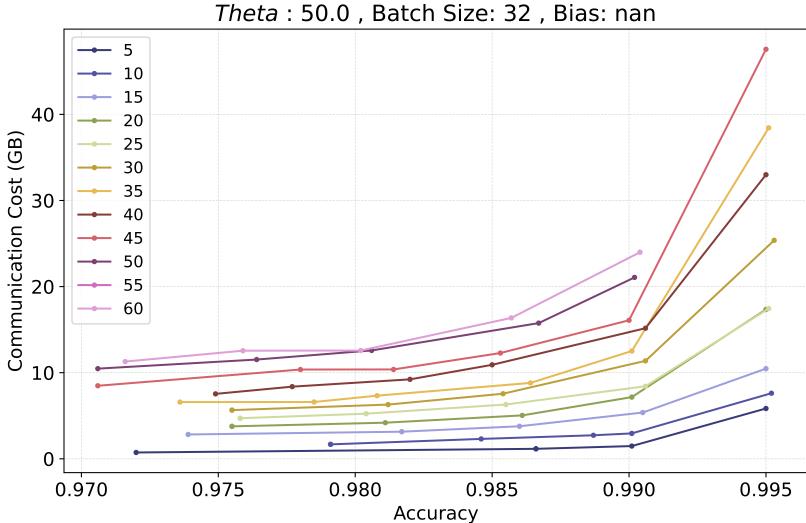


linear

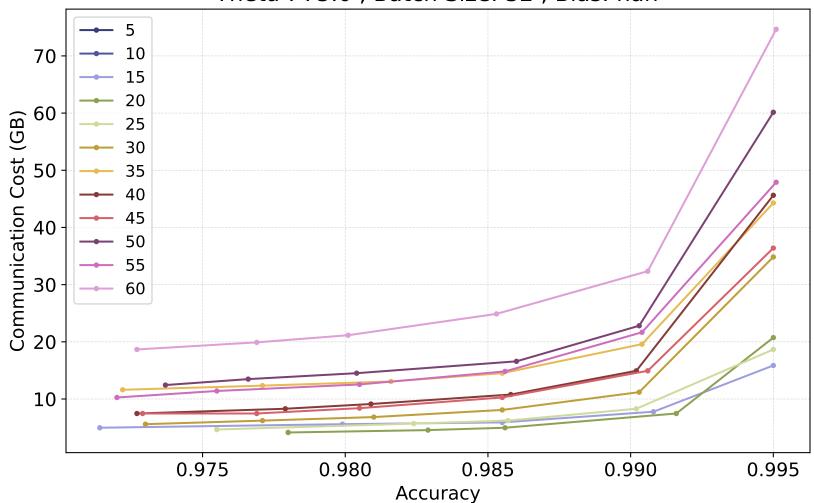
Theta: 50.0, Batch Size: 32, Bias: nan



sketch

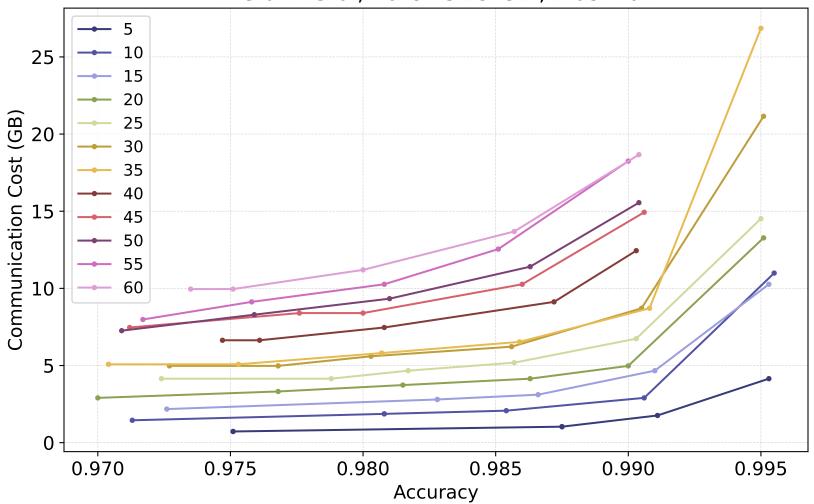


gm *Theta*: 75.0 , Batch Size: 32 , Bias: nan

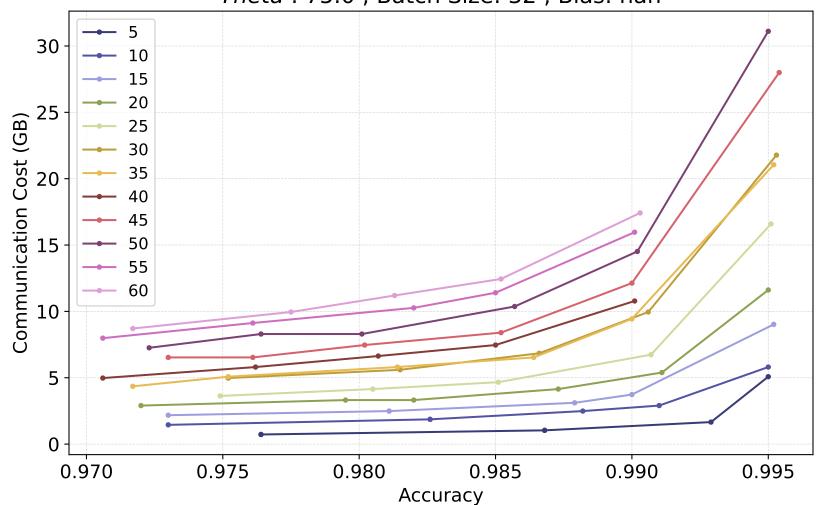


naive



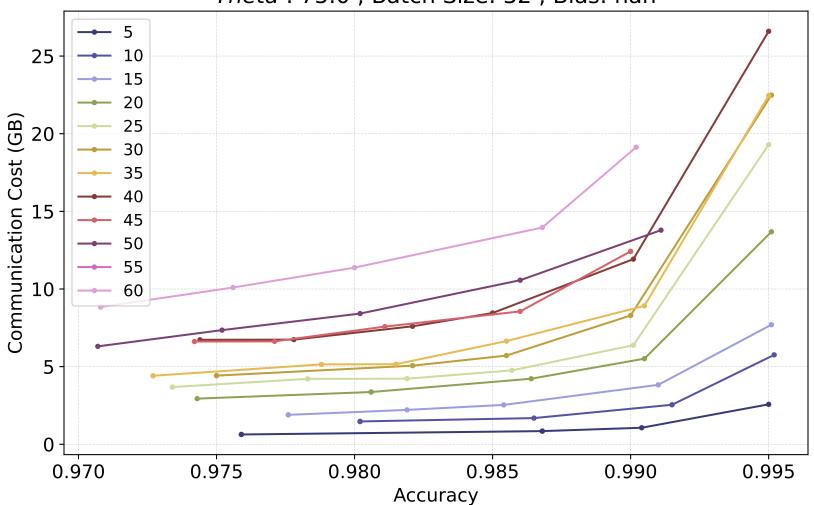


linear
Theta: 75.0, Batch Size: 32, Bias: nan

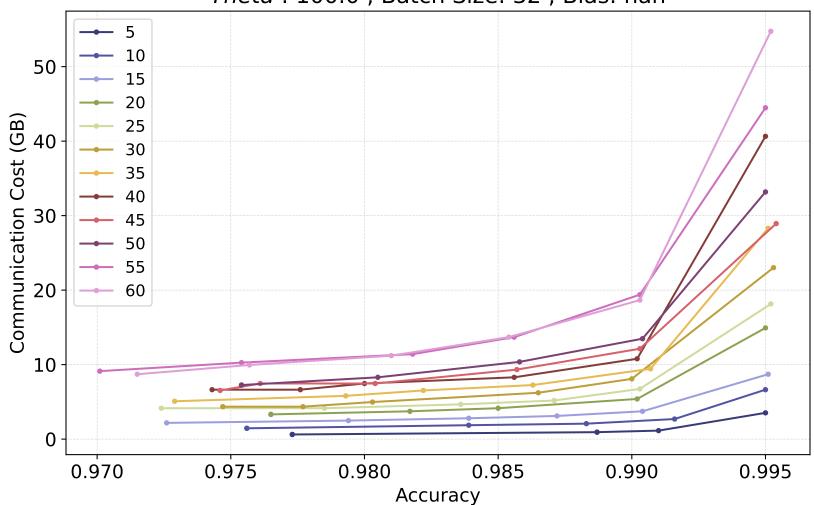


sketch

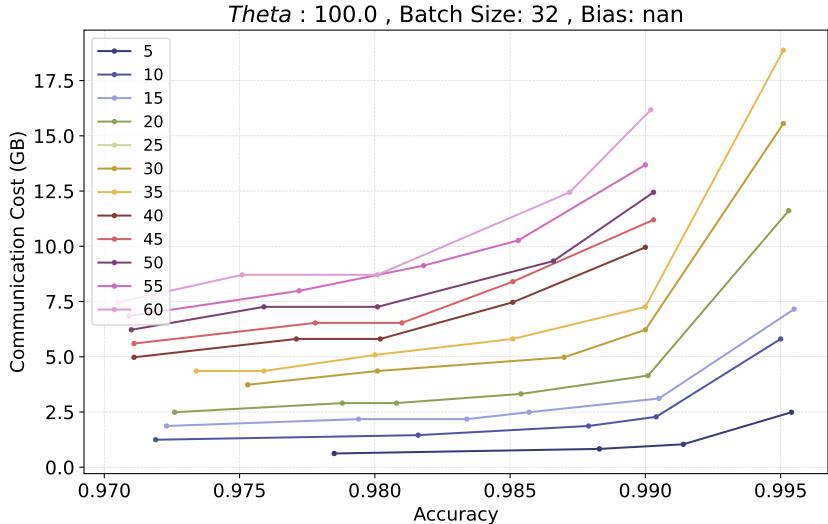




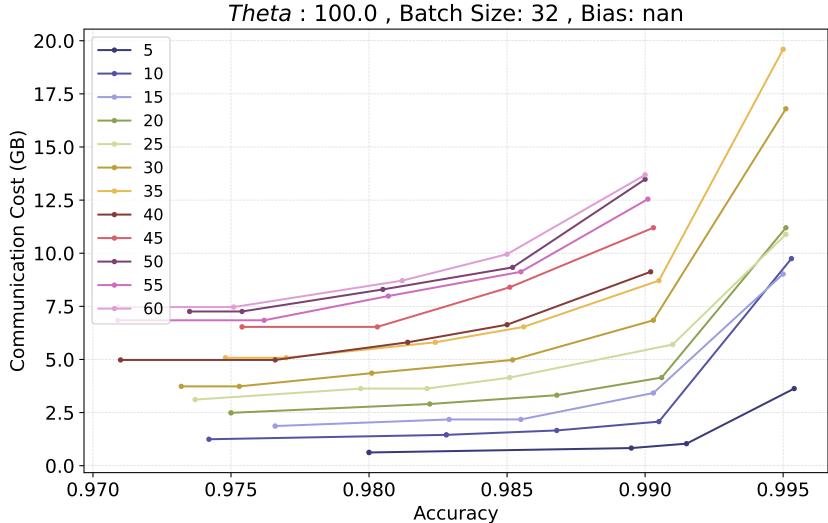
gm *Theta*: 100.0, Batch Size: 32, Bias: nan



naive

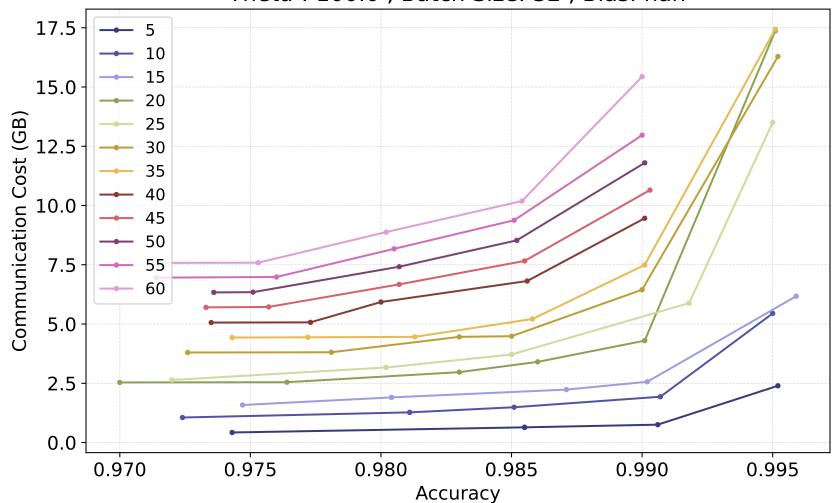


linear

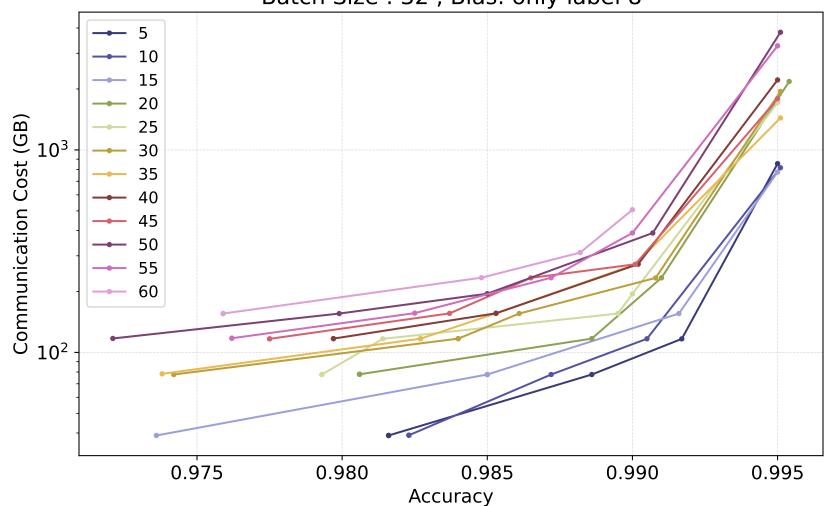


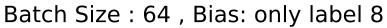
sketch

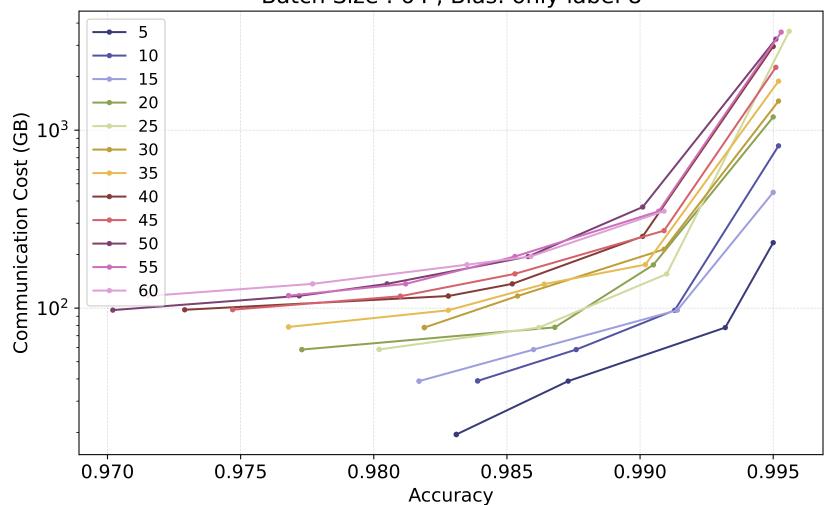
Theta: 100.0, Batch Size: 32, Bias: nan



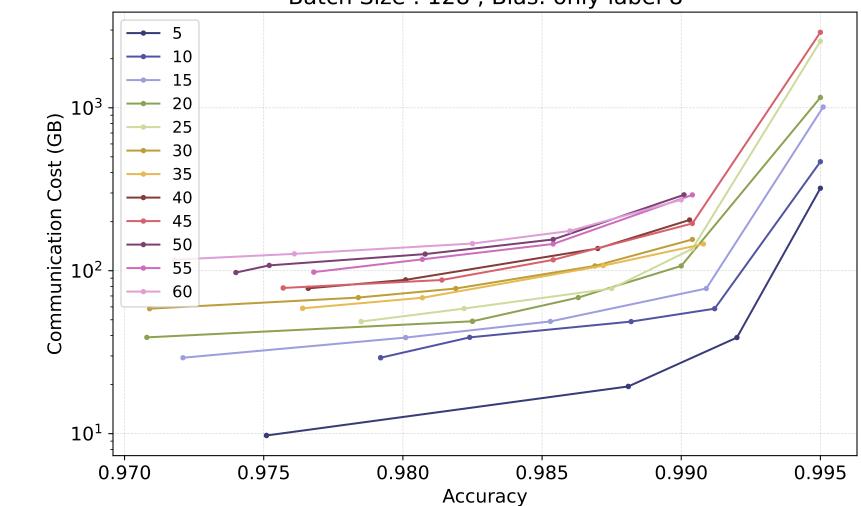
Batch Size: 32, Bias: only label 8

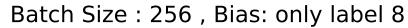


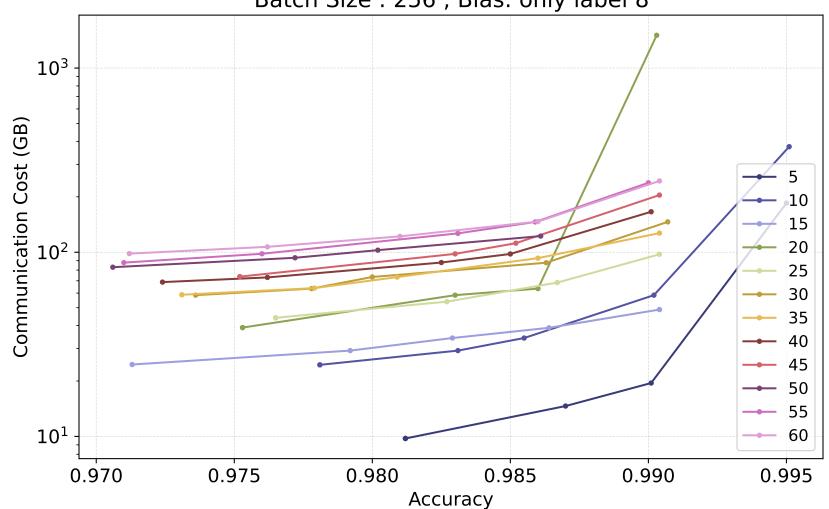




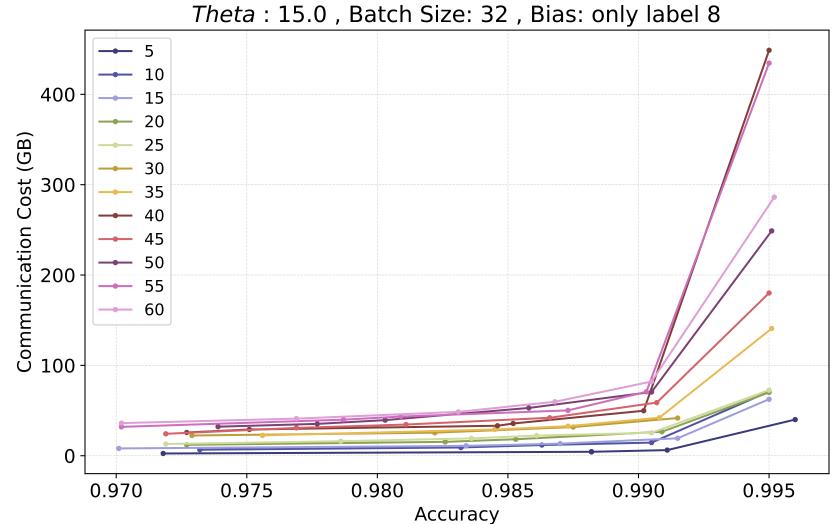
Batch Size: 128, Bias: only label 8



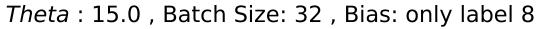


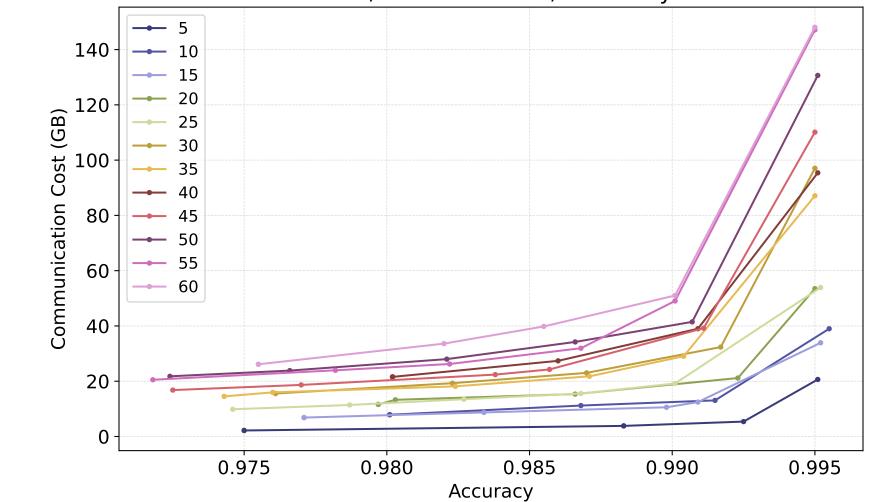


gm

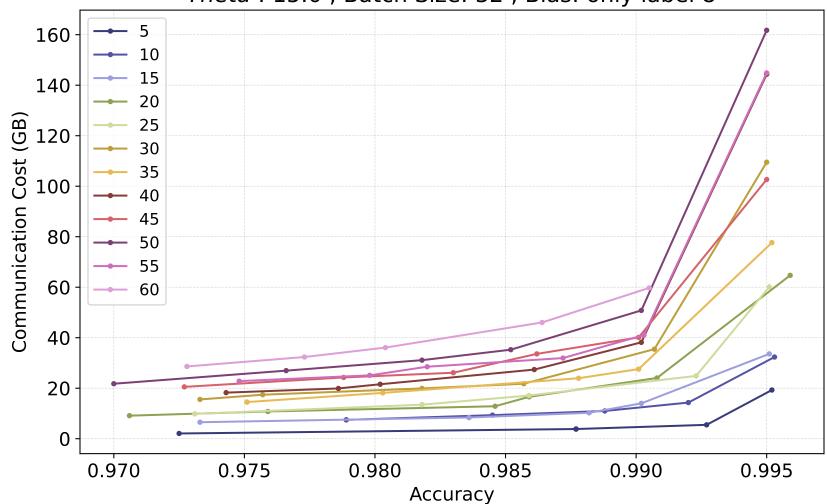


naive

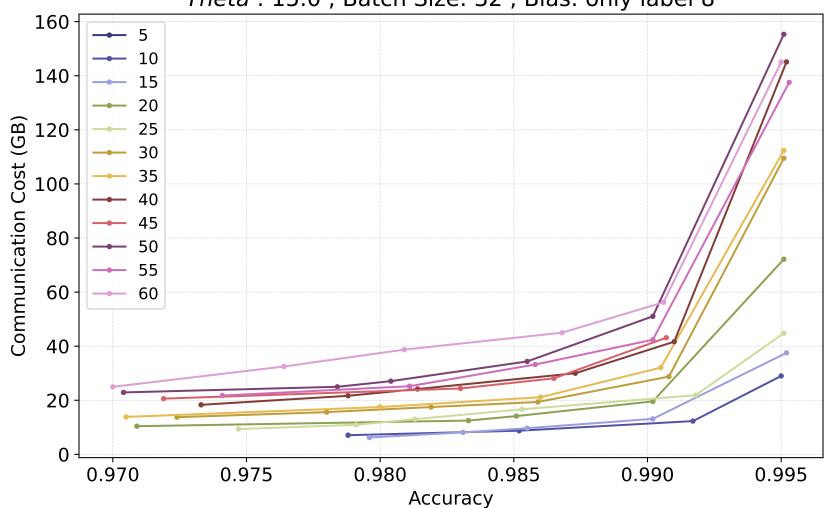




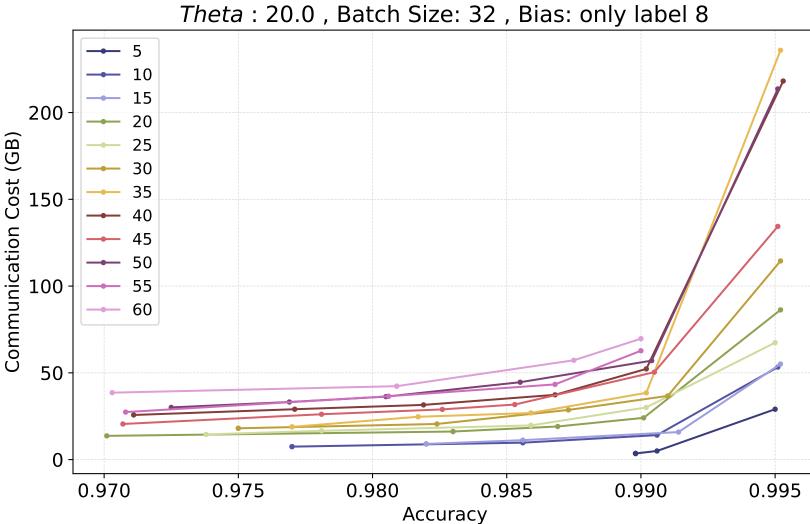
linear
Theta: 15.0, Batch Size: 32, Bias: only label 8



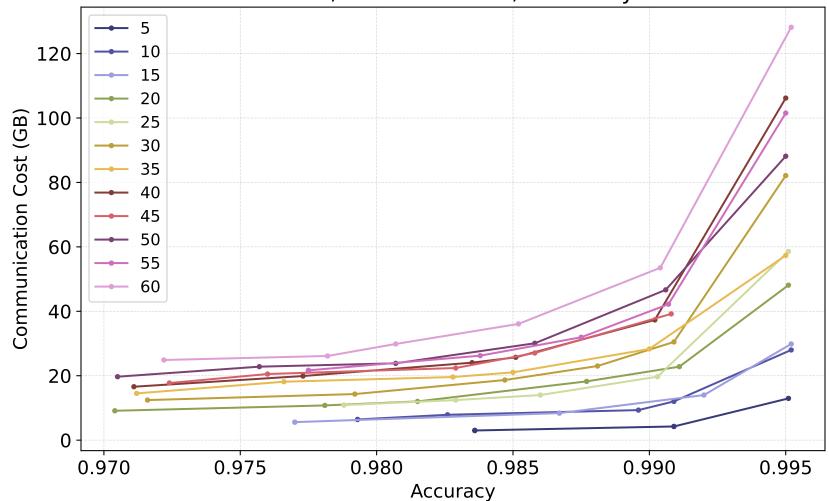
sketch



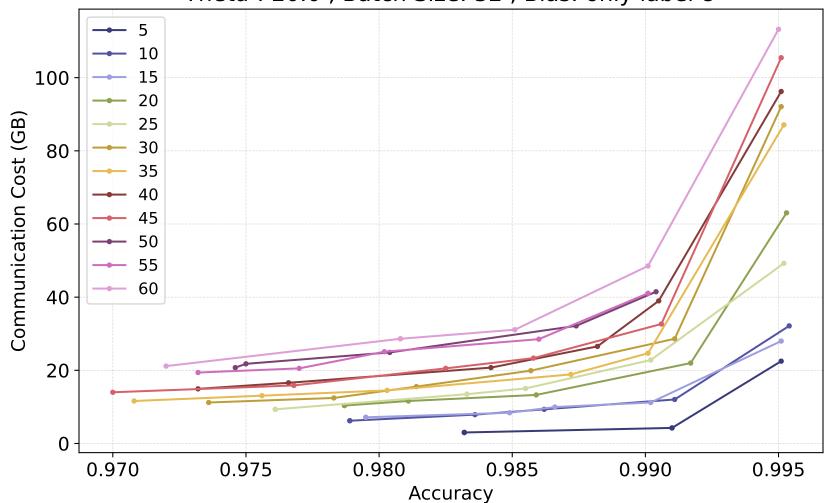
gm



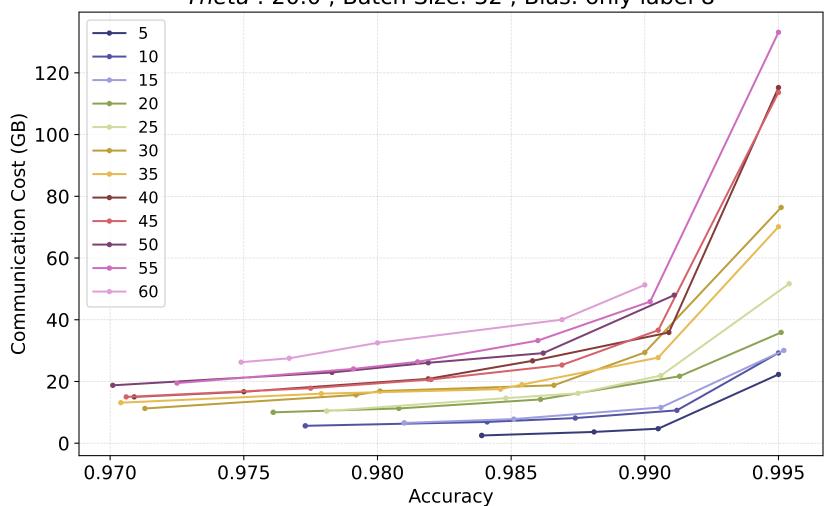
naive



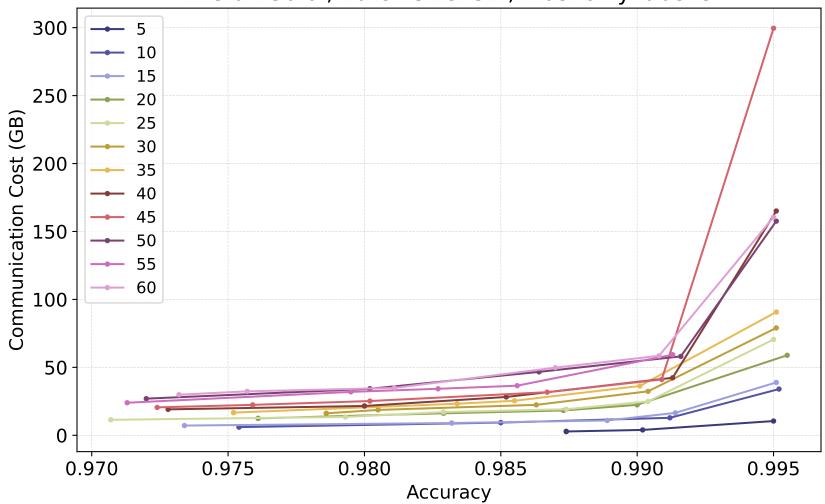
linear



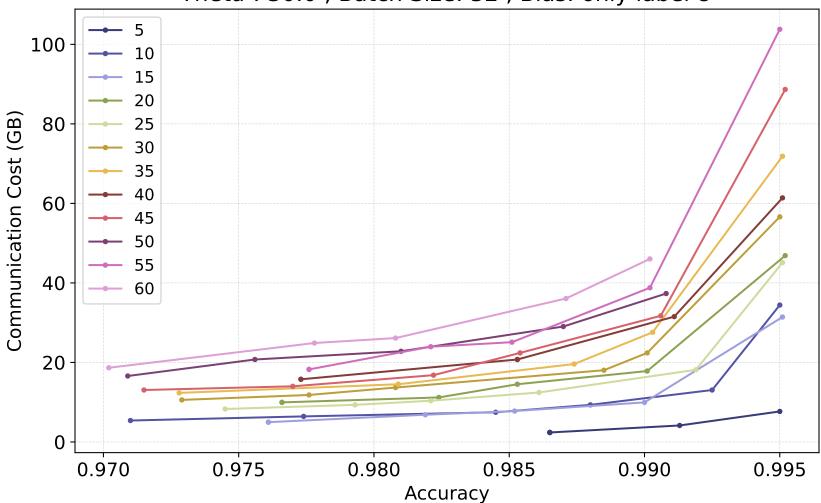
sketch



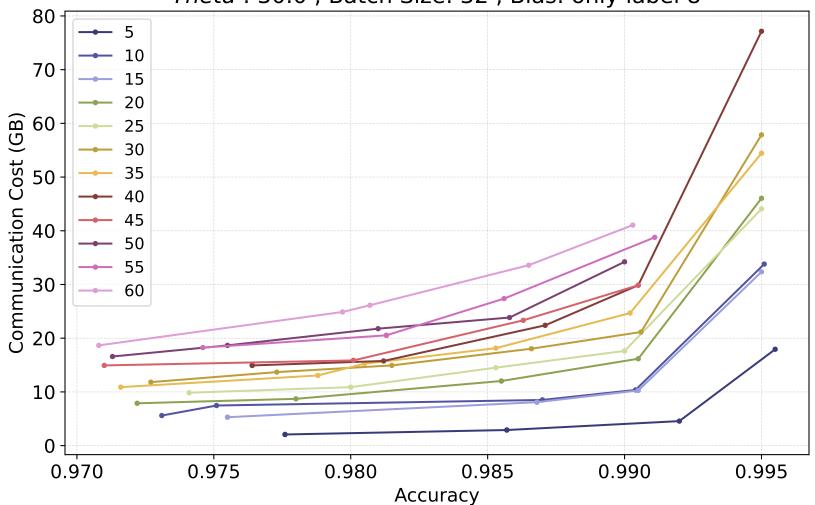
gm



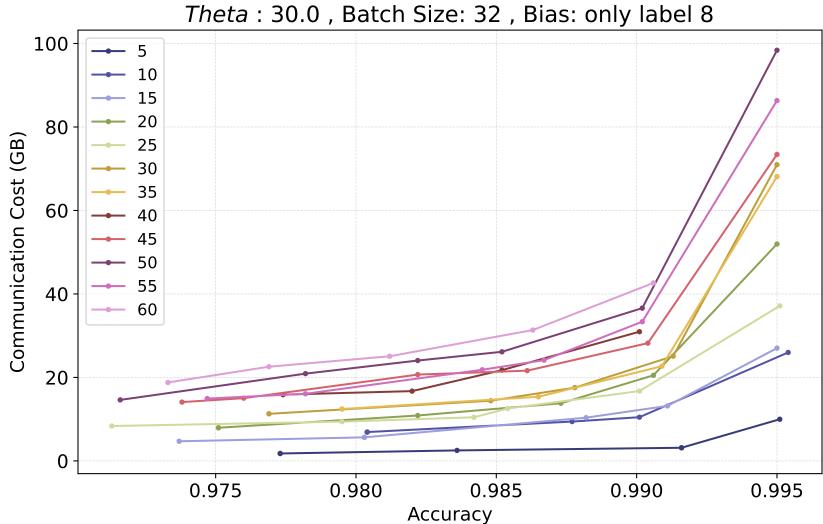
naive



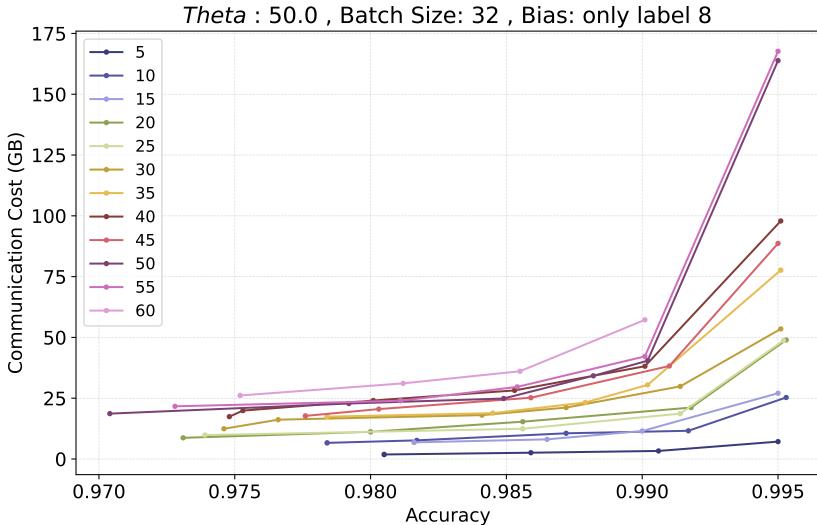
linear



sketch

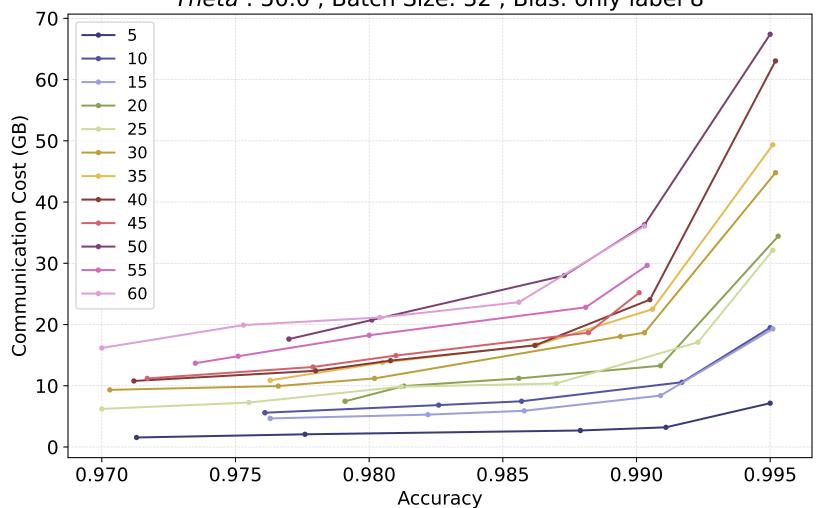


gm



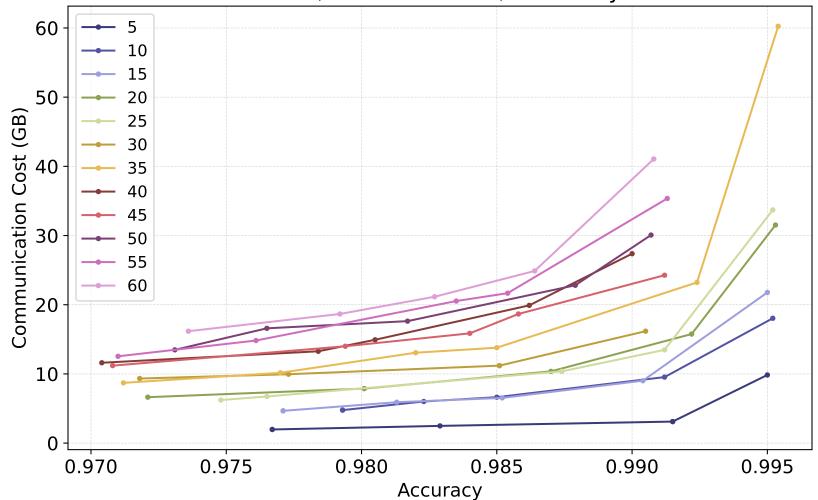
naive



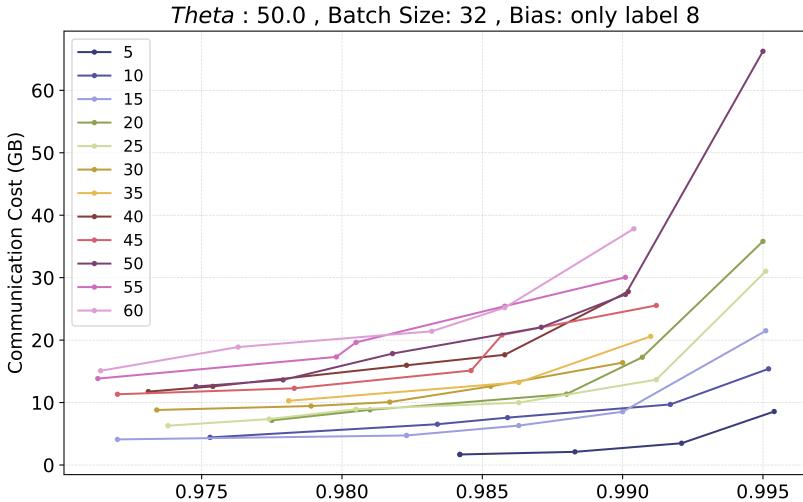


linear



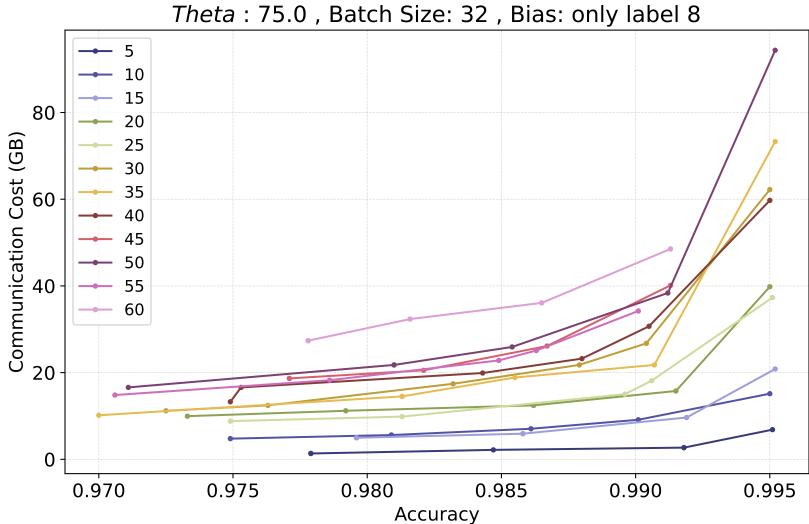


sketch

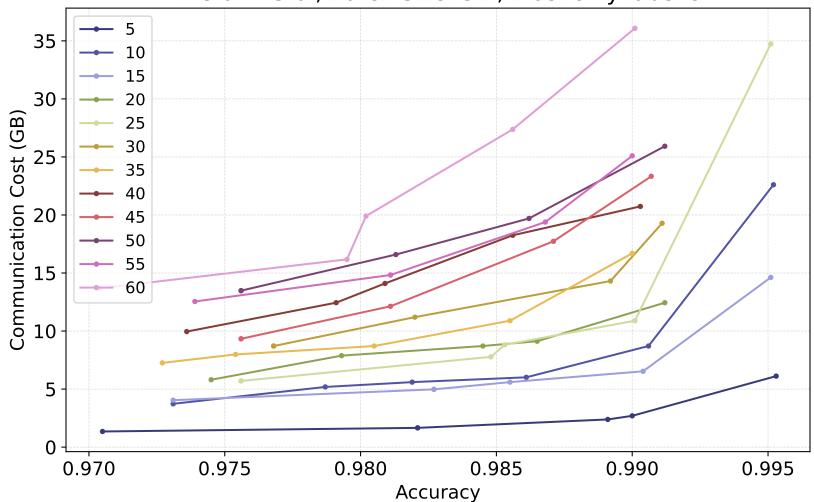


Accuracy

gm

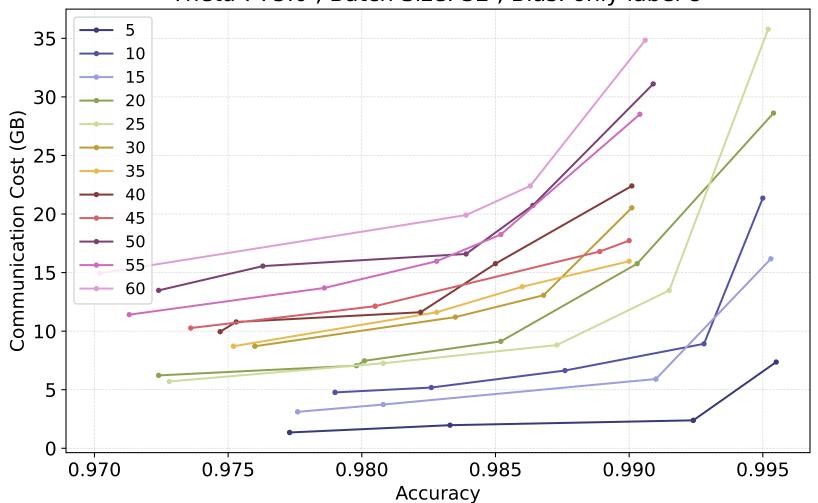


naive

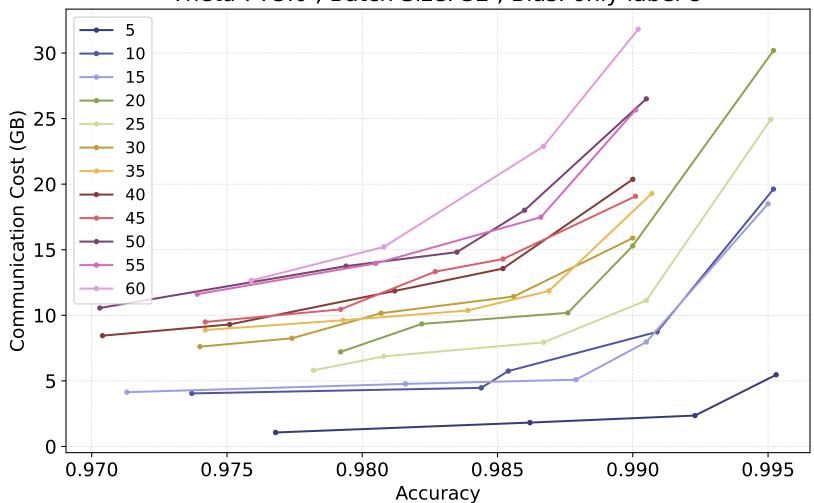


linear

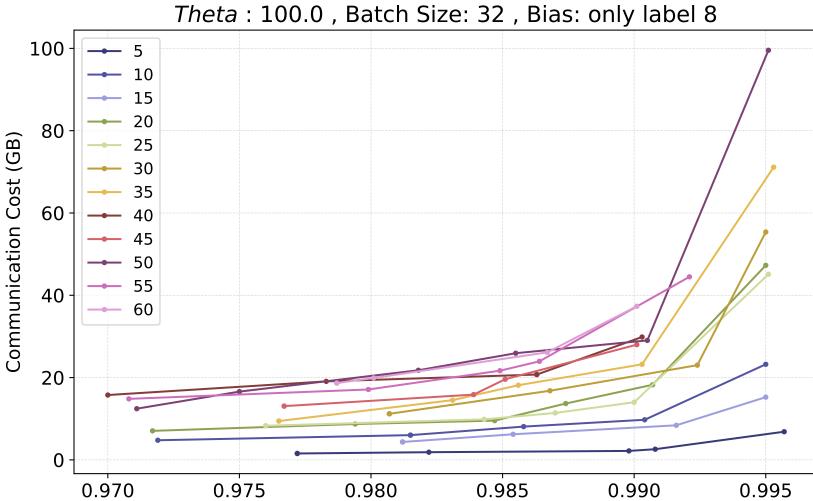




sketch

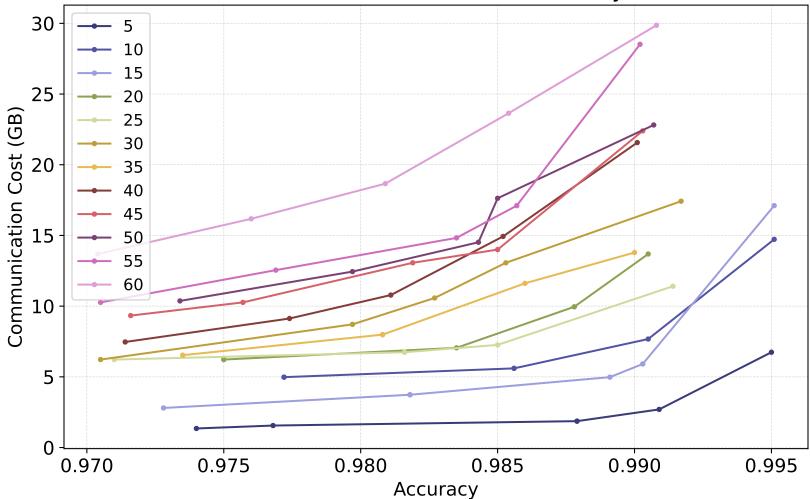


gm

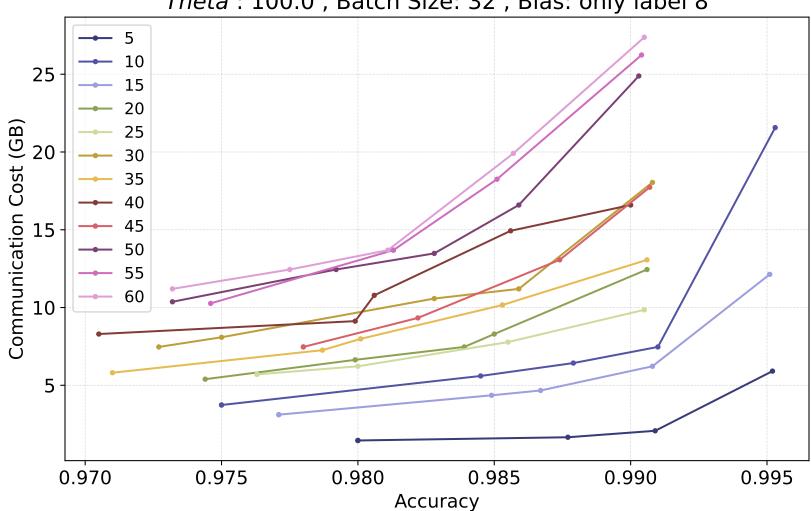


Accuracy

naive

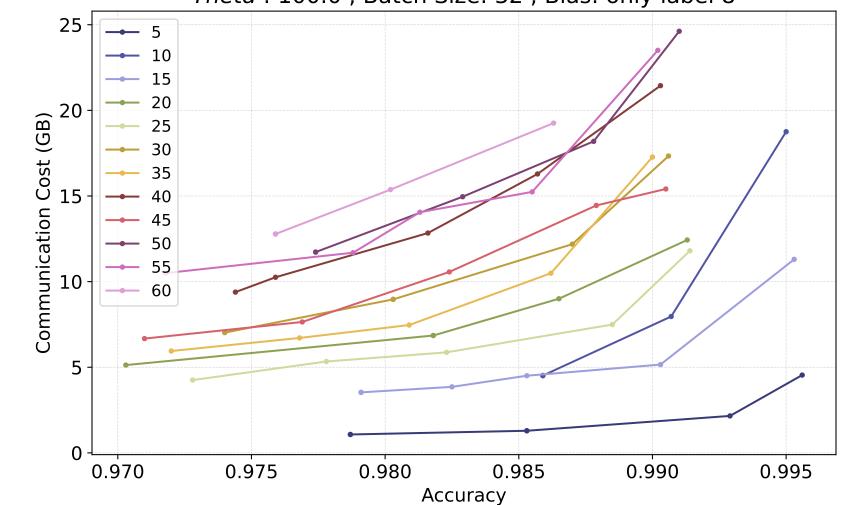


linear
Theta: 100.0, Batch Size: 32, Bias: only label 8

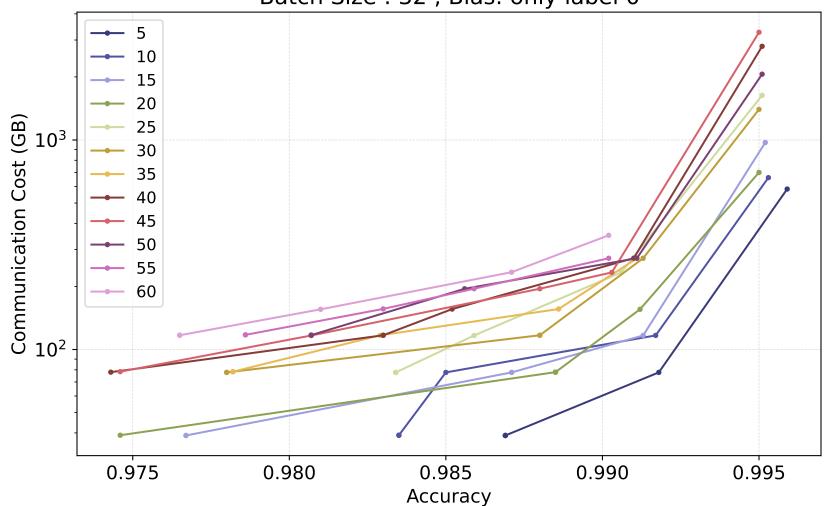


sketch

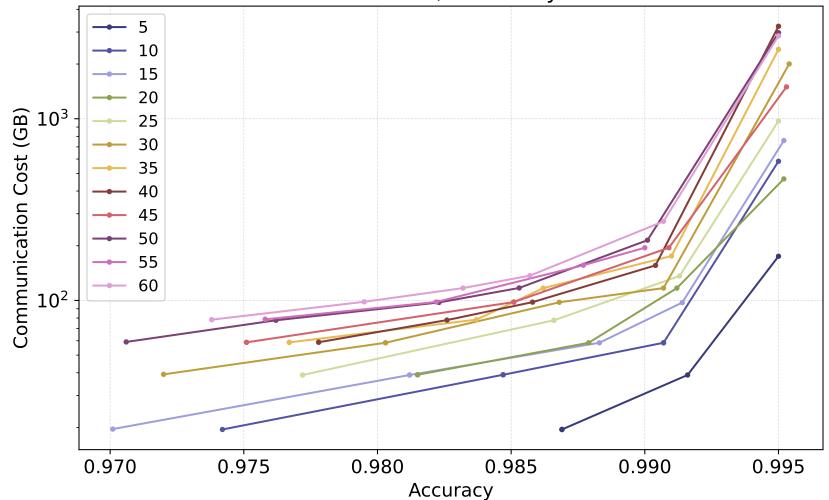




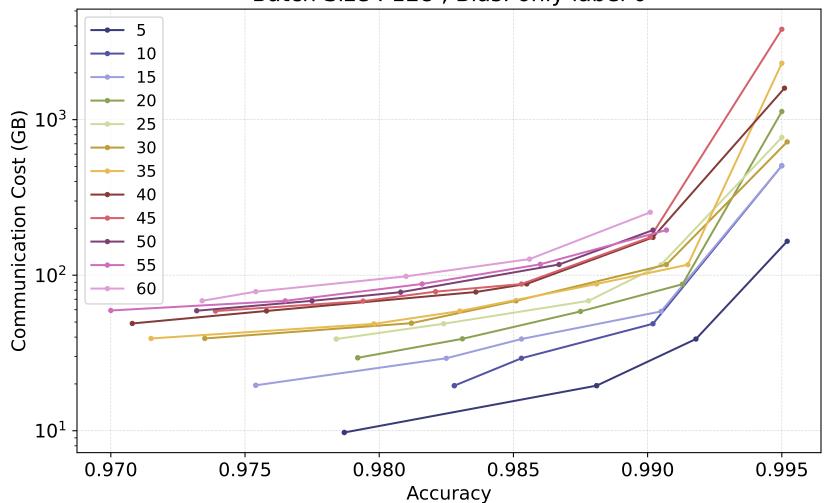
Batch Size: 32, Bias: only label 0



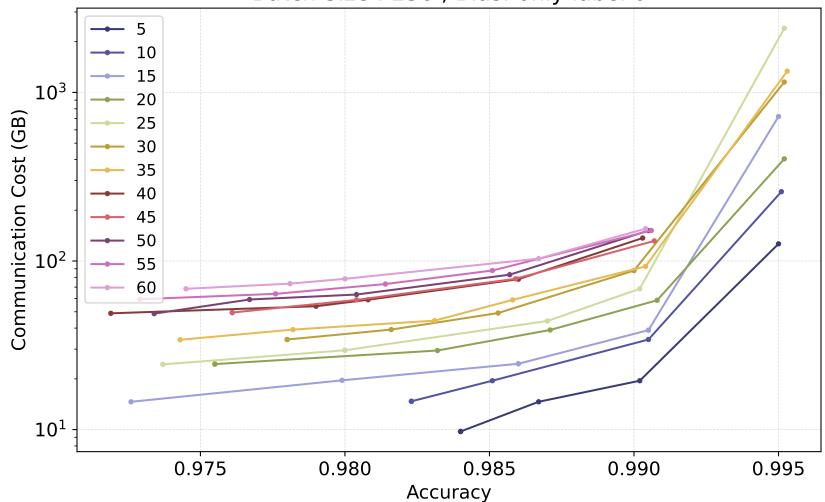
Batch Size: 64, Bias: only label 0



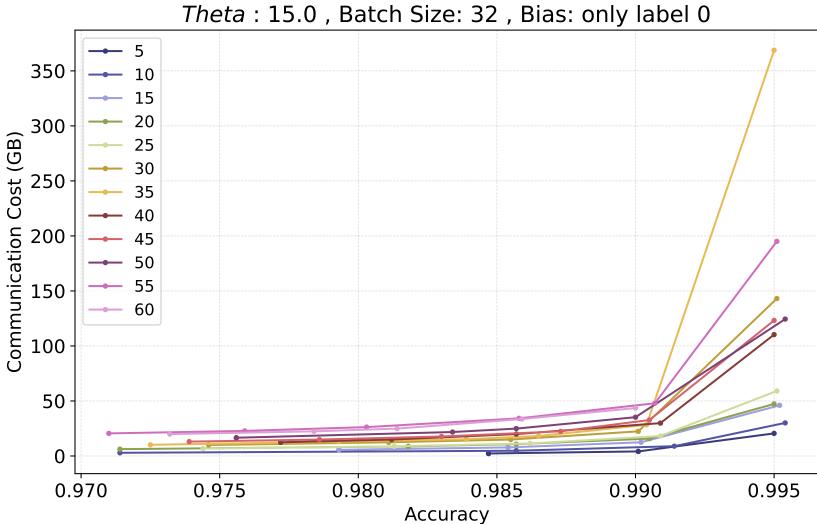
Batch Size: 128, Bias: only label 0



Batch Size: 256, Bias: only label 0

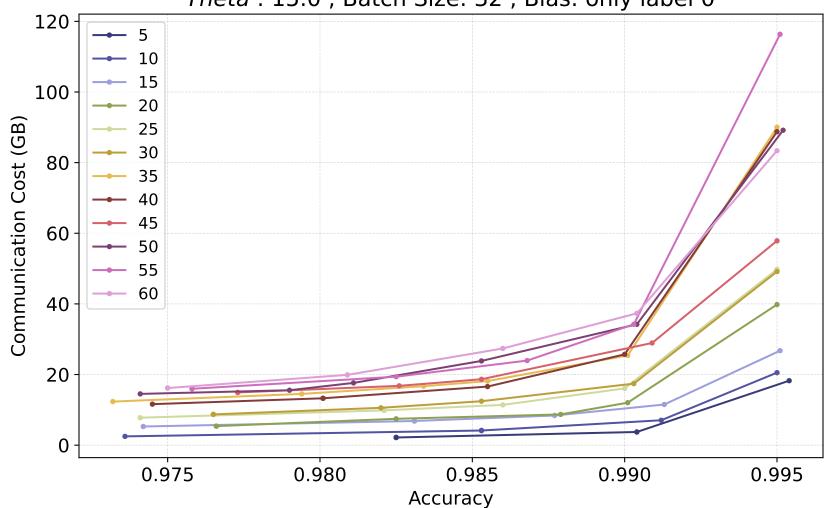


gm

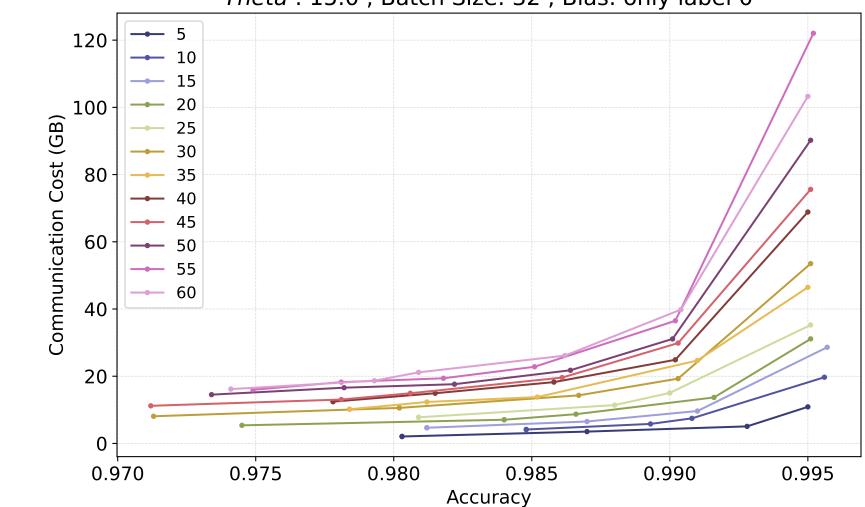


naive

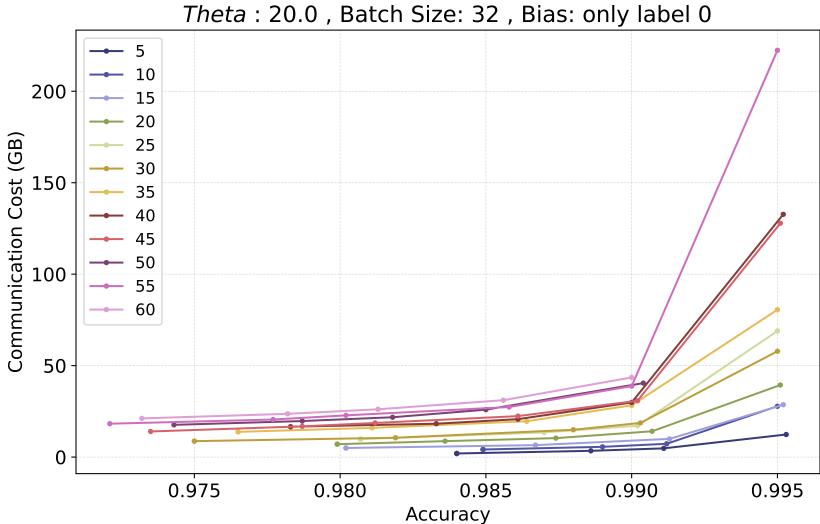




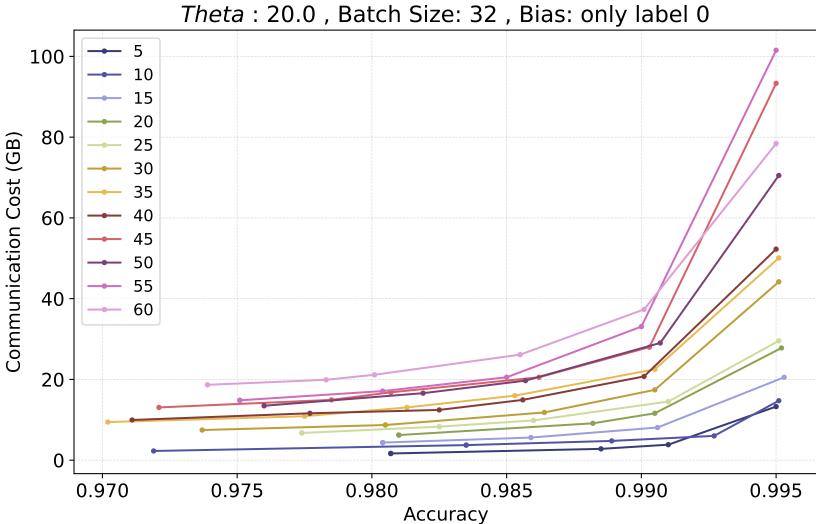
linear
Theta: 15.0, Batch Size: 32, Bias: only label 0



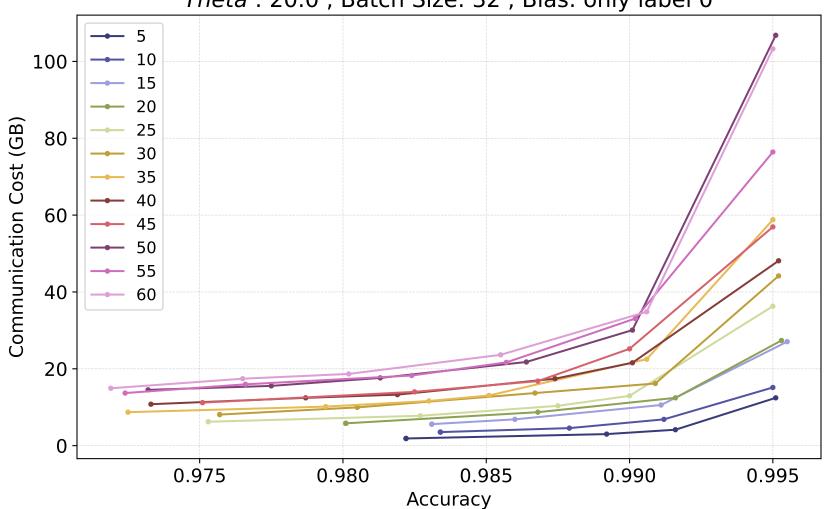
gm



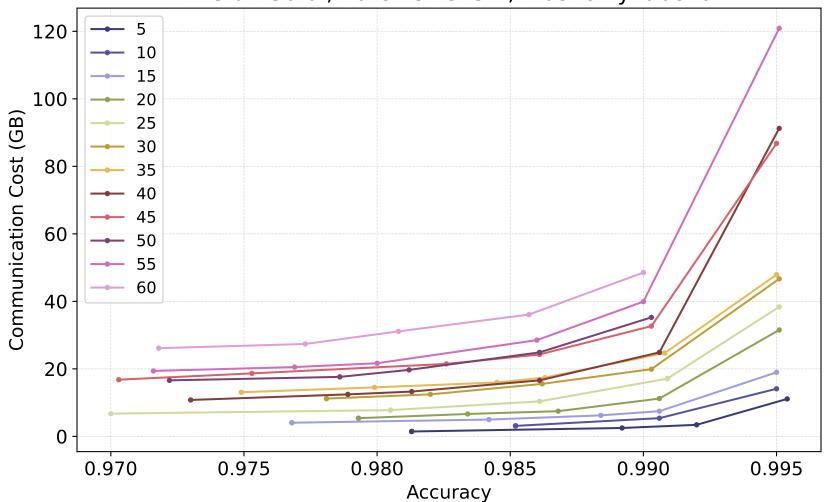
naive



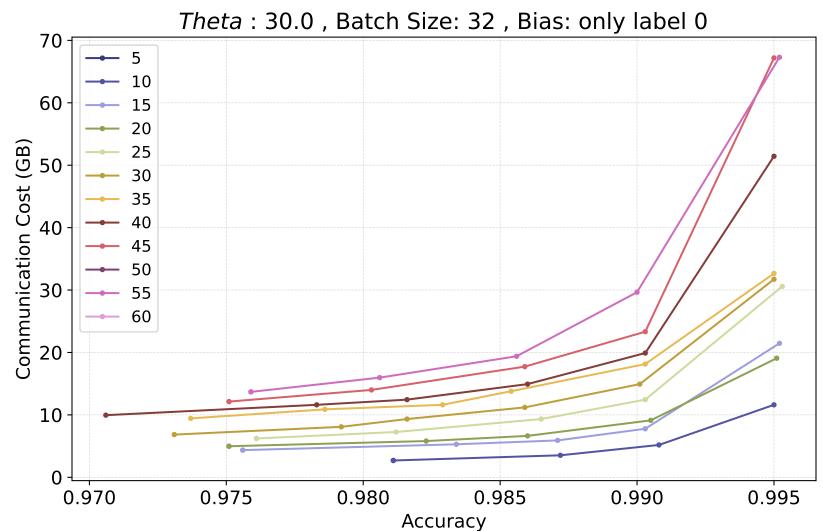
linear
Theta: 20.0, Batch Size: 32, Bias: only label 0



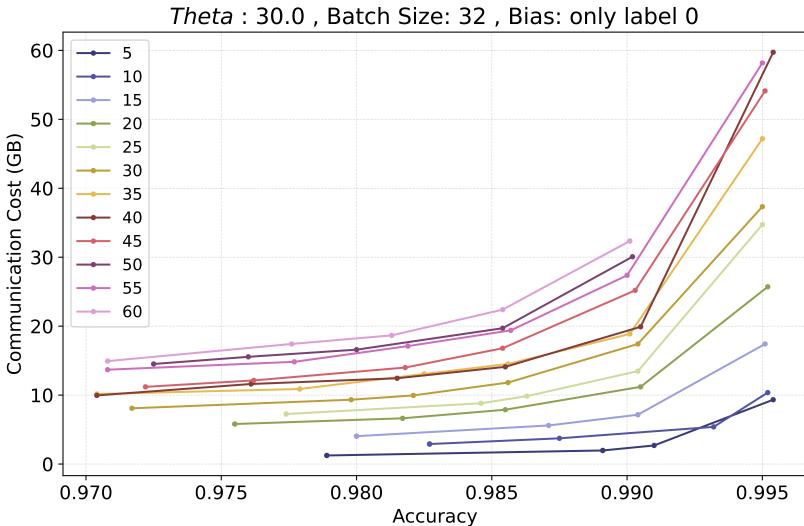
gm



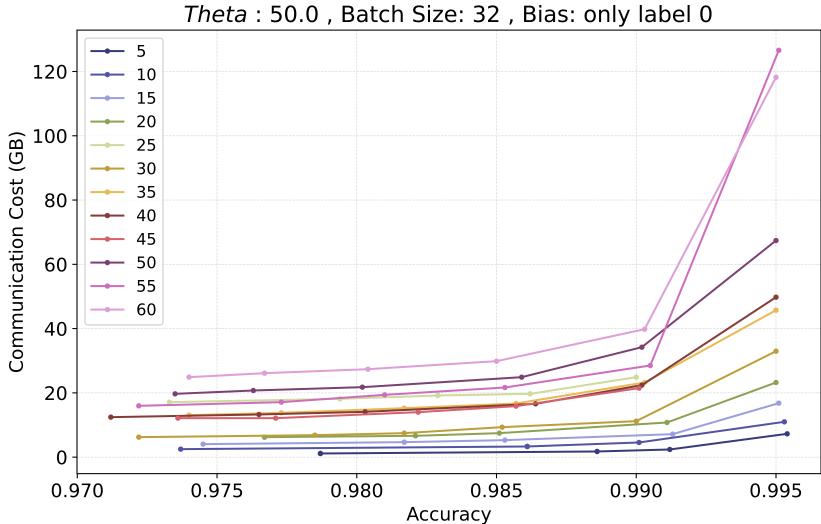
naive



linear

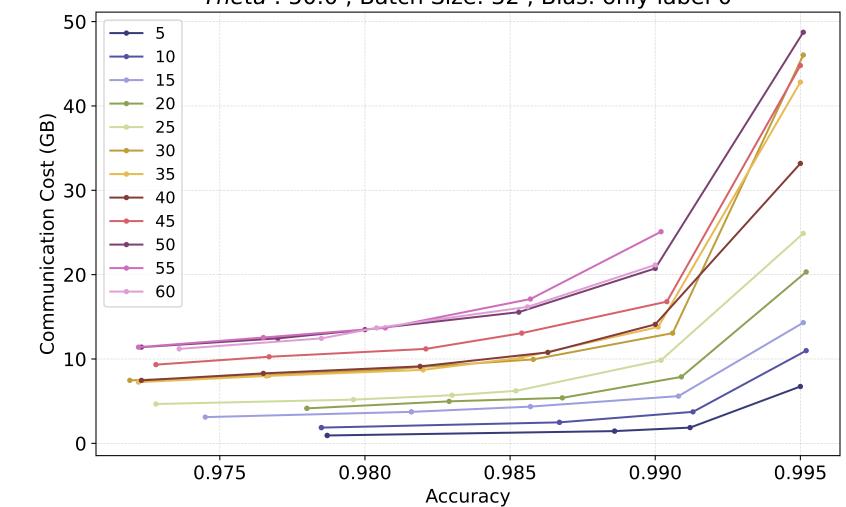


gm



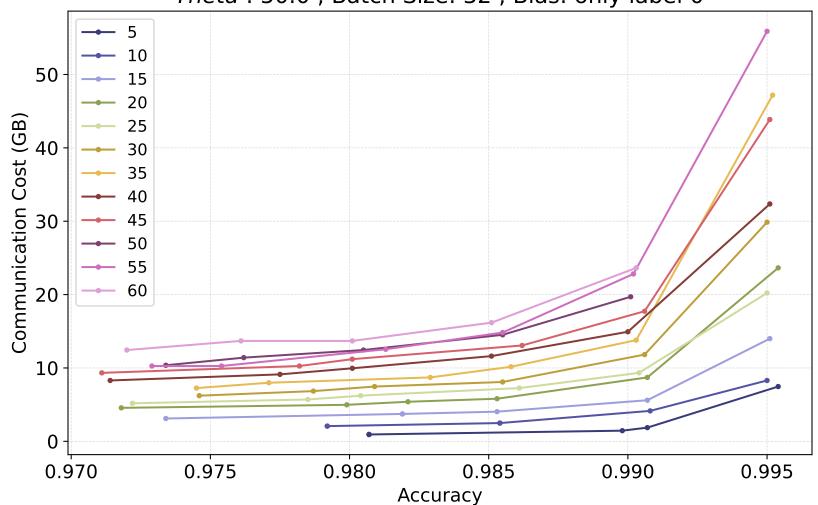
naive



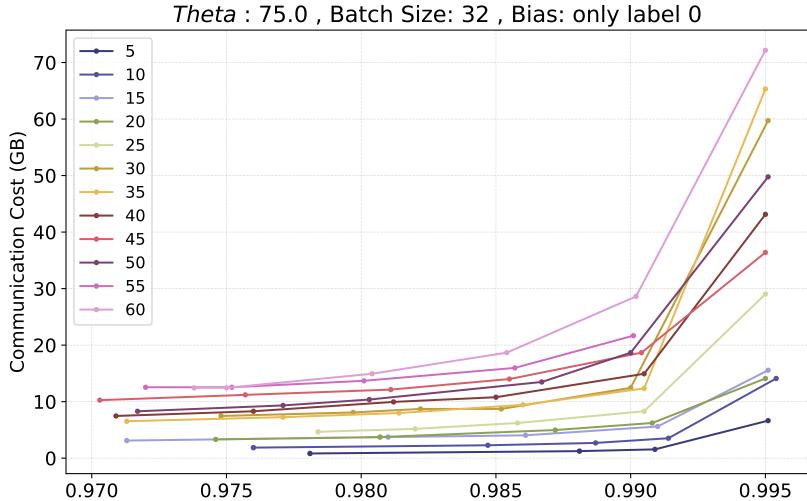


linear

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



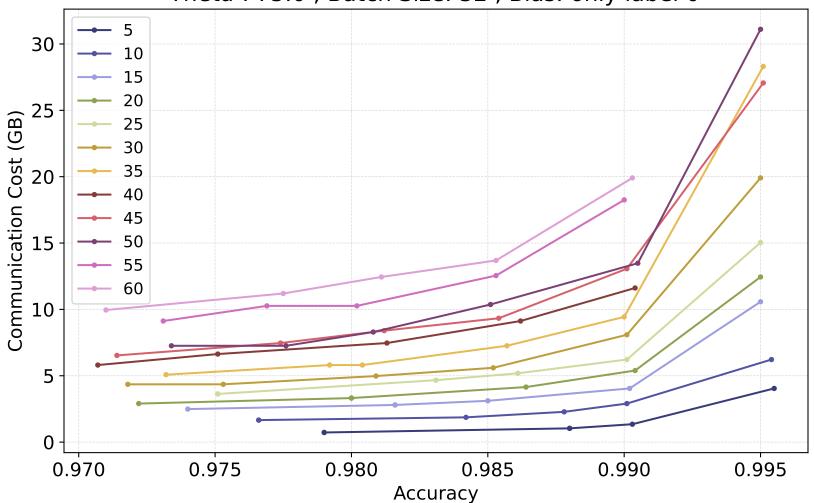
gm



Accuracy

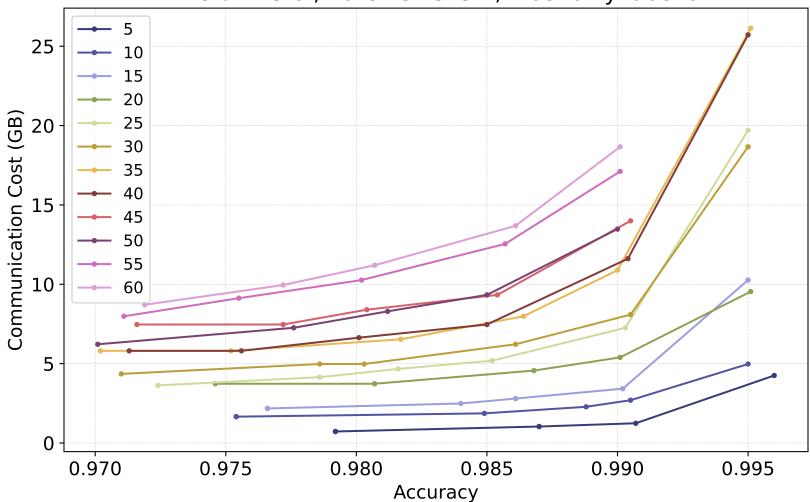
naive

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

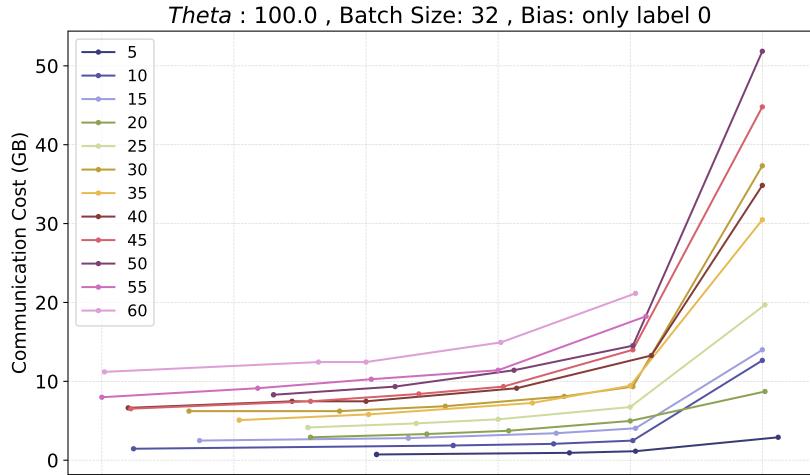


linear

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



gm



0.980

0.985

Accuracy

0.990

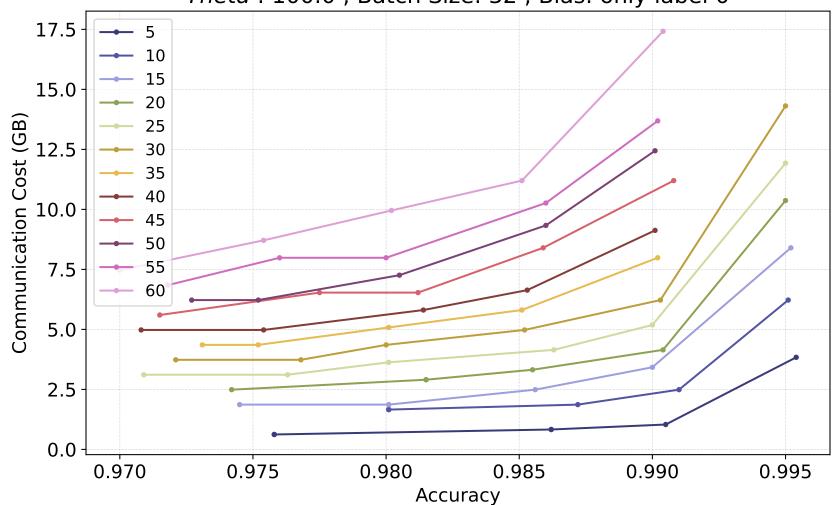
0.995

0.975

0.970

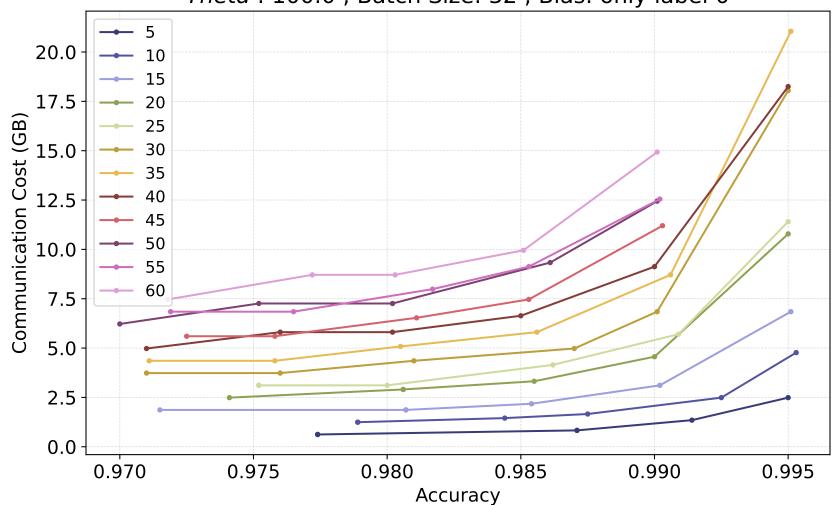
naive

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



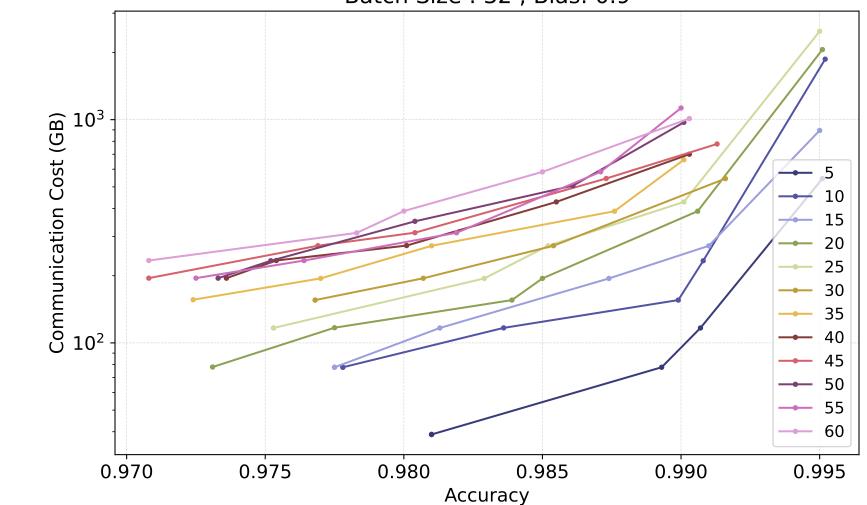
linear

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

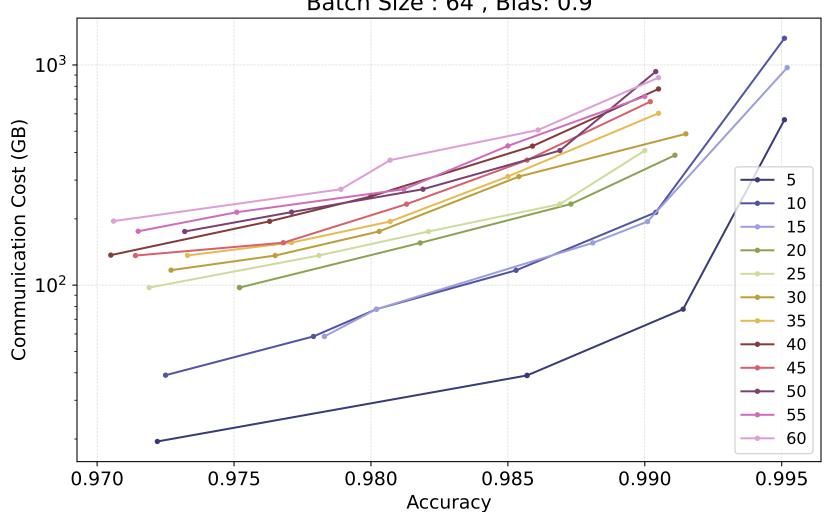


synchronous

Batch Size: 32, Bias: 0.9



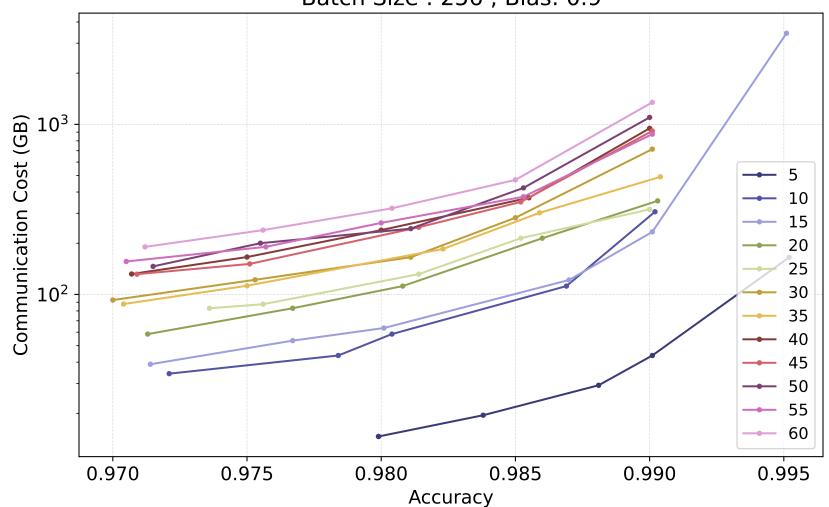
synchronous Batch Size: 64, Bias: 0.9



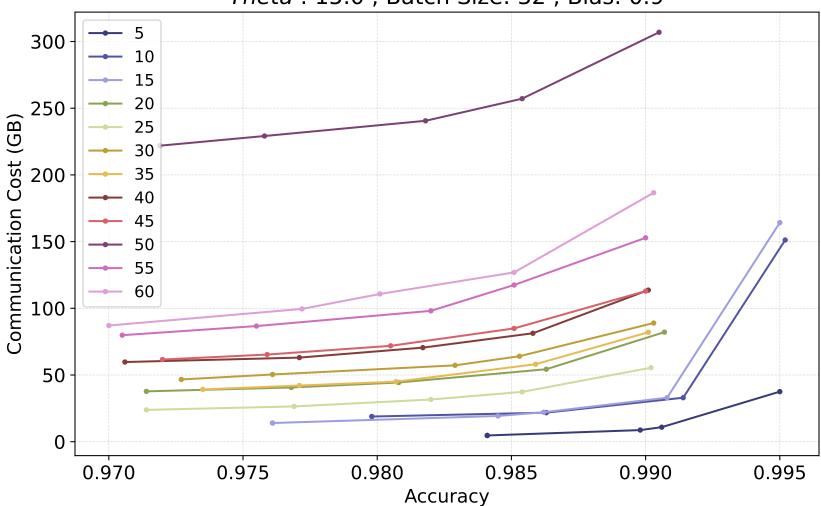
synchronous Batch Size: 128, Bias: 0.9 10³ 10 15 20 Communication Cost (GB) 25 30 35 40 45 50 10² 55 60 0.980 0.995 0.970 0.975 0.985 0.990 Accuracy

synchronous

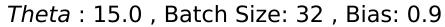
Batch Size: 256, Bias: 0.9

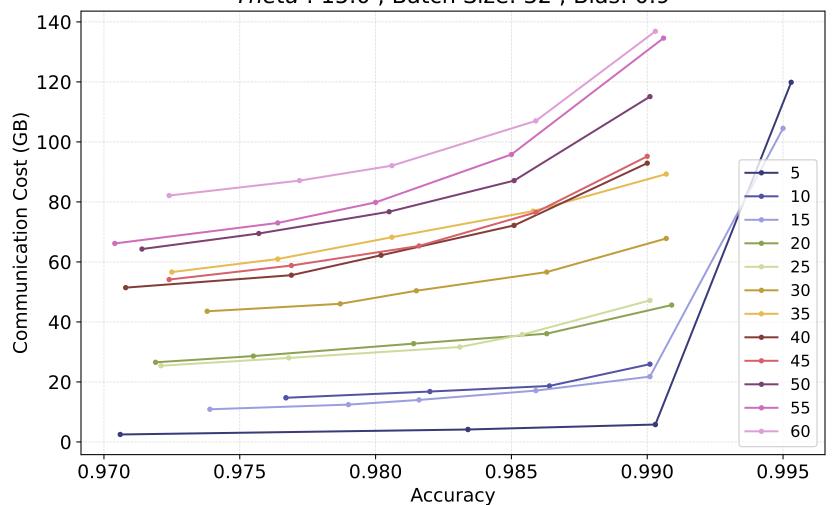


gm
Theta: 15.0, Batch Size: 32, Bias: 0.9



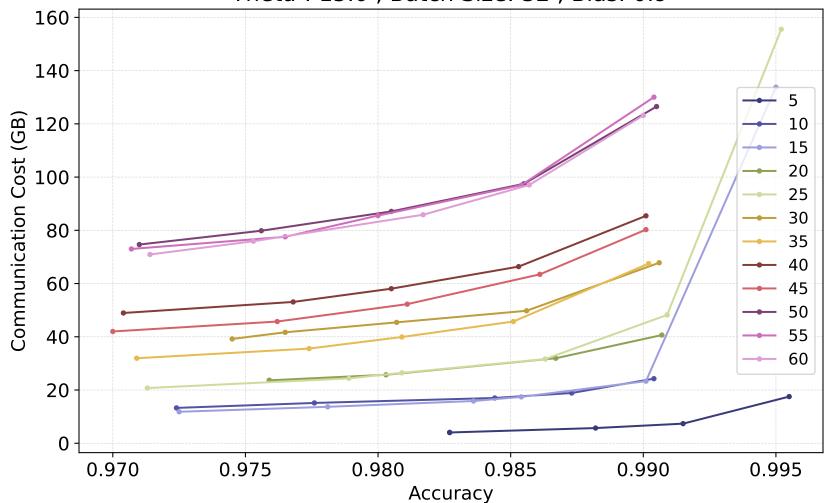
naive





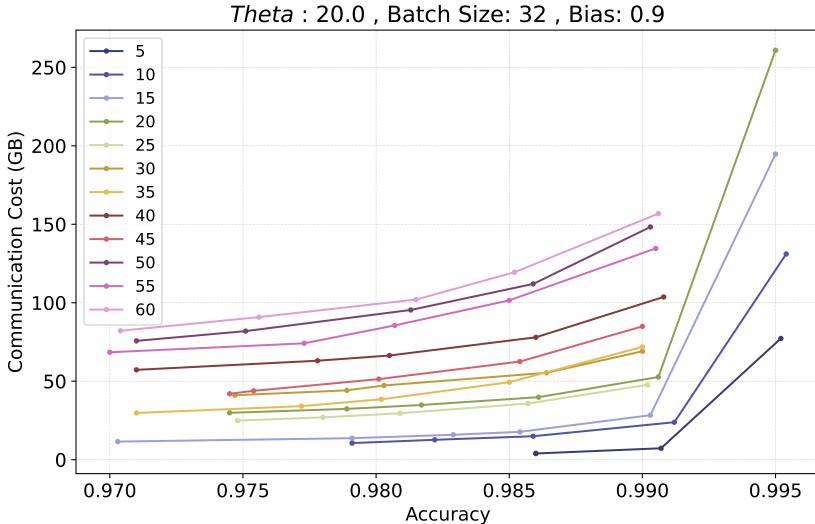
linear

Theta: 15.0, Batch Size: 32, Bias: 0.9



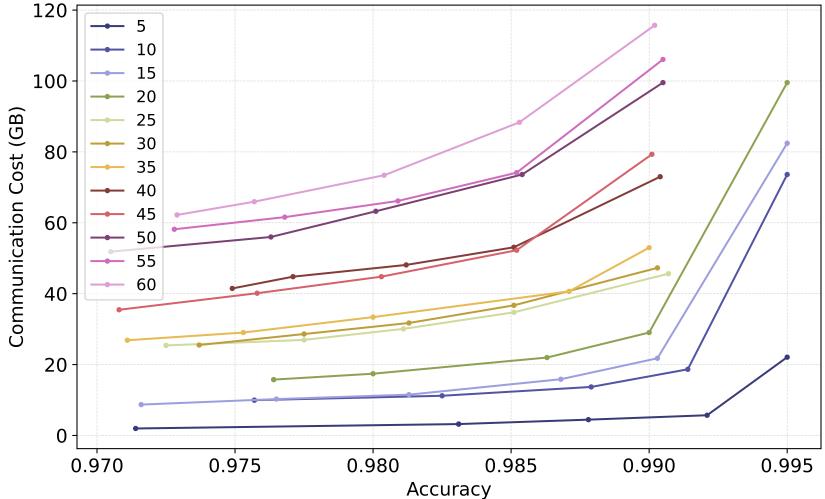
sketch Theta: 15.0, Batch Size: 32, Bias: 0.9 160 10 140 15 20 Communication Cost (GB) 100 80 60 40 25 30 35 40 45 50 55 60 20 0 0.970 0.975 0.980 0.985 0.990 0.995 Accuracy

gm

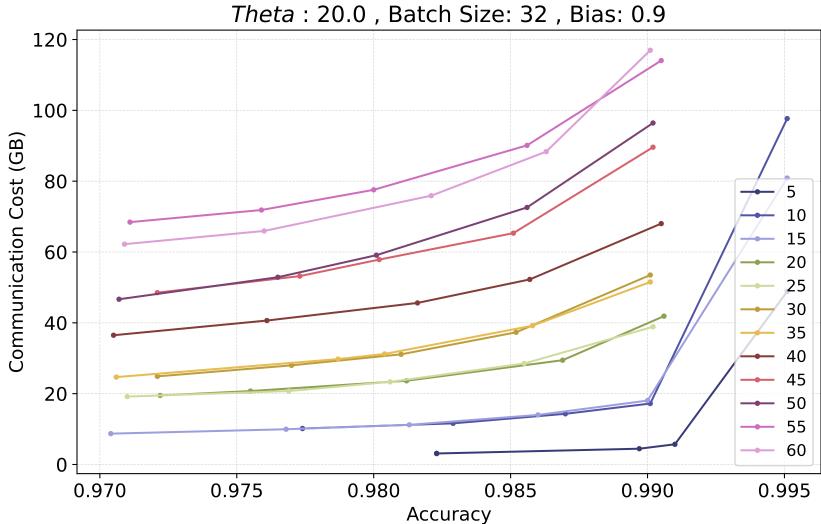


naive



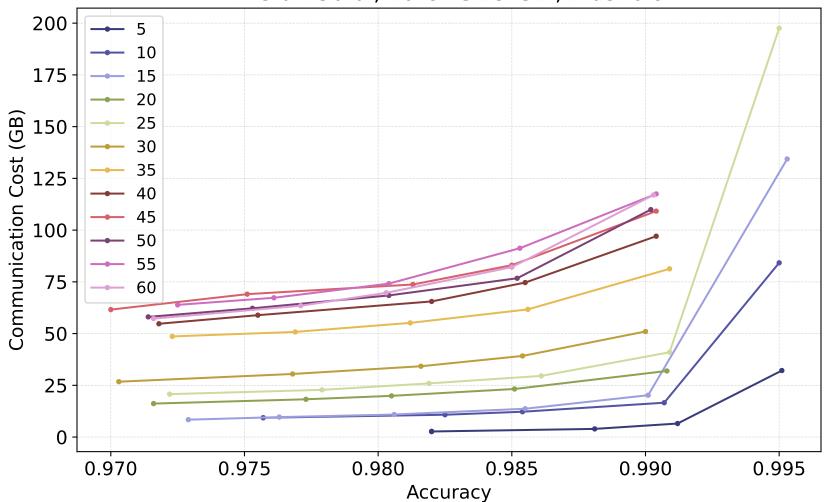


linear

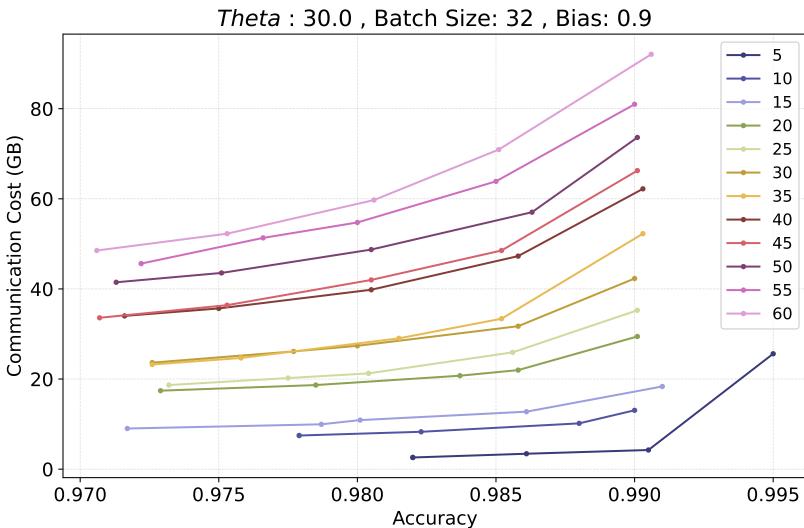


gm

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

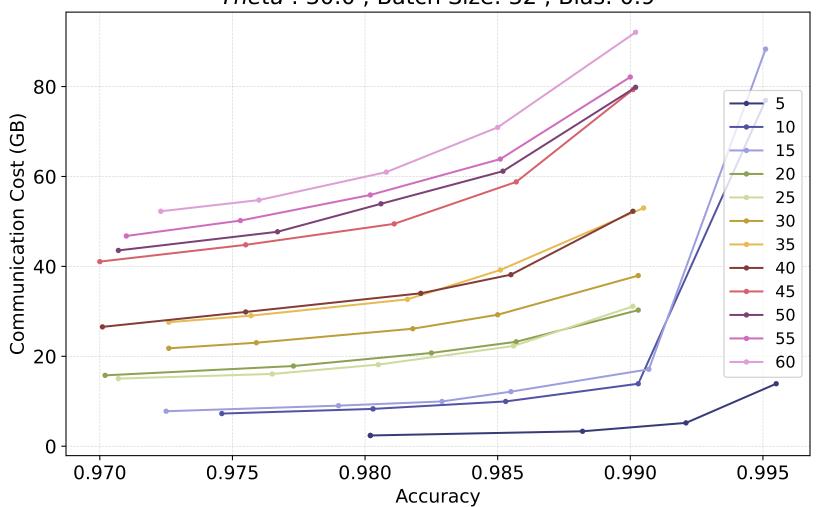


naive

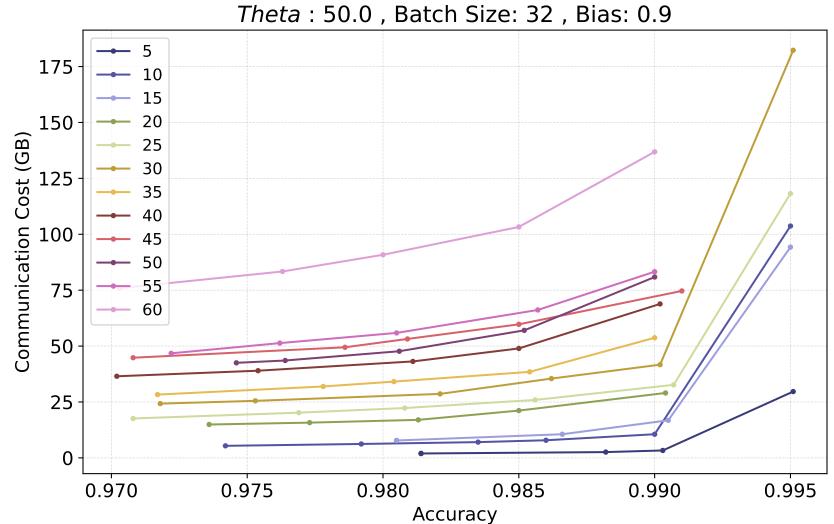


linear





gm



naive Theta: 50.0, Batch Size: 32, Bias: 0.9 Communication Cost (GB)

0.980

0.985

Accuracy

0.990

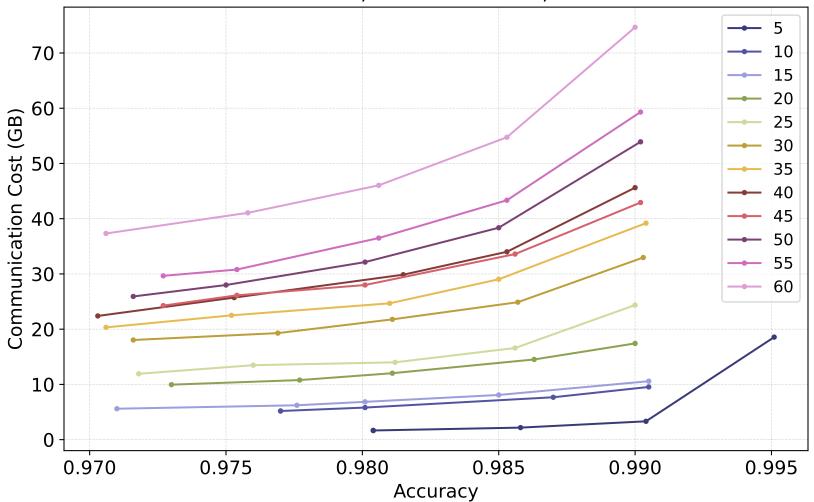
0.995

0.970

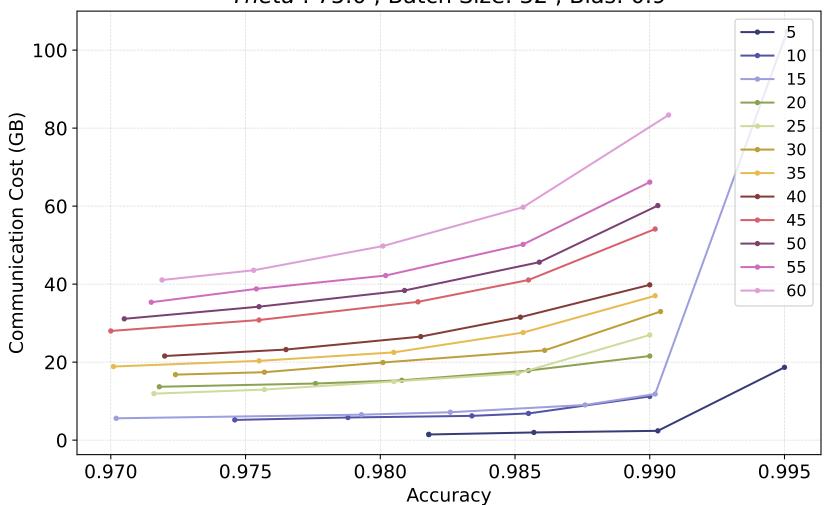
0.975

linear

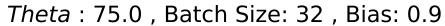


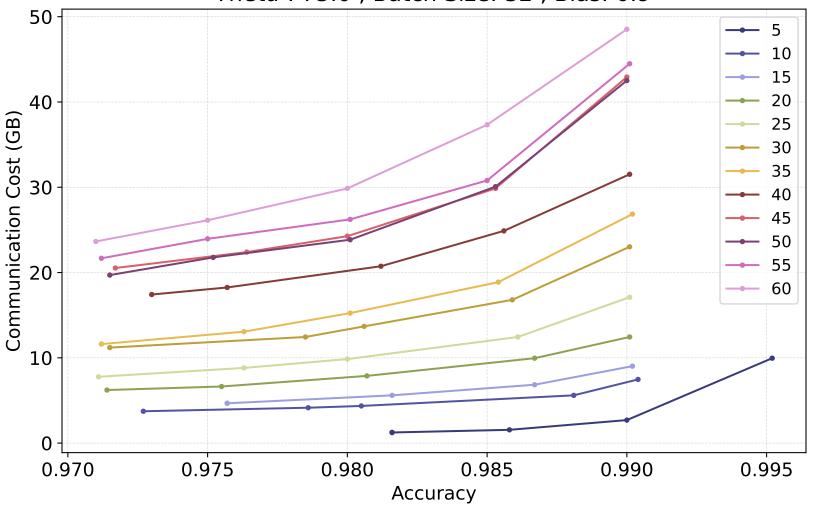


gm *Theta*: 75.0, Batch Size: 32, Bias: 0.9



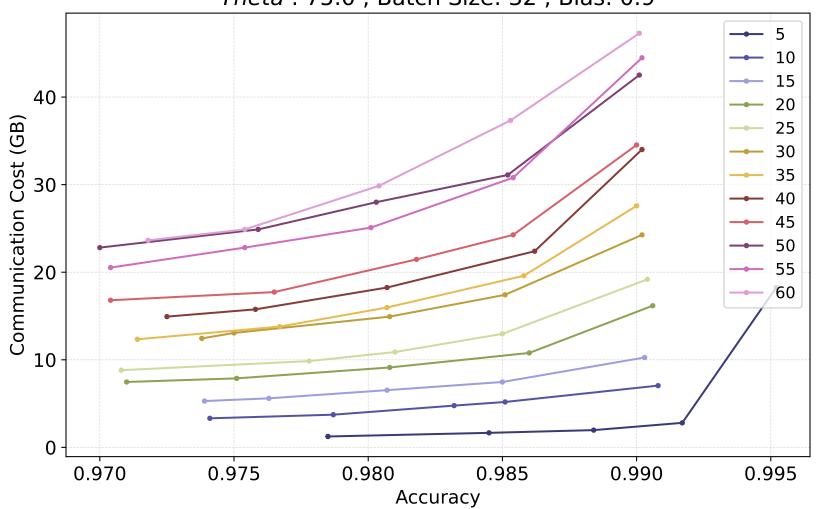
naive



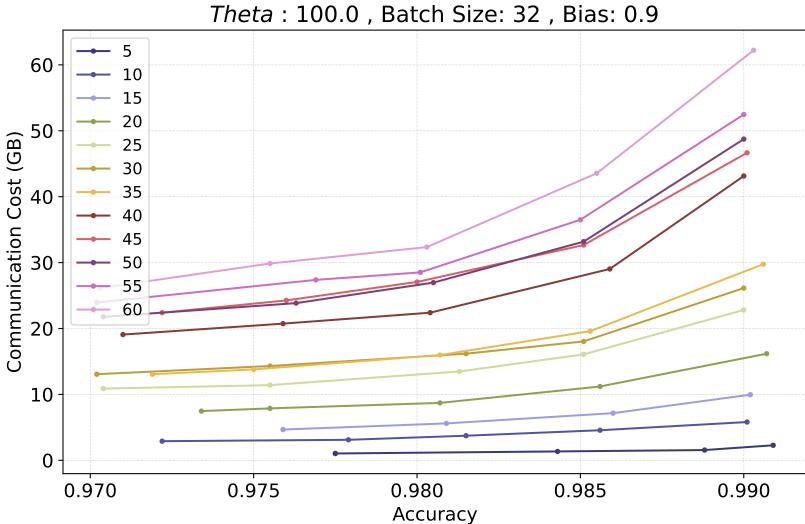


linear



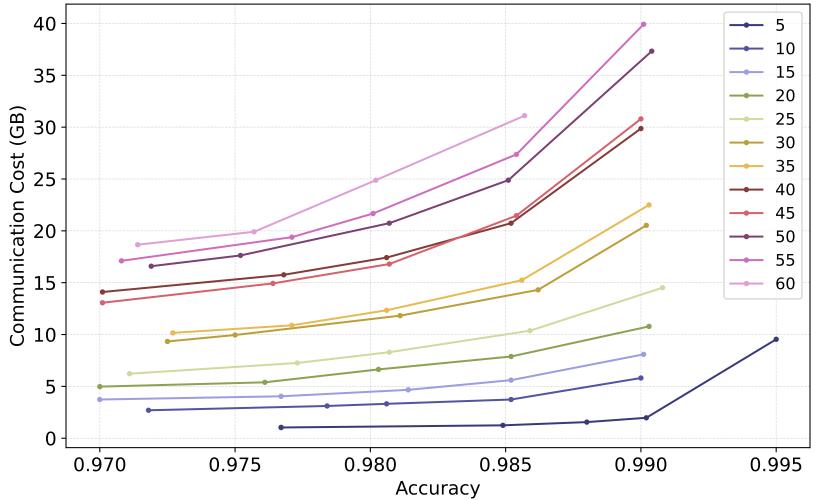


gm



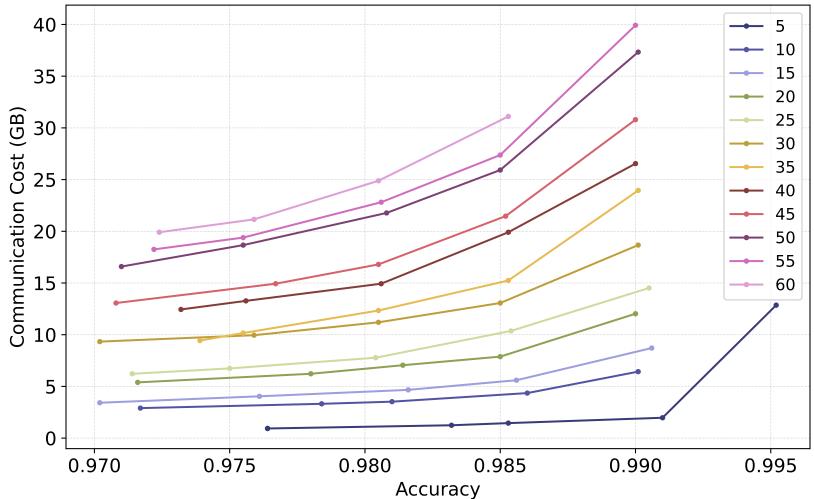
naive





linear

Theta: 100.0, Batch Size: 32, Bias: 0.9



sketch Theta: 100.0, Batch Size: 32, Bias: 0.9 Communication Cost (GB)

0.985

Accuracy

0.990

0.995

0.980

0.970

0.975