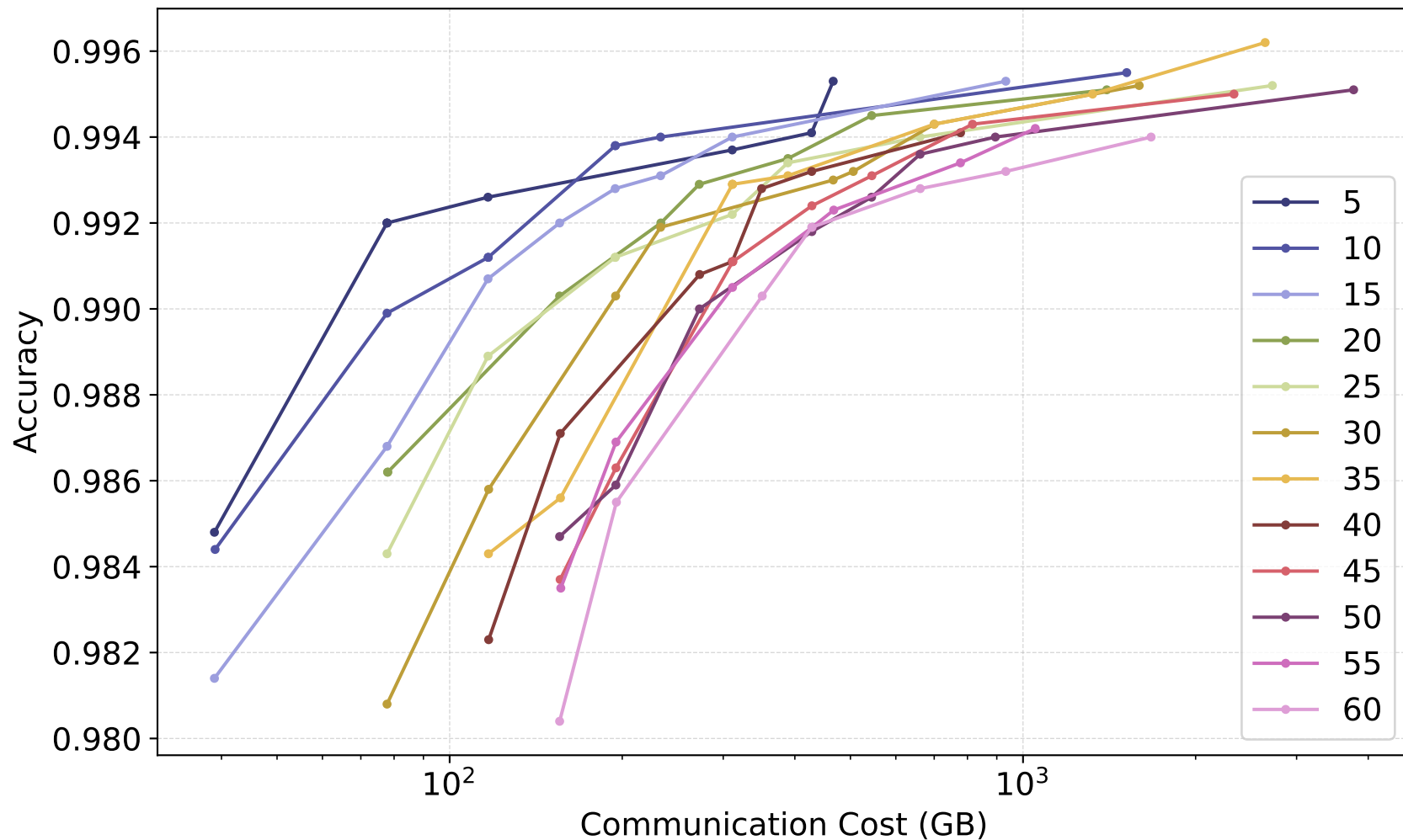


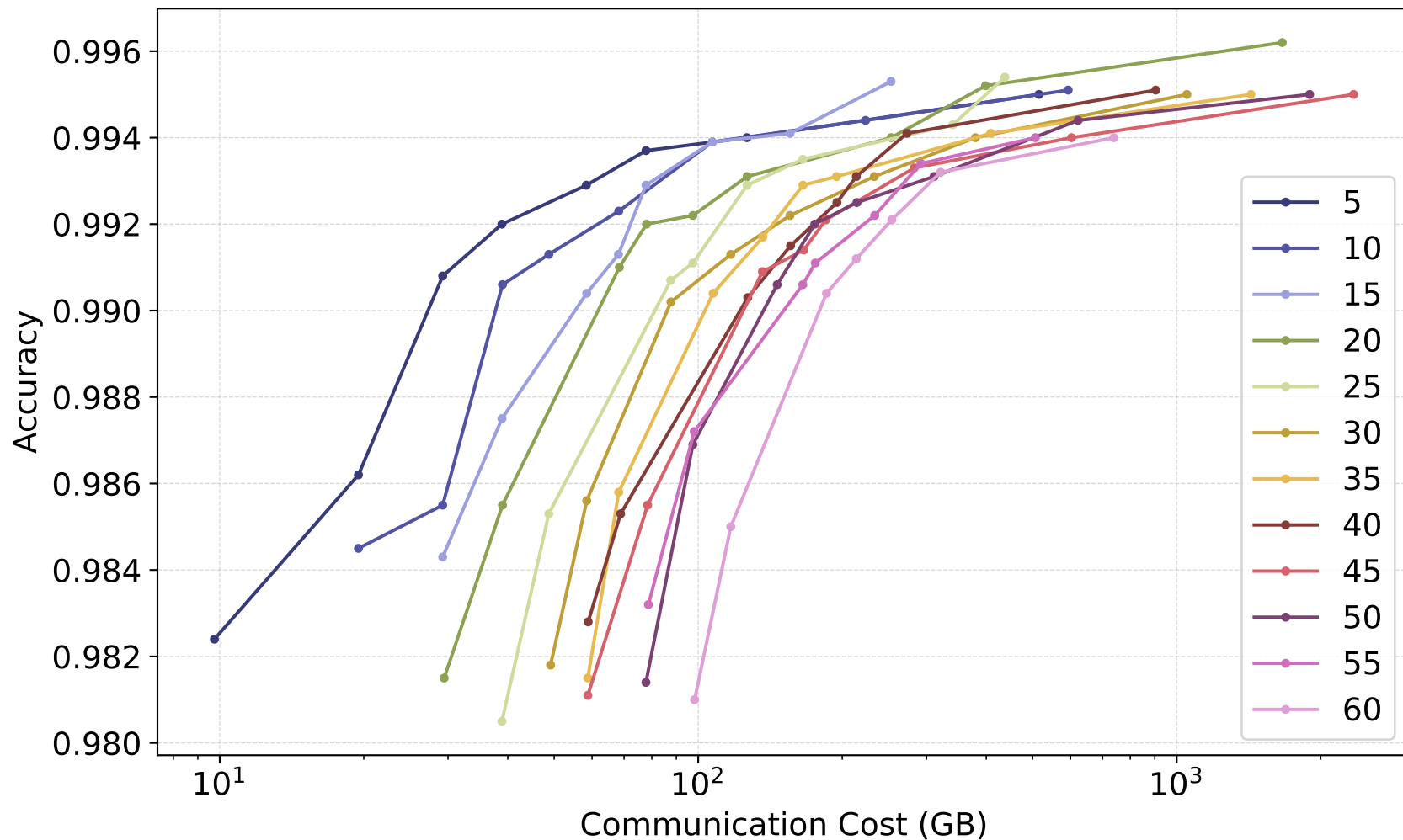
synchronous  
Batch Size : 32 , Bias: nan



Batch Size : 64 , Bias: nan



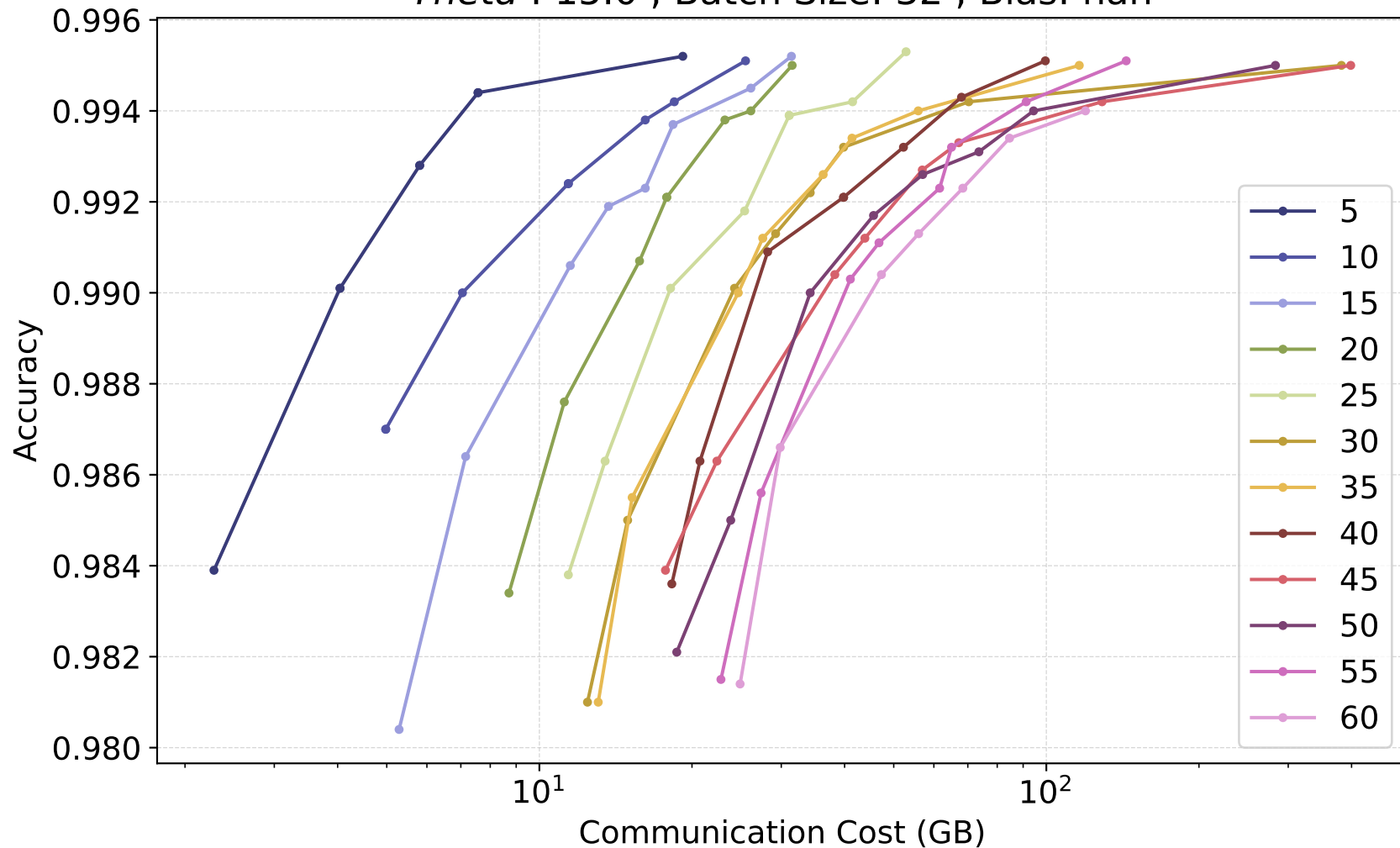
synchronous  
Batch Size : 128 , Bias: nan



Batch Size : 256 , Bias: nan

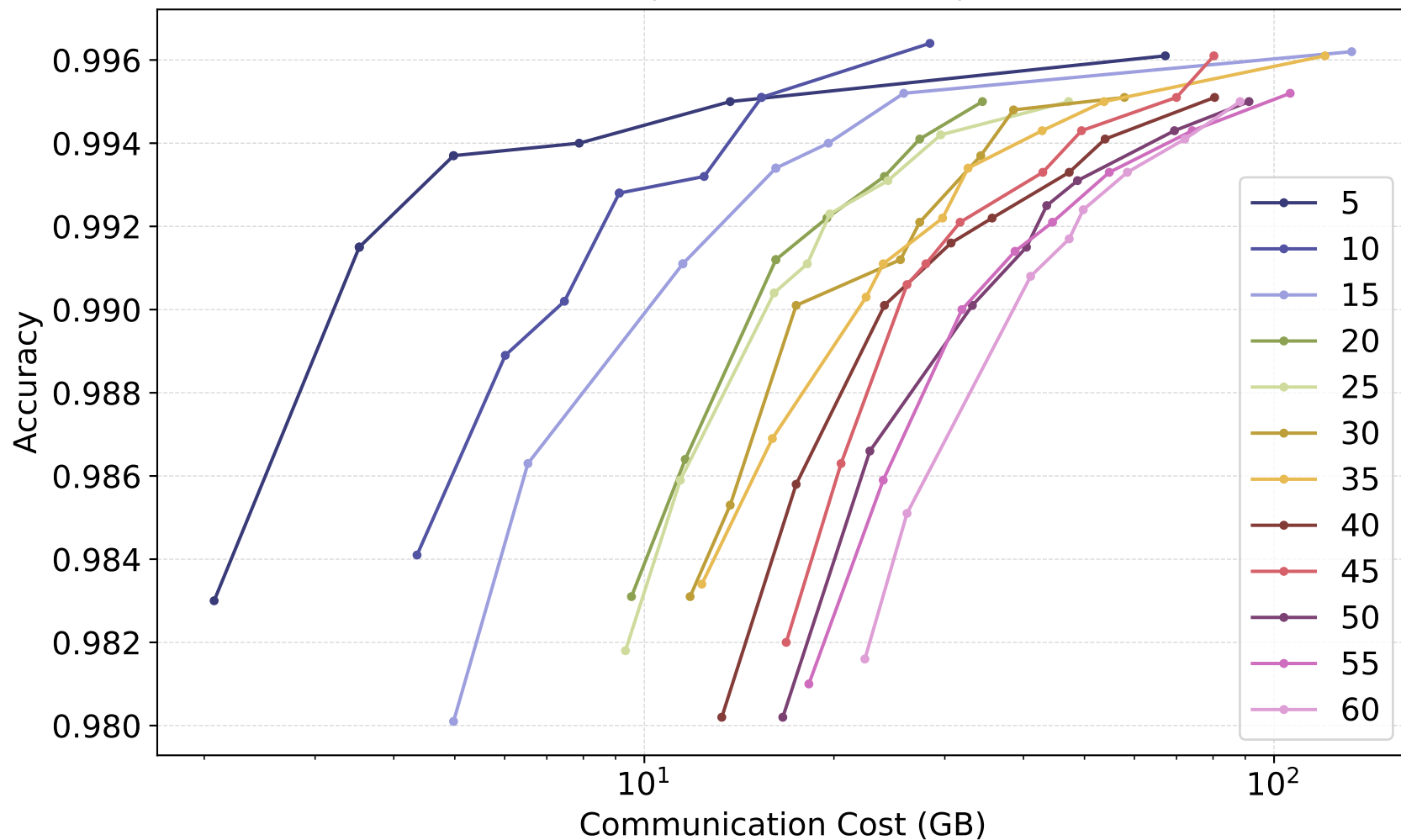


gm  
*Theta* : 15.0 , Batch Size: 32 , Bias: nan



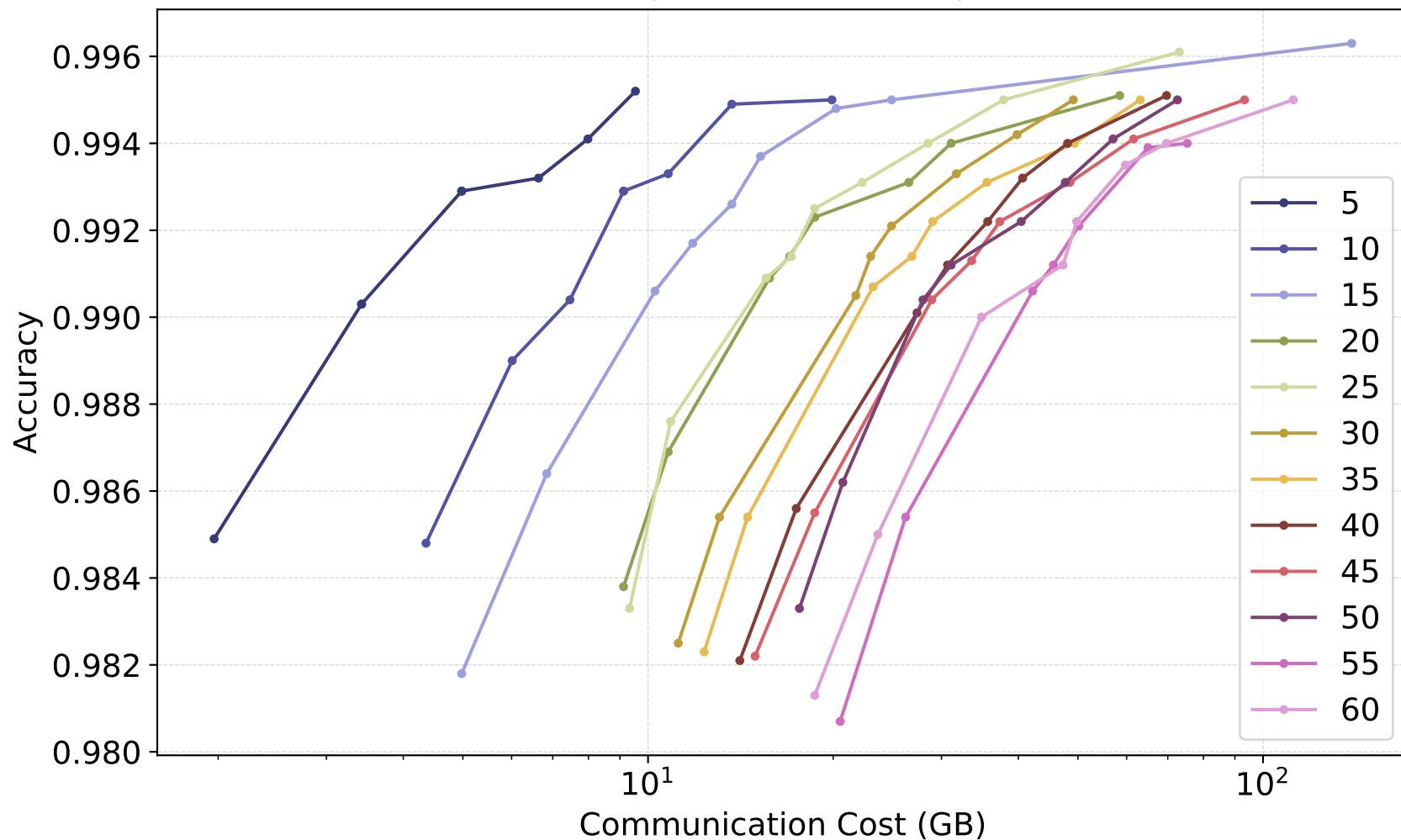
naive

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

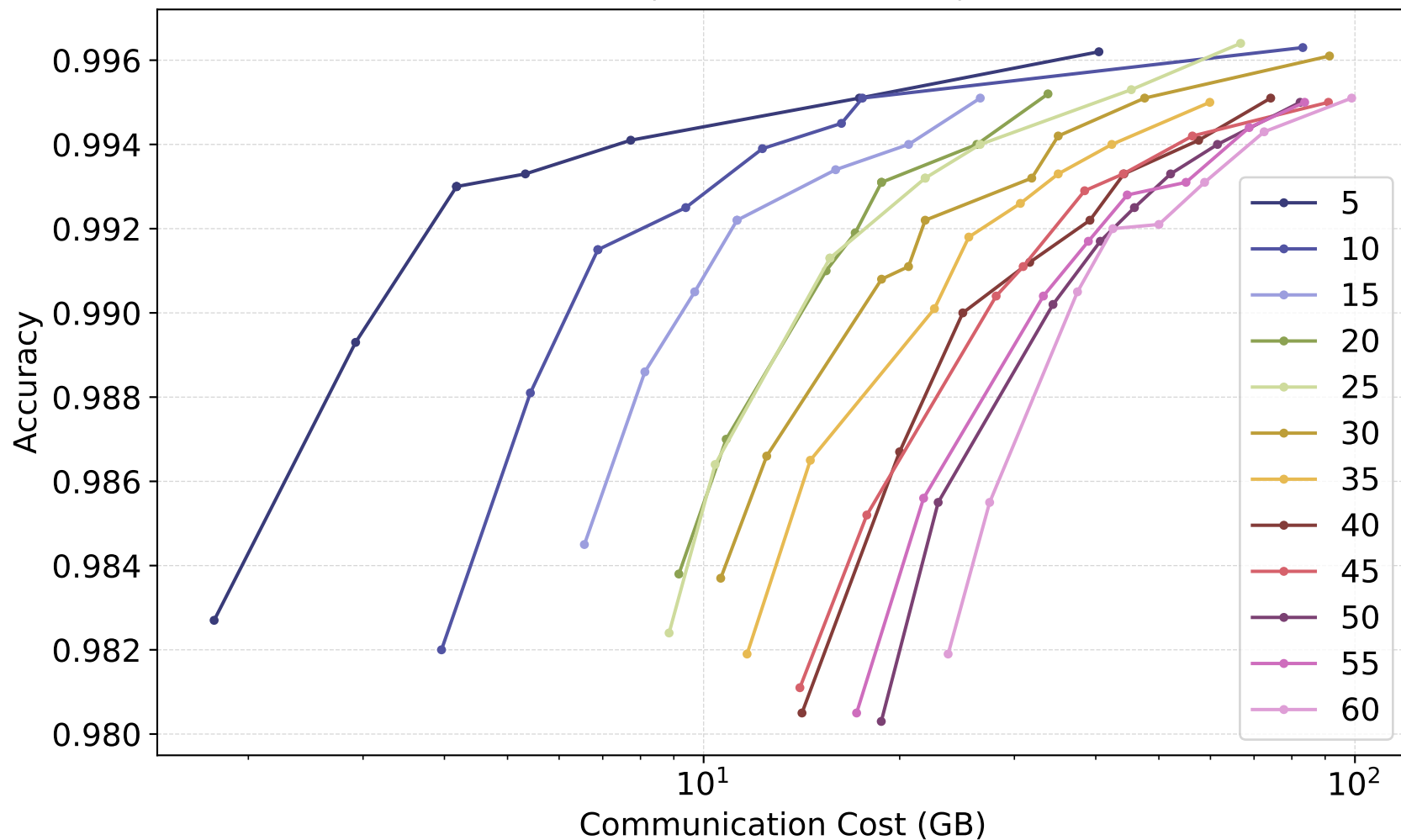


linear

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

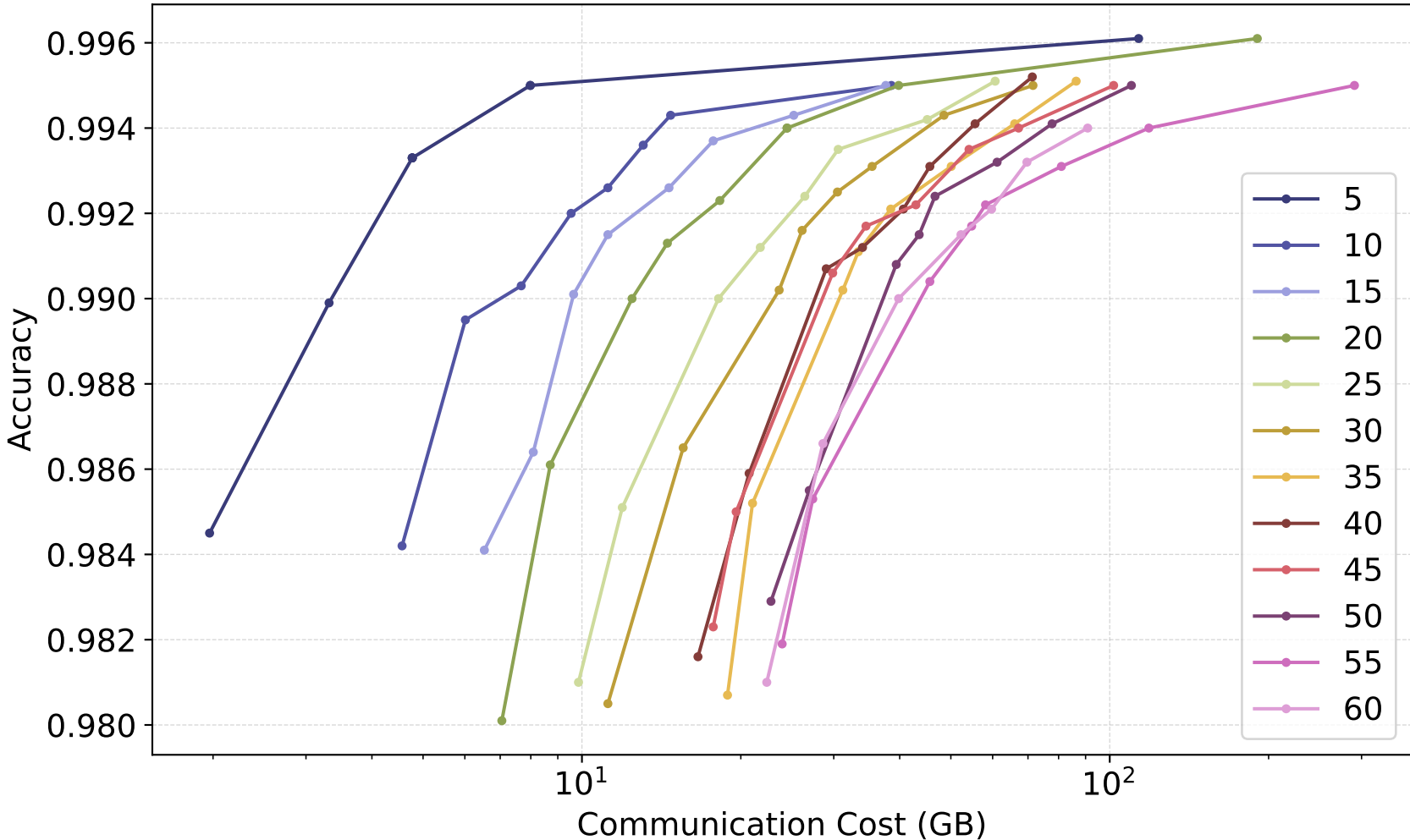


*Theta* : 15.0 , Batch Size: 32 , Bias: nan



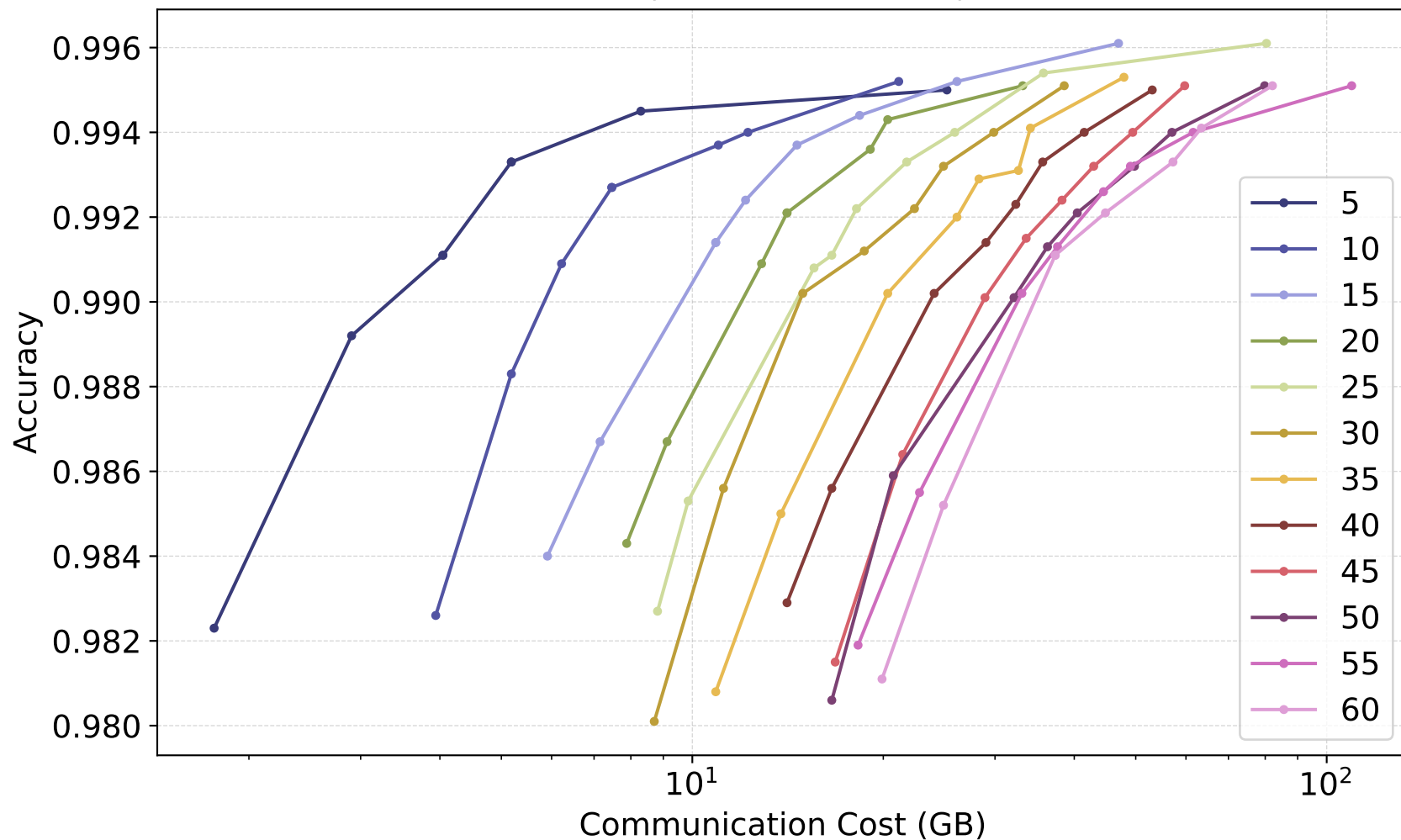


gm  
Theta : 20.0 , Batch Size: 32 , Bias: nan

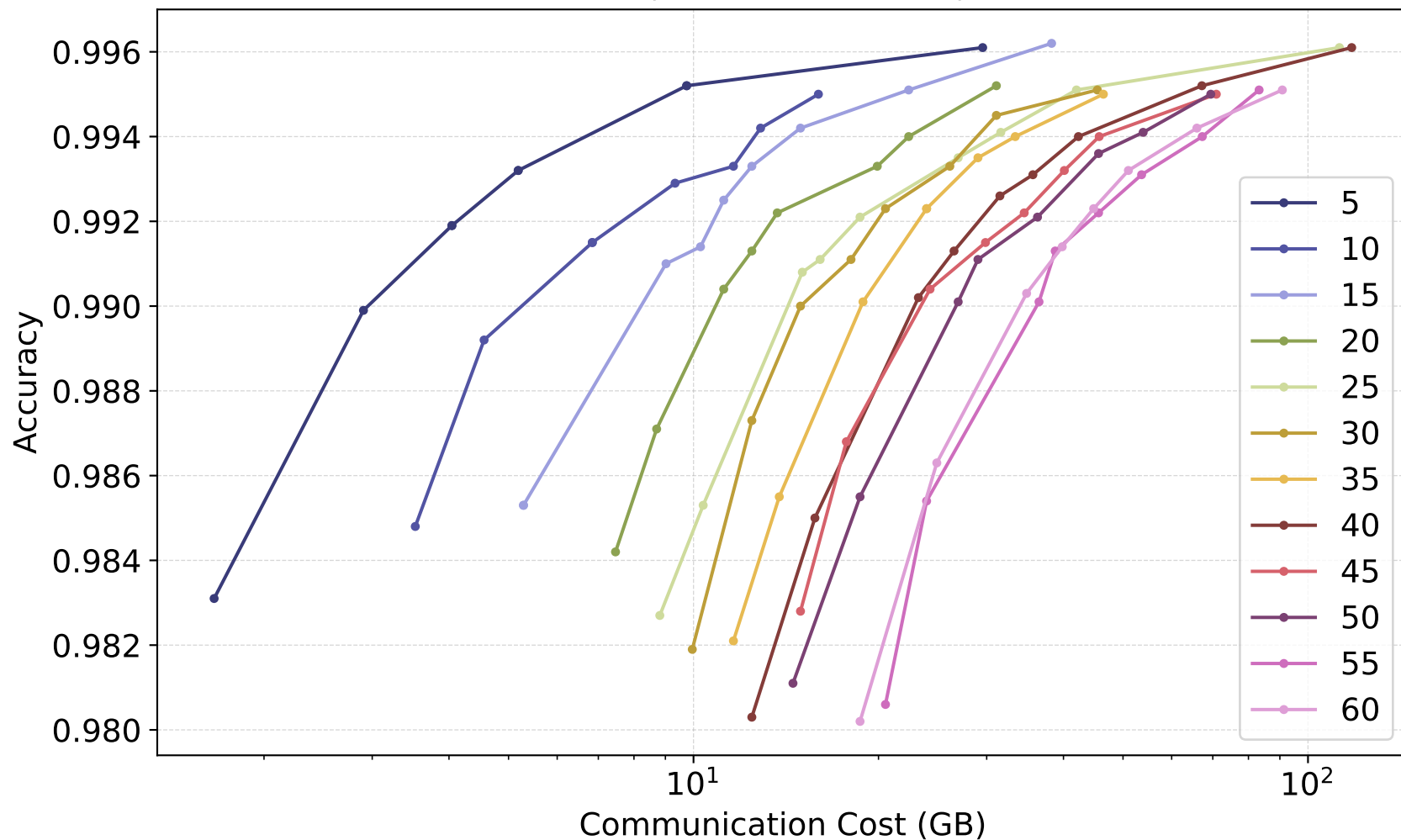


naive

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

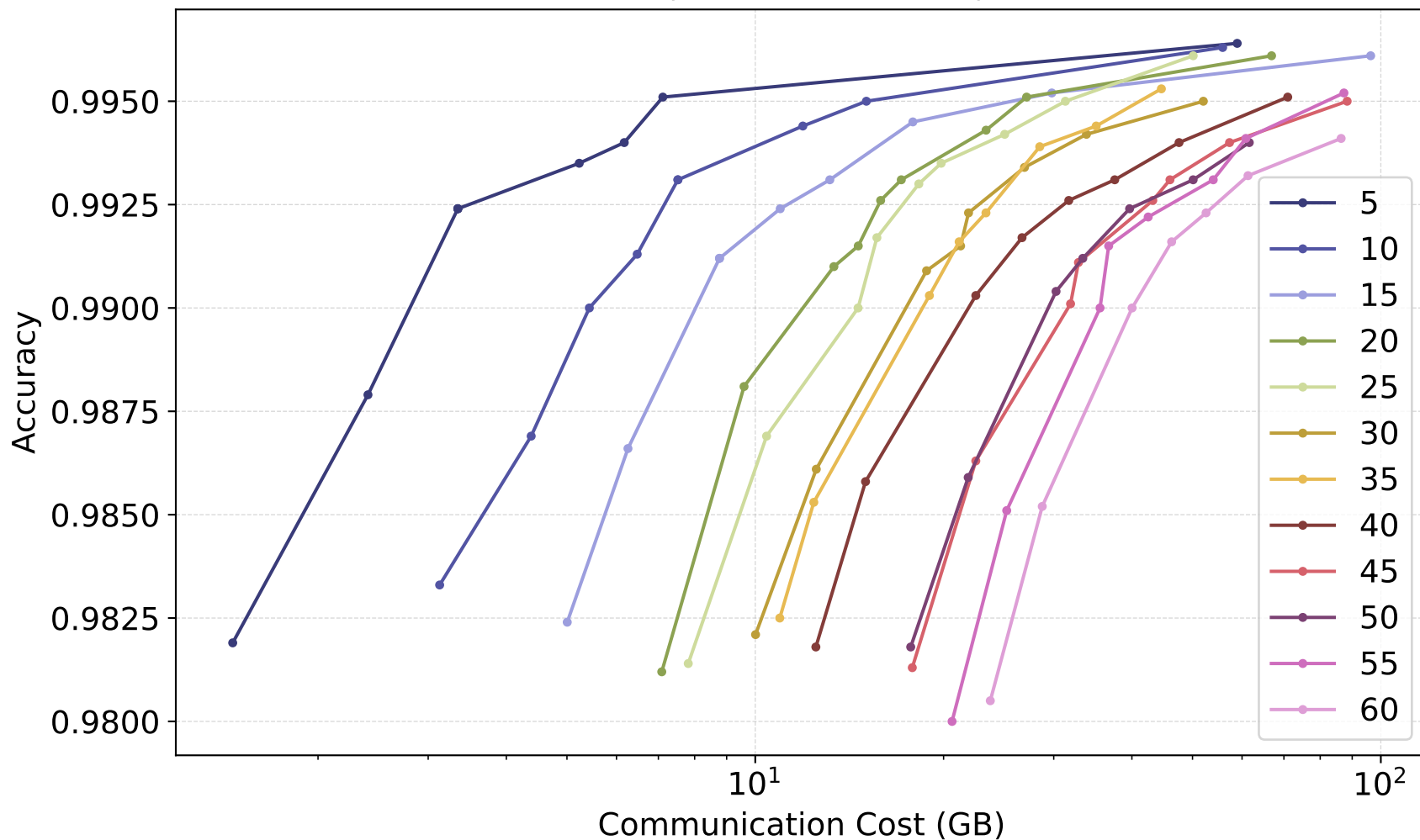


*Theta* : 20.0 , Batch Size: 32 , Bias: nan

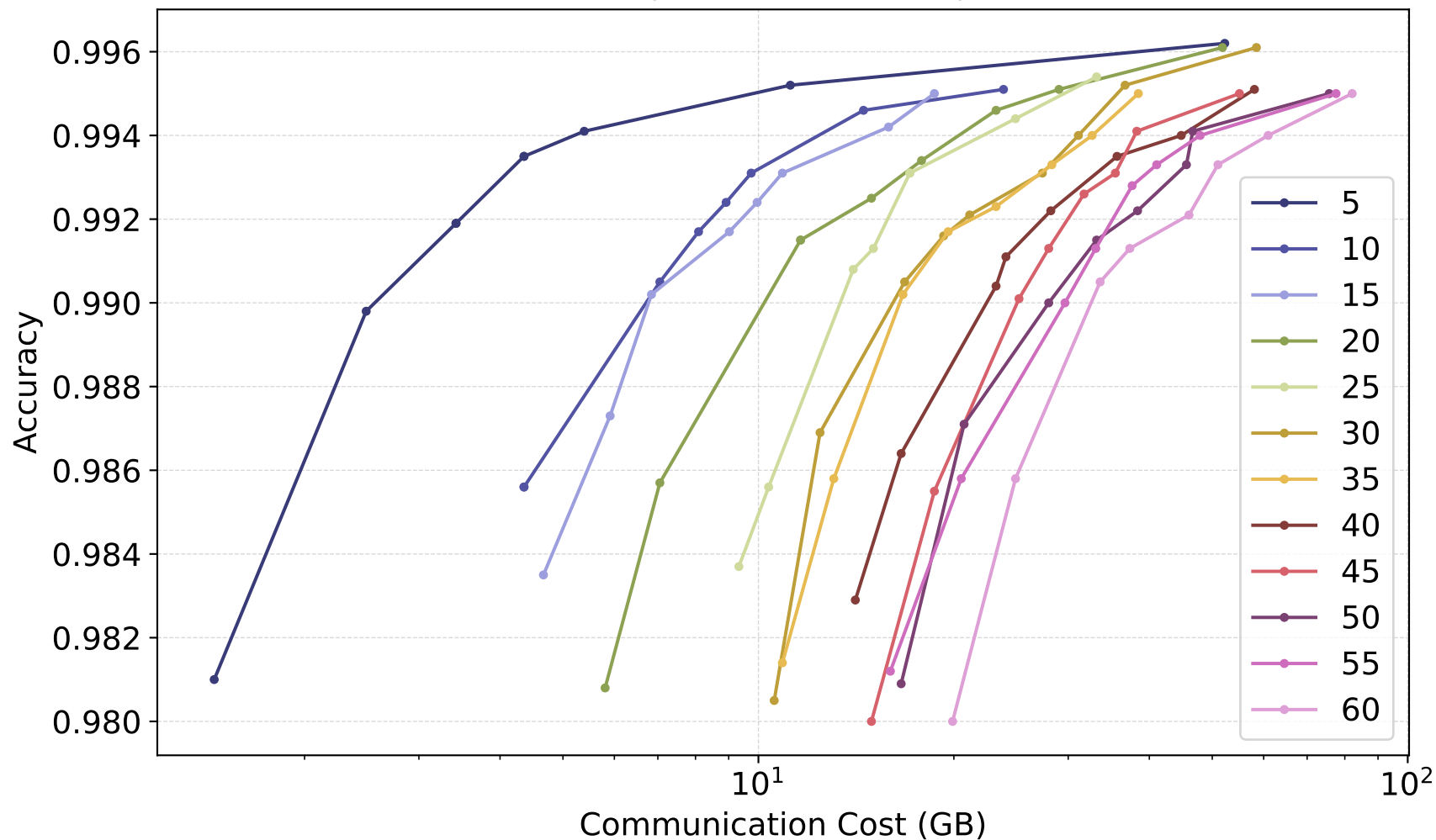


sketch

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

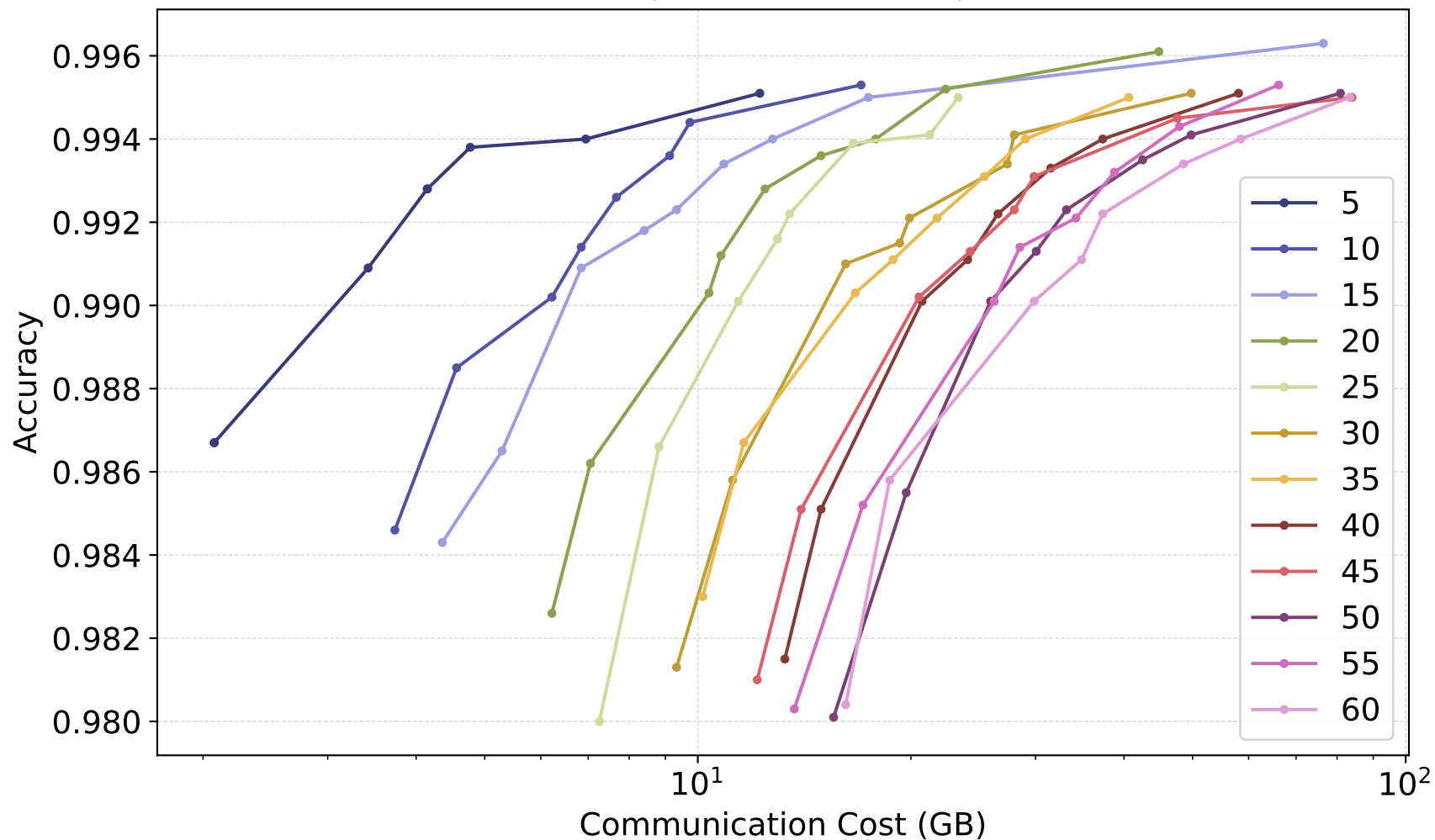


*Theta* : 25.0 , Batch Size: 32 , Bias: nan

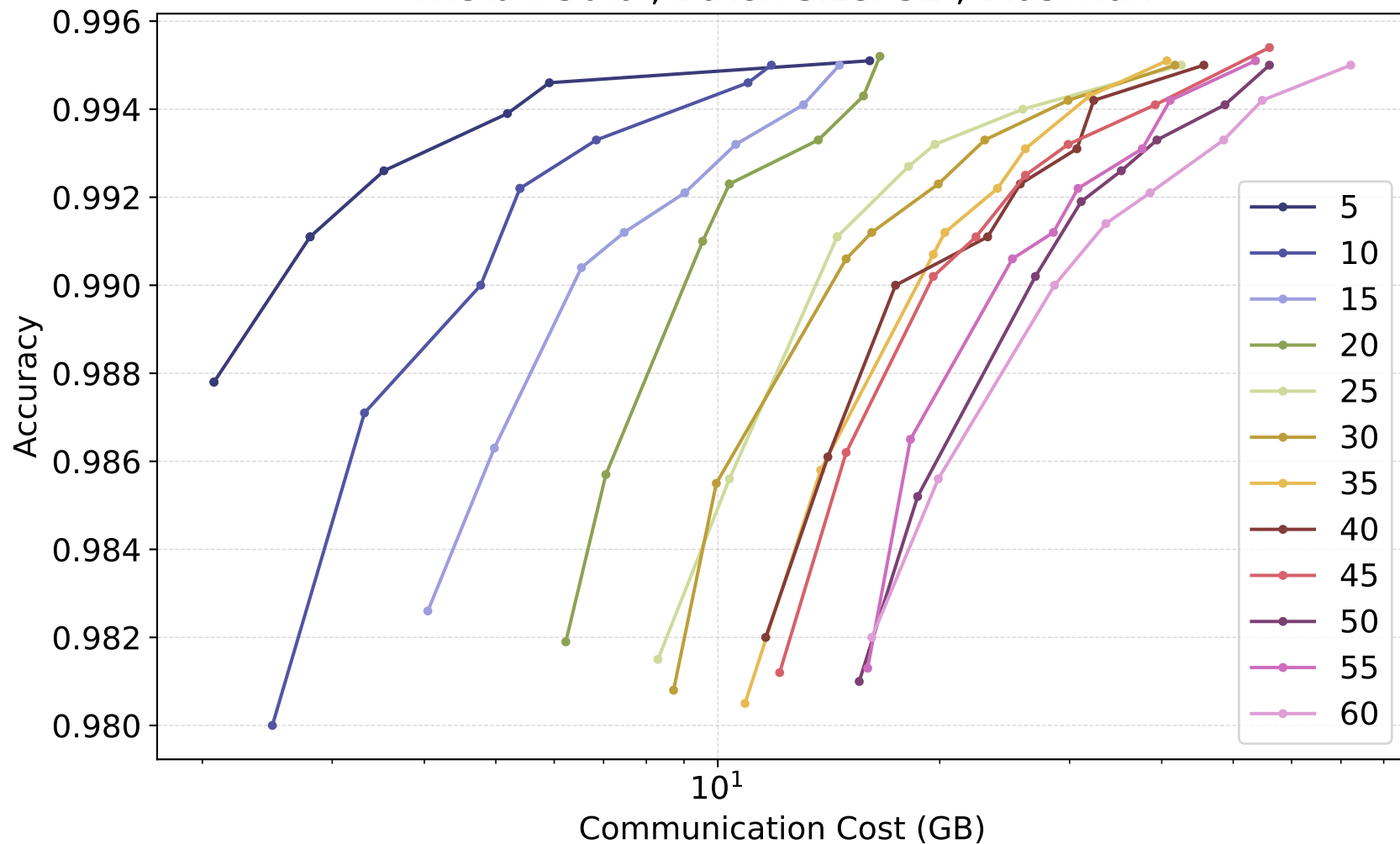




*Theta* : 30.0 , Batch Size: 32 , Bias: nan



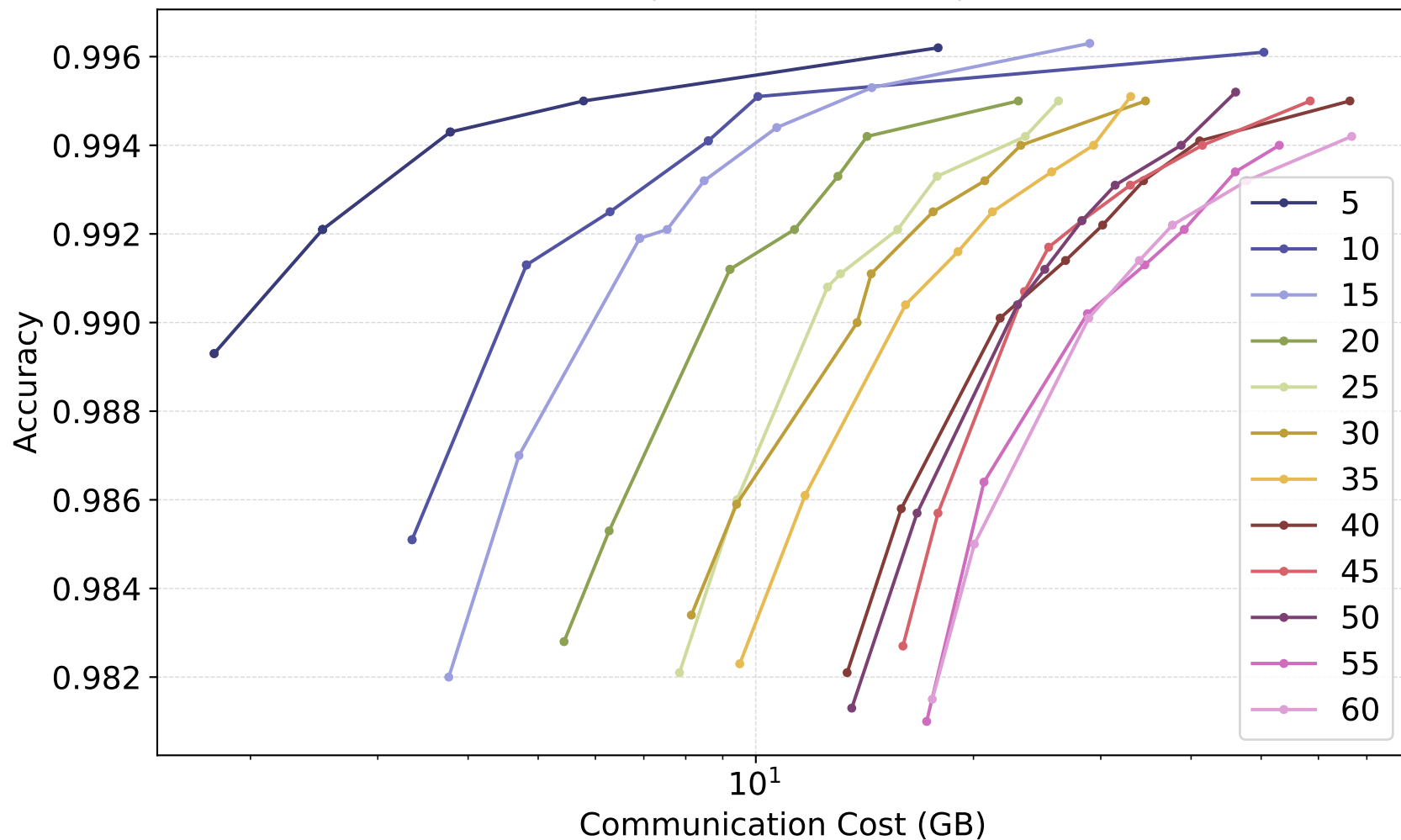
*Theta* : 30.0 , Batch Size: 32 , Bias: nan



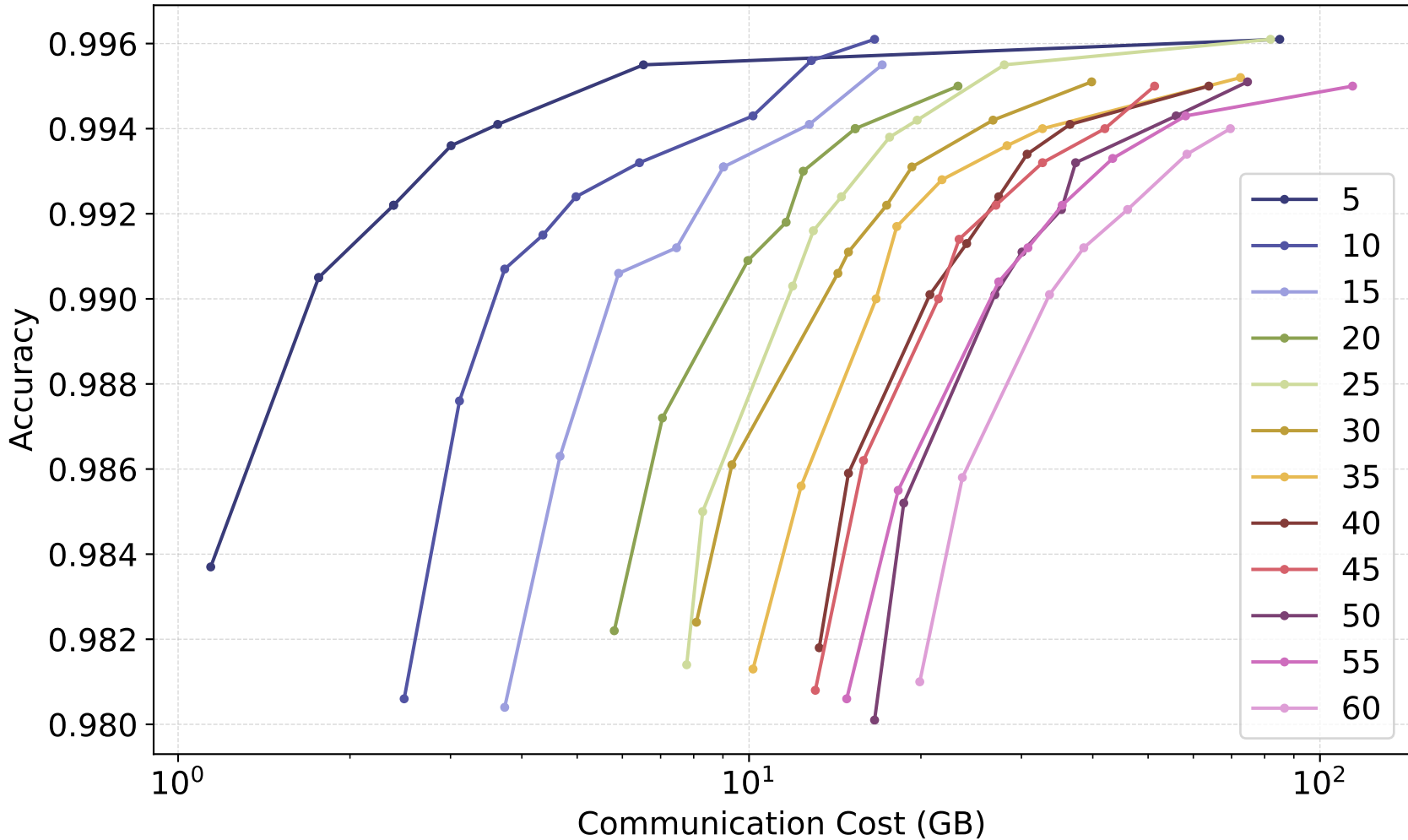


sketch

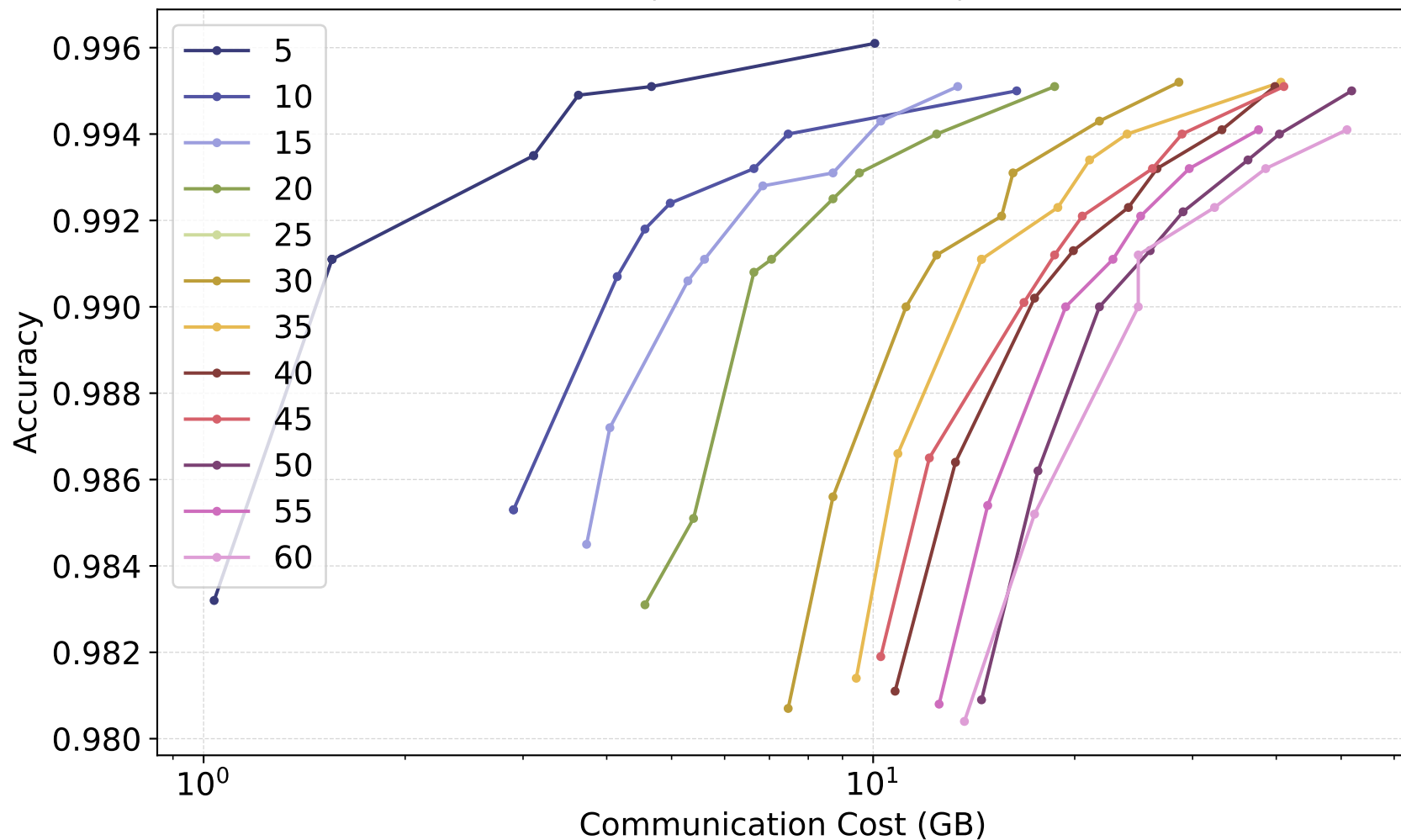
*Theta* : 30.0 , Batch Size: 32 , Bias: nan



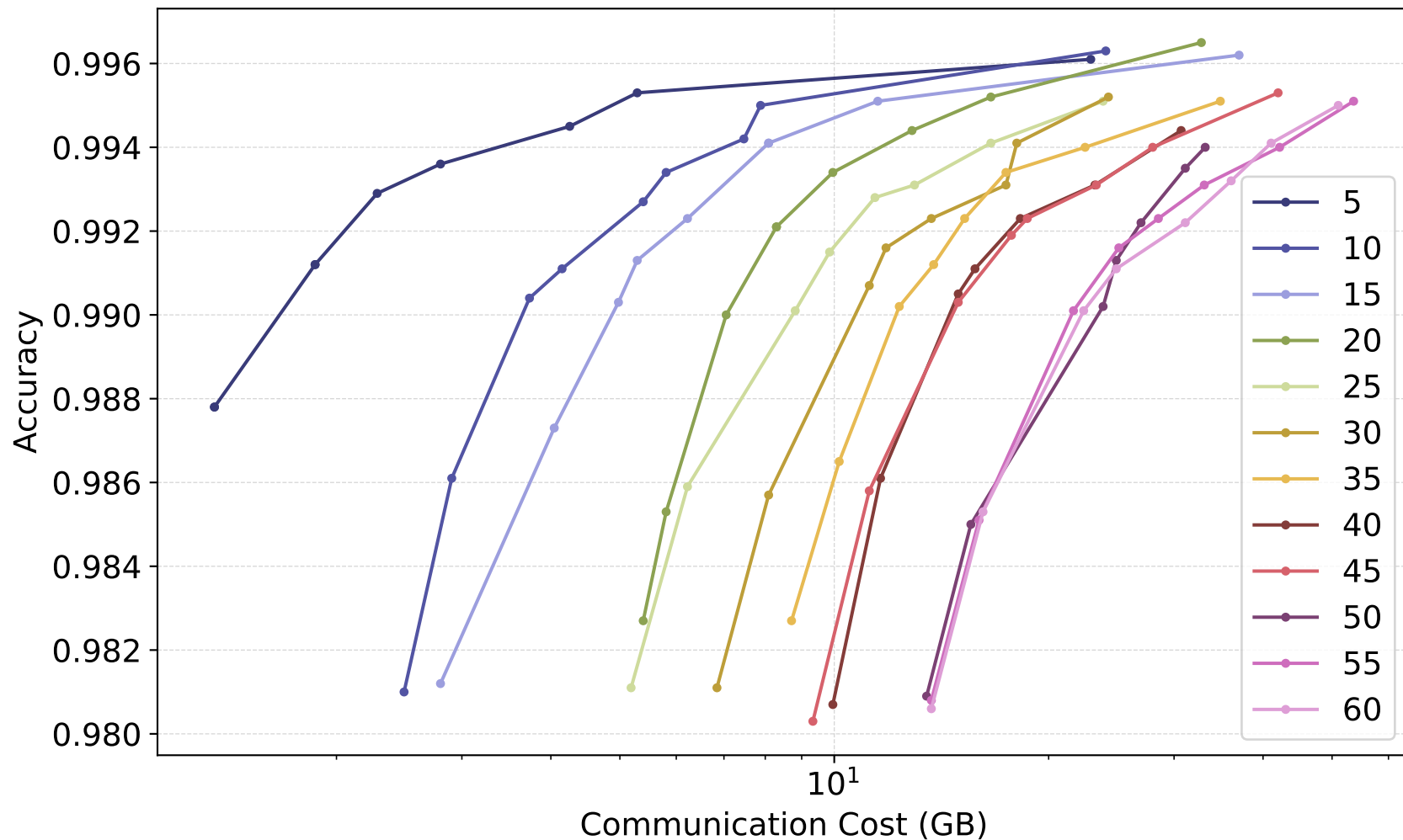
gm  
Theta : 50.0 , Batch Size: 32 , Bias: nan



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

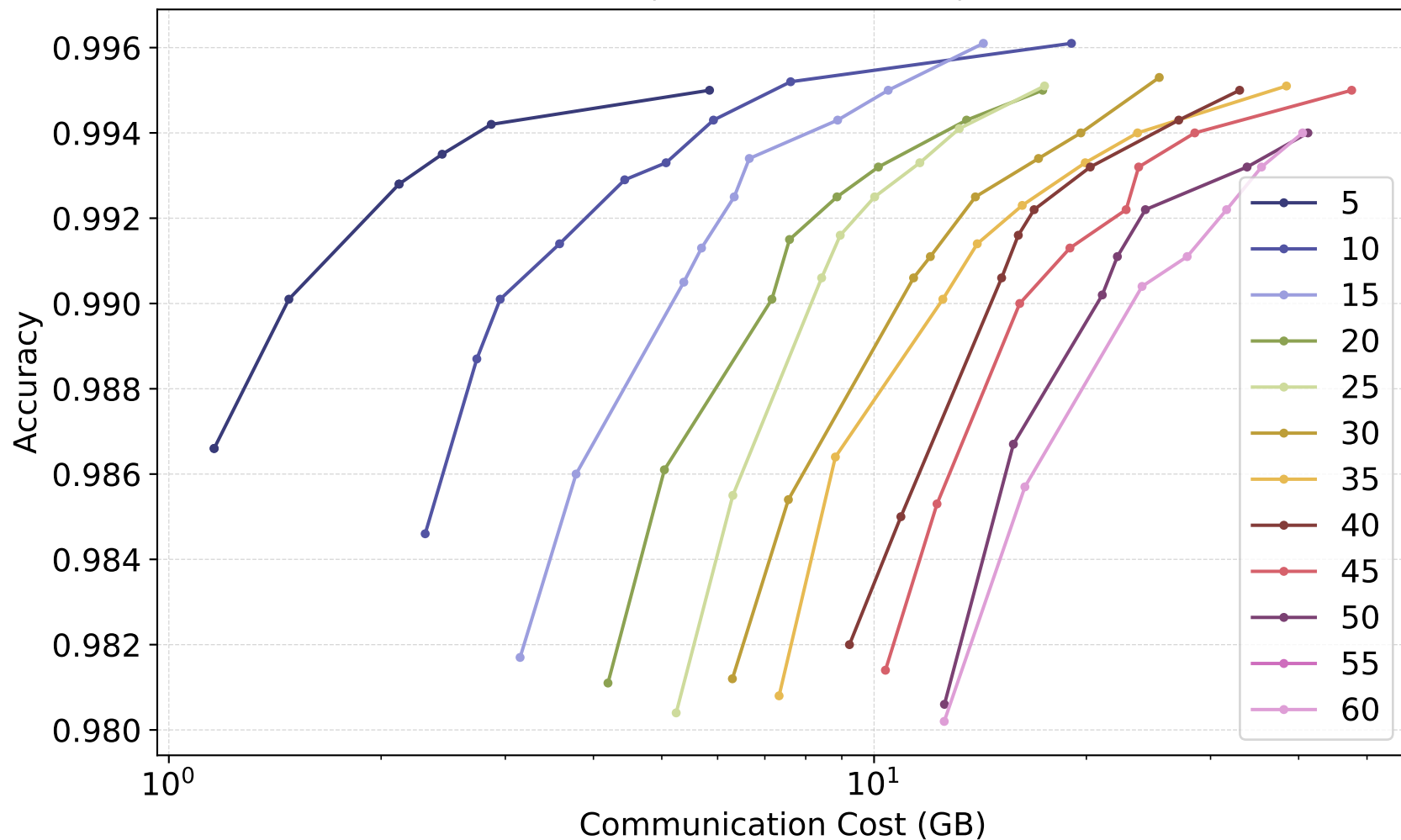


*Theta* : 50.0 , Batch Size: 32 , Bias: nan



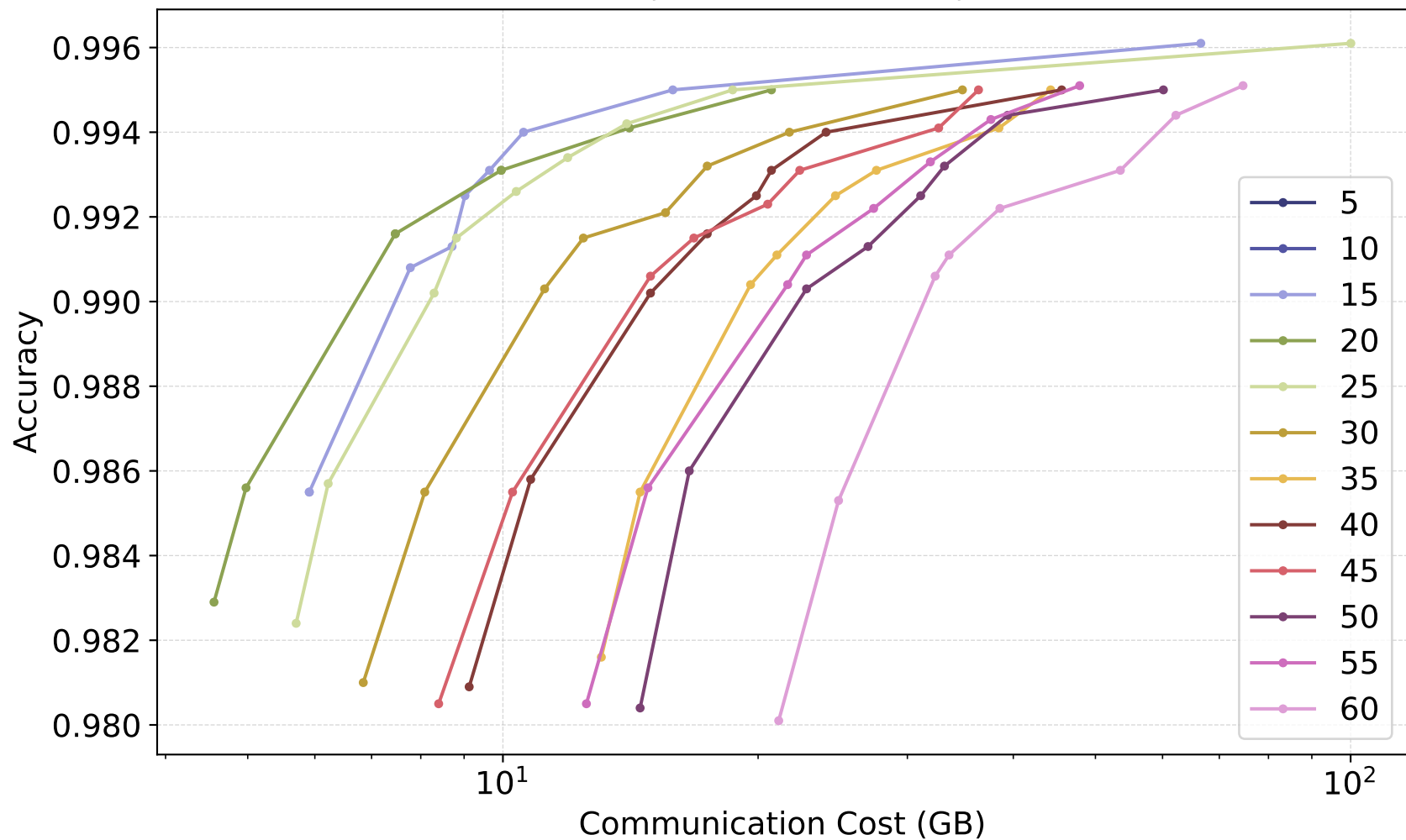
sketch

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

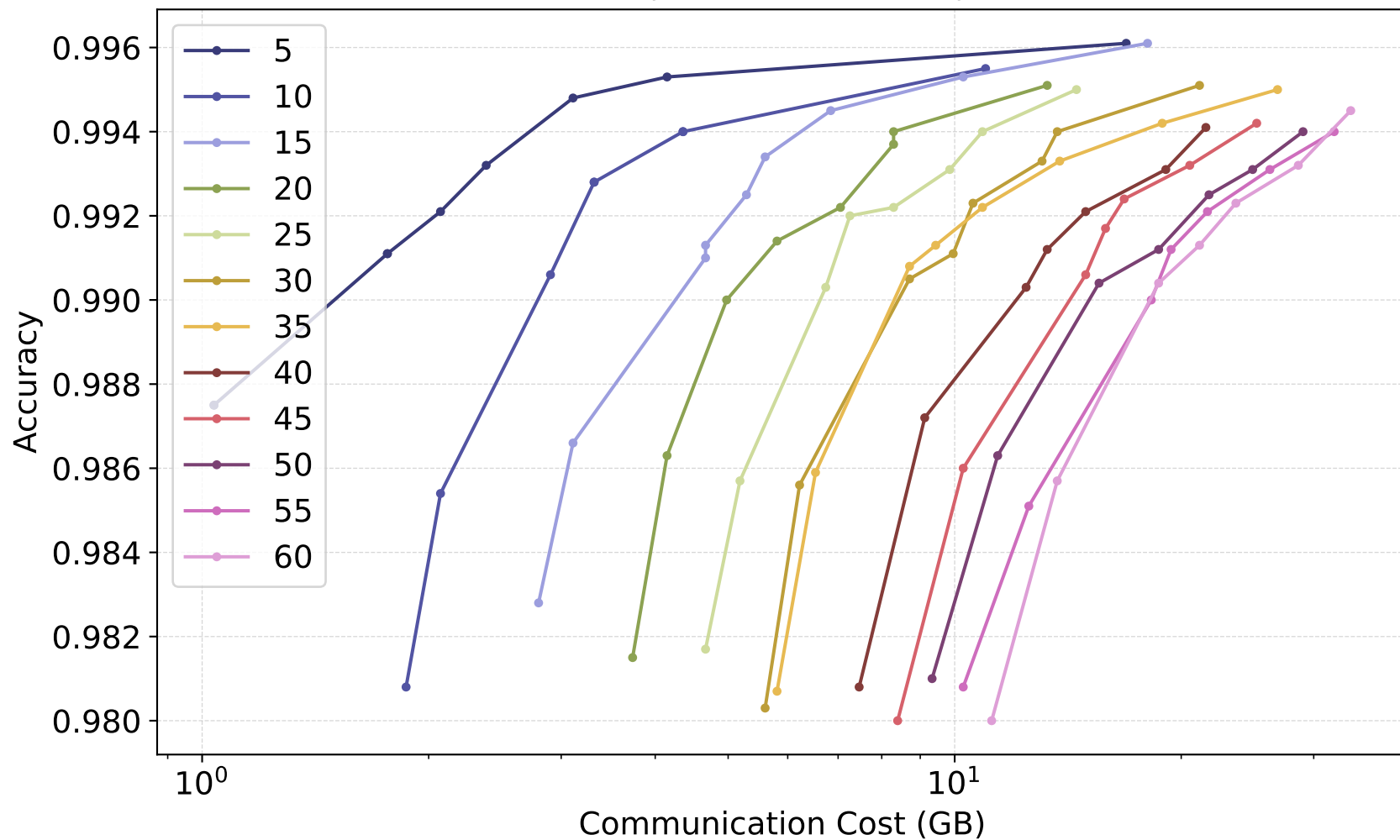


gm

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

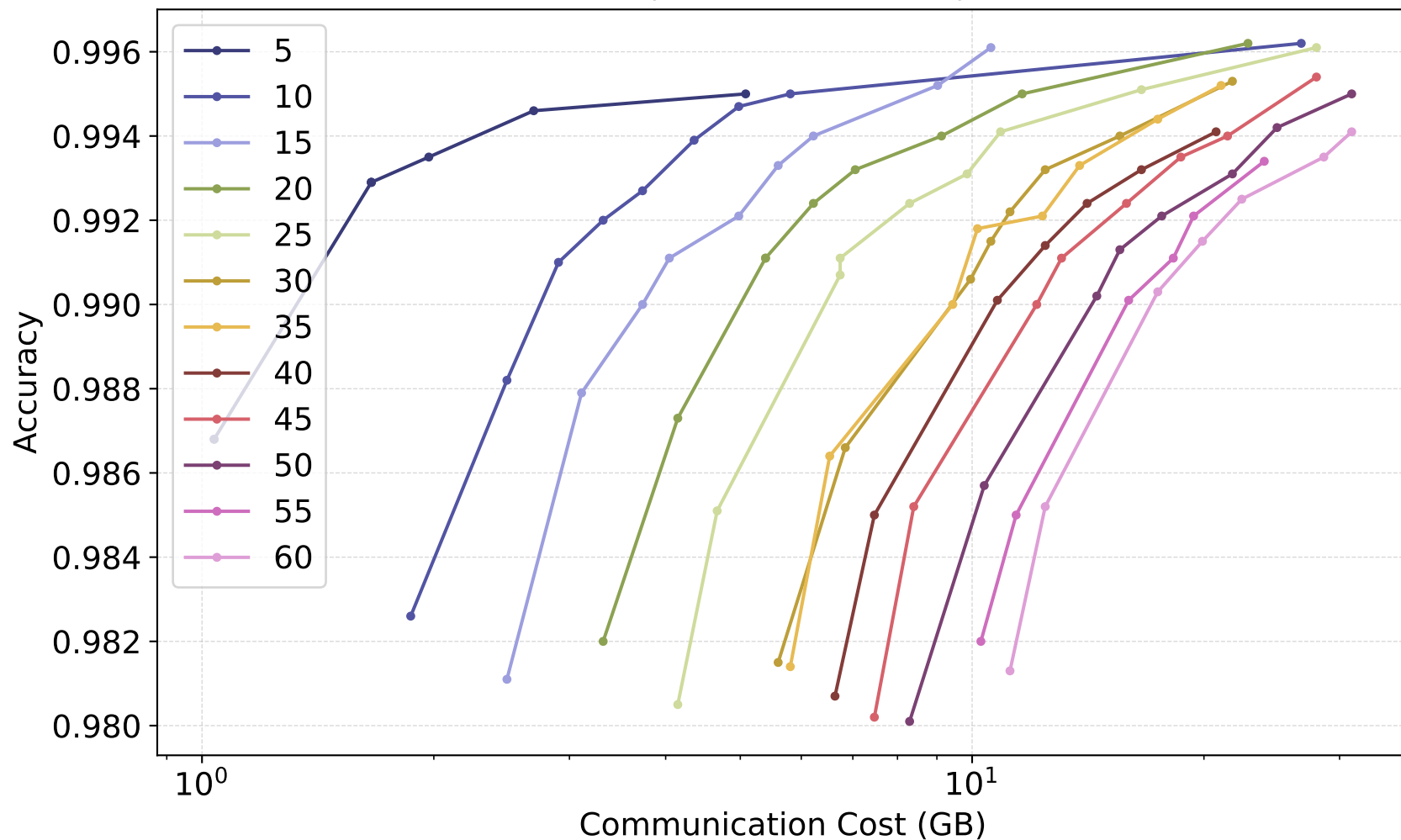


*Theta* : 75.0 , Batch Size: 32 , Bias: nan



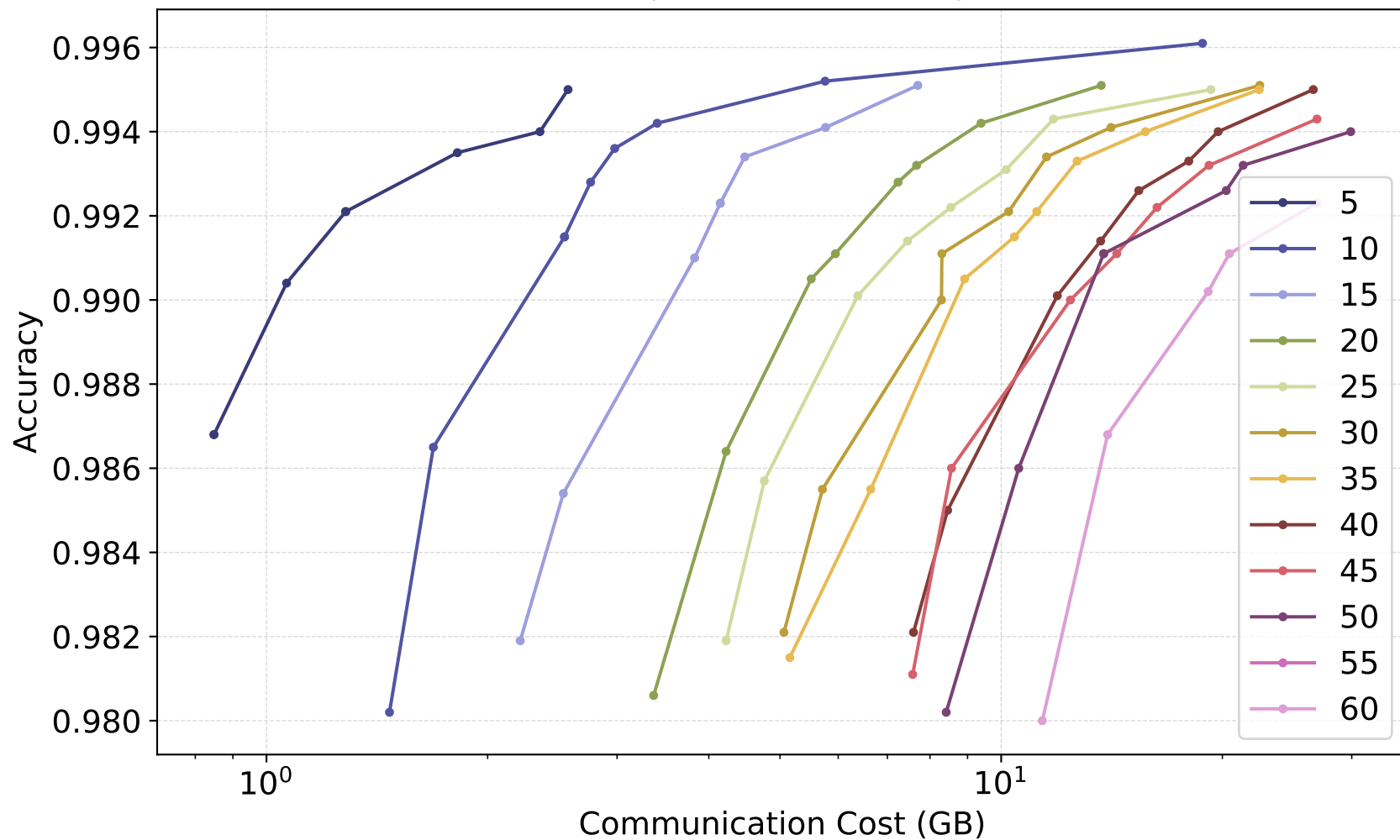
linear

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



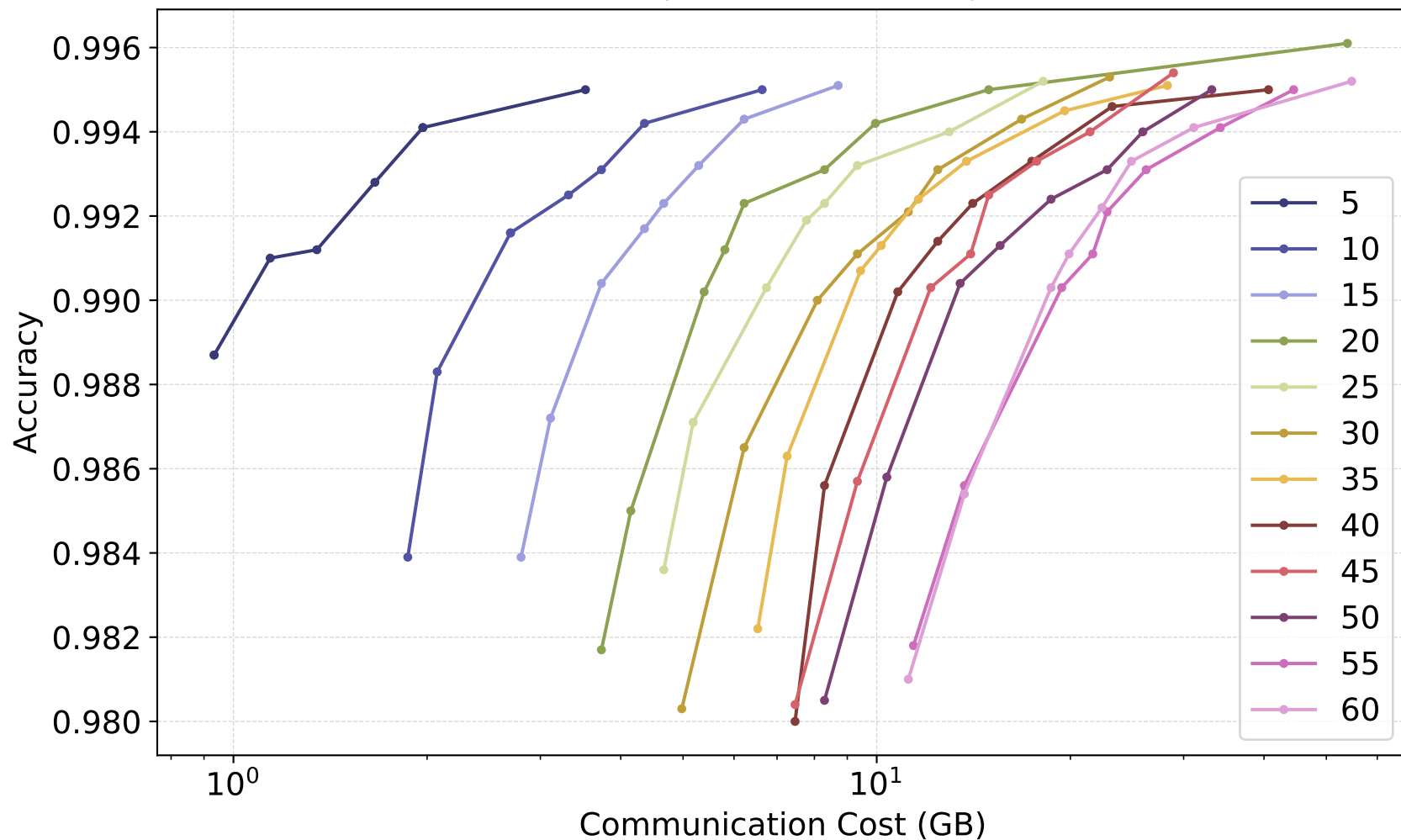


*Theta* : 75.0 , Batch Size: 32 , Bias: nan



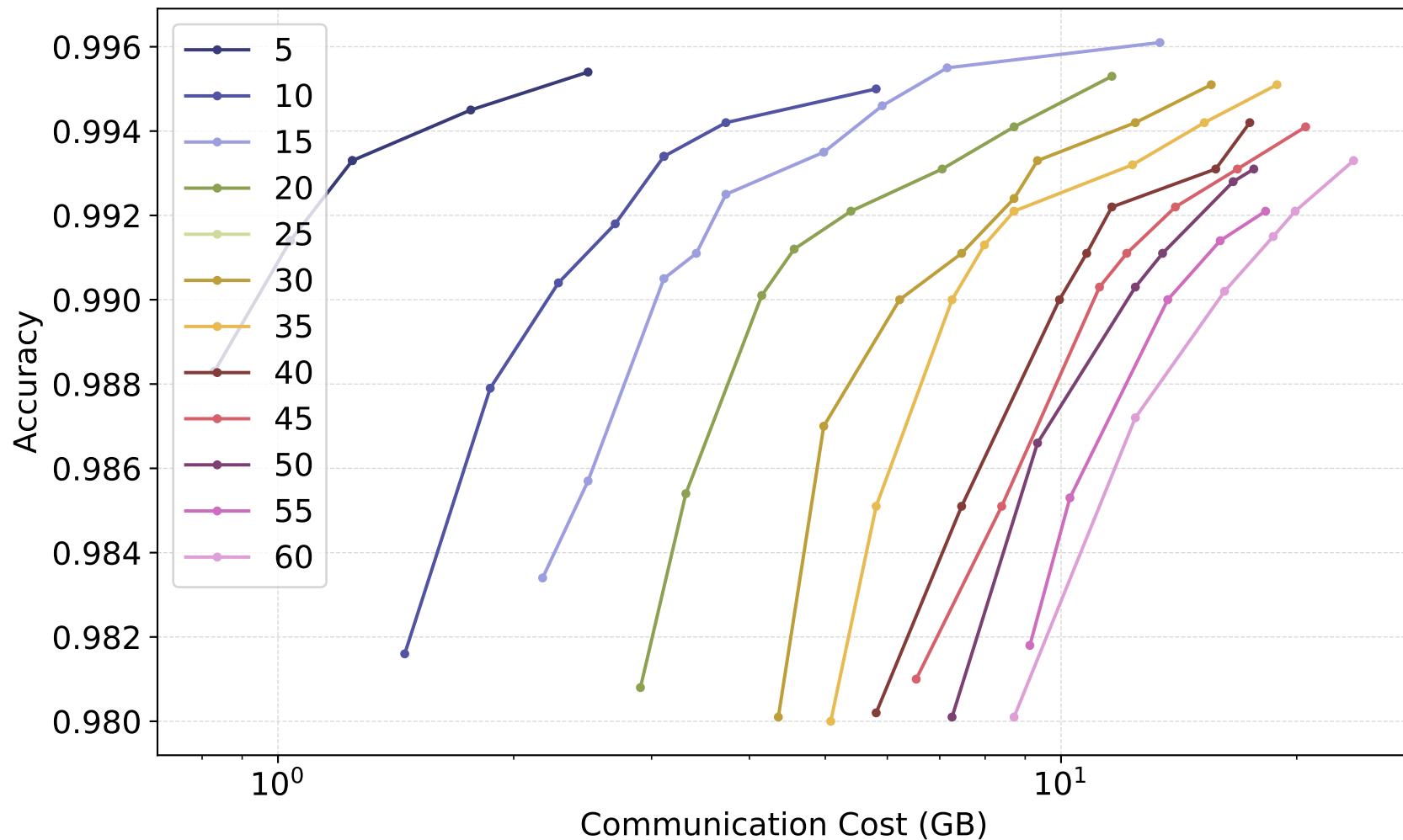
gm

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



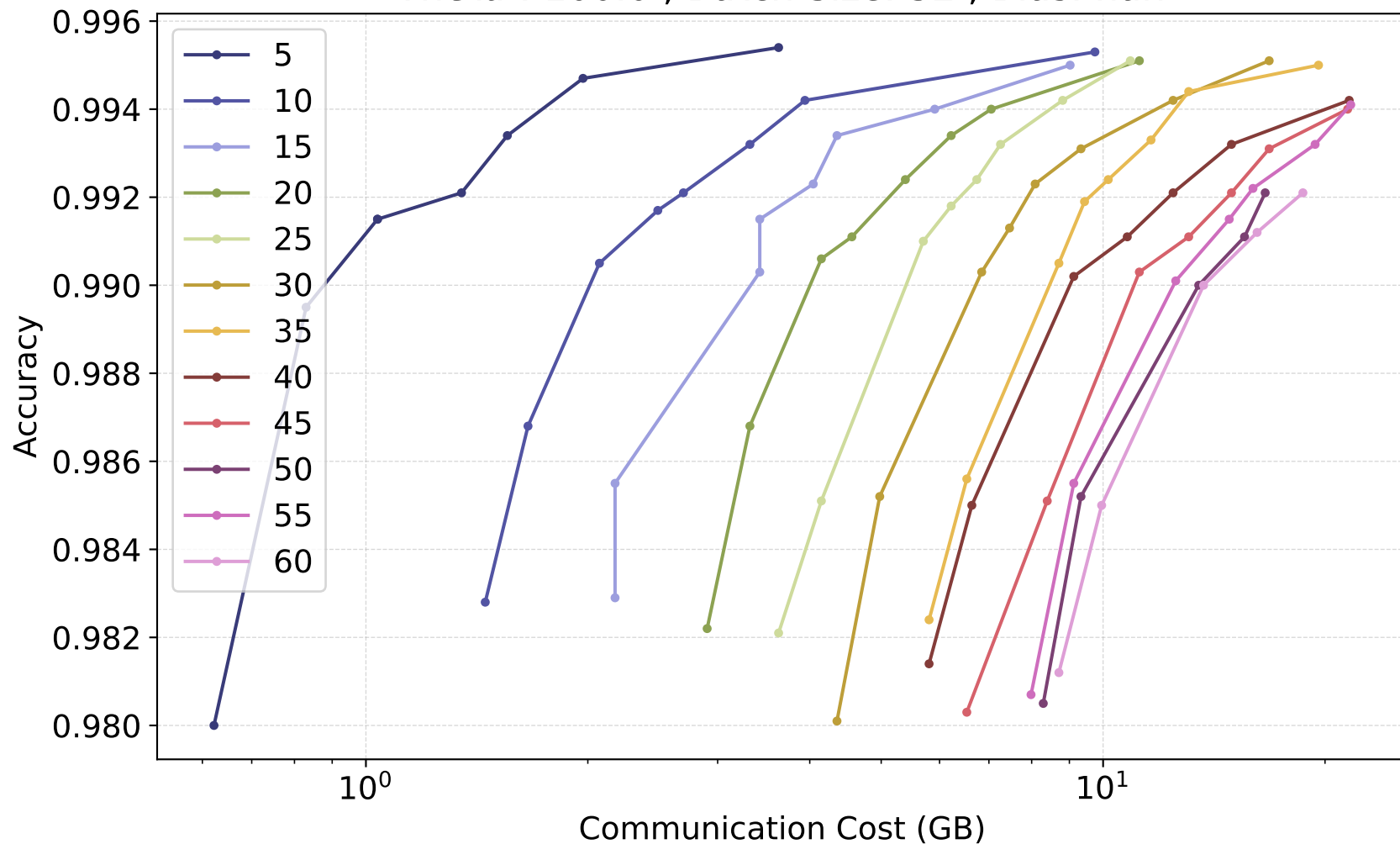
naive

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



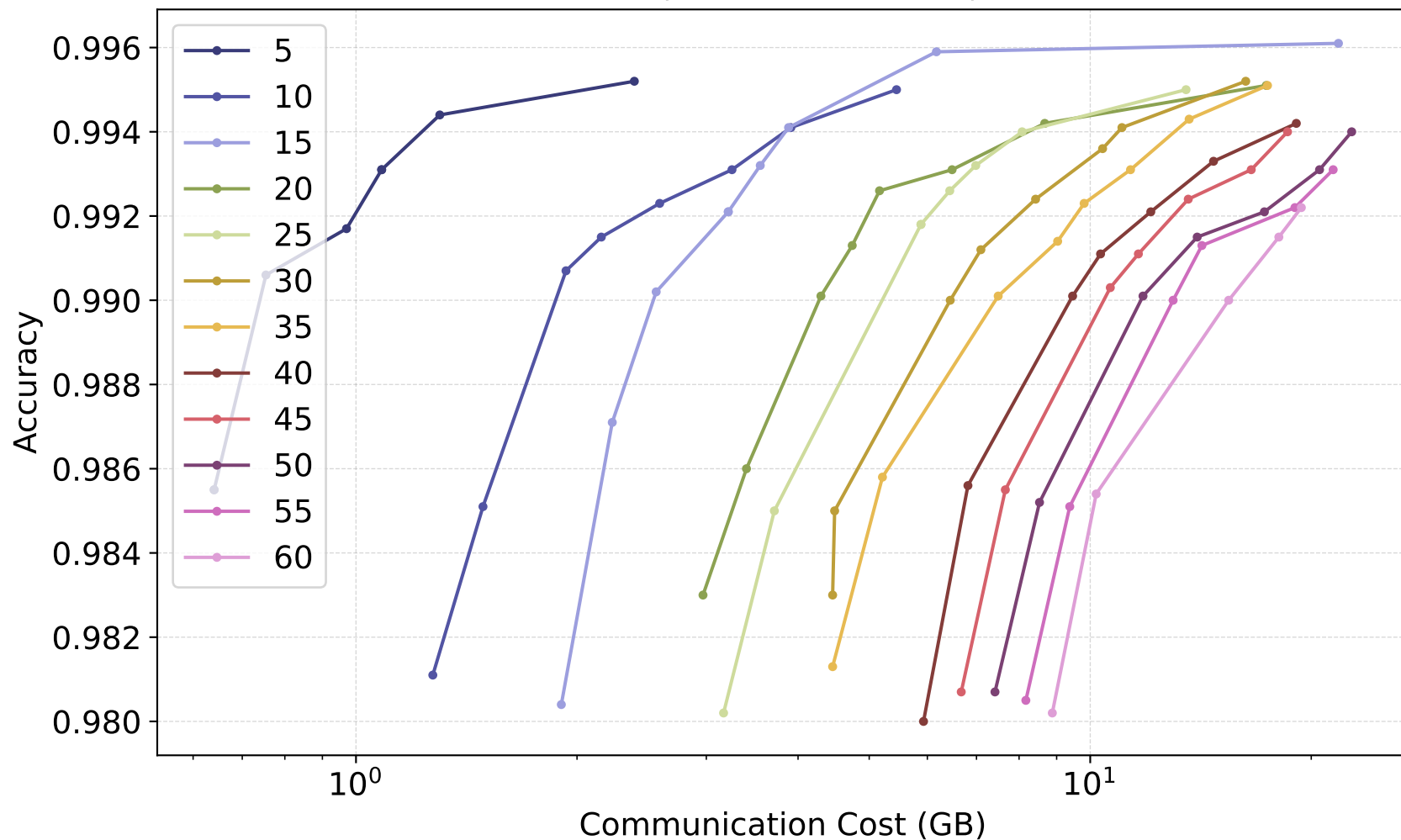
linear

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

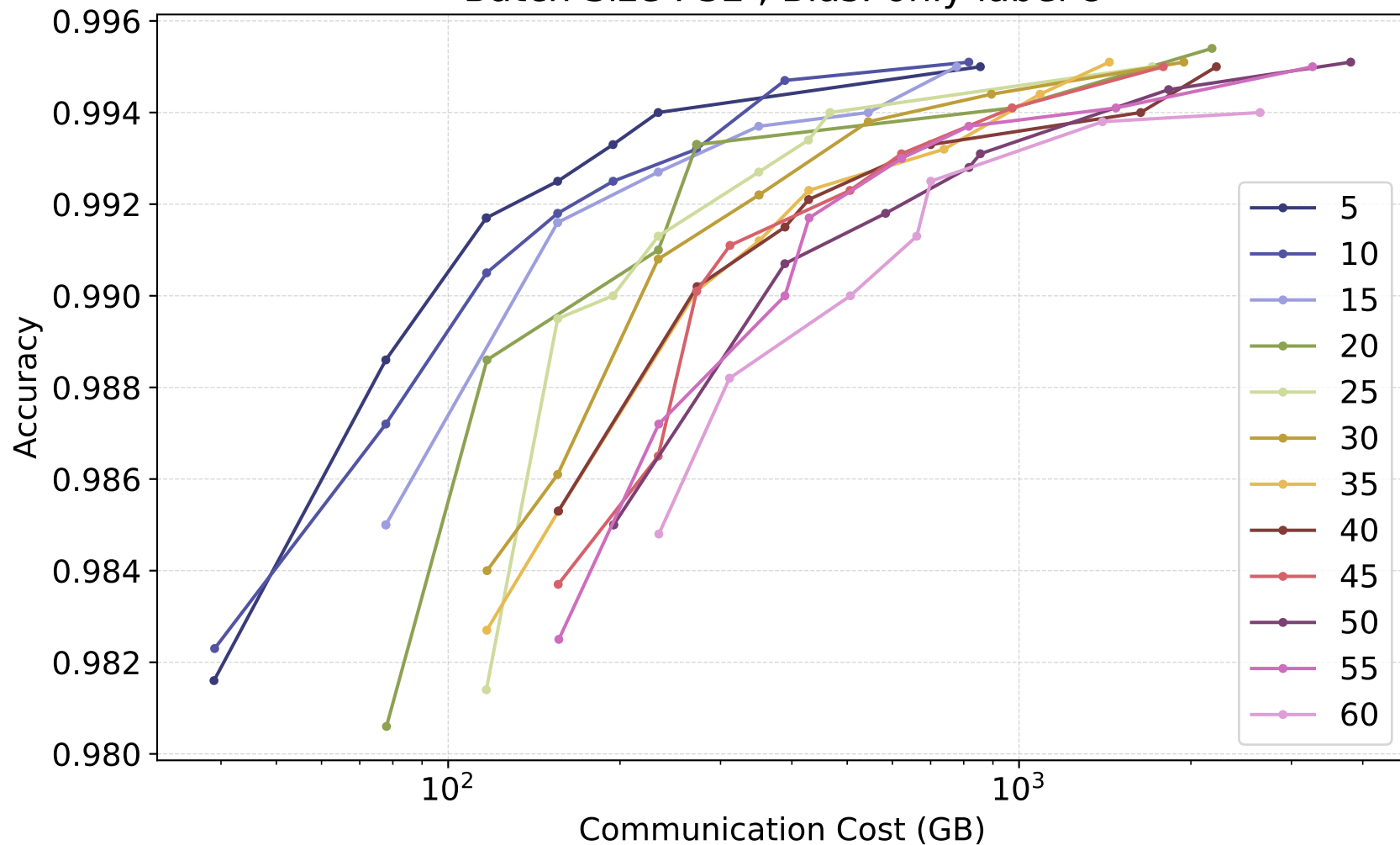


sketch

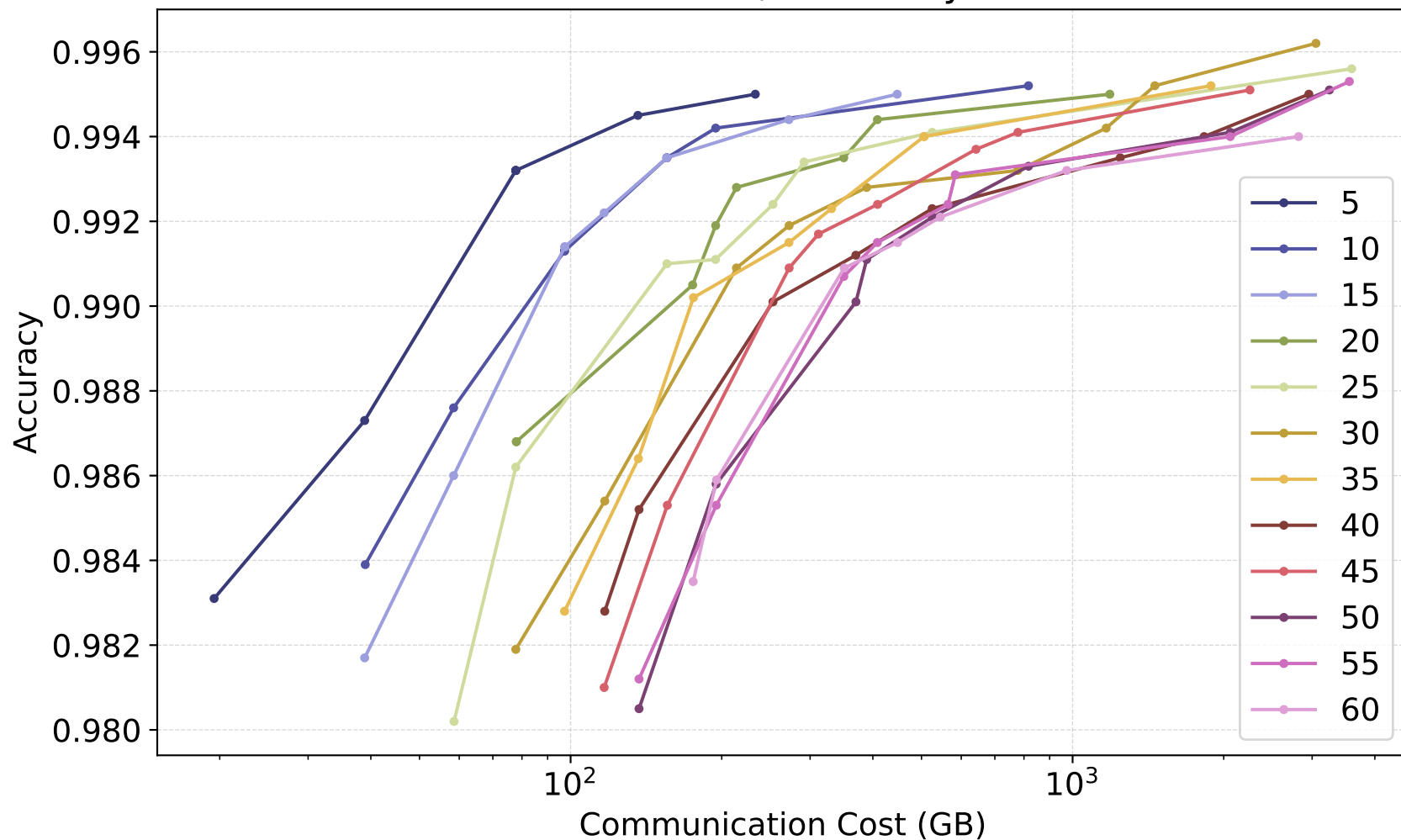
*Theta* : 100.0 , Batch Size: 32 , Bias: nan



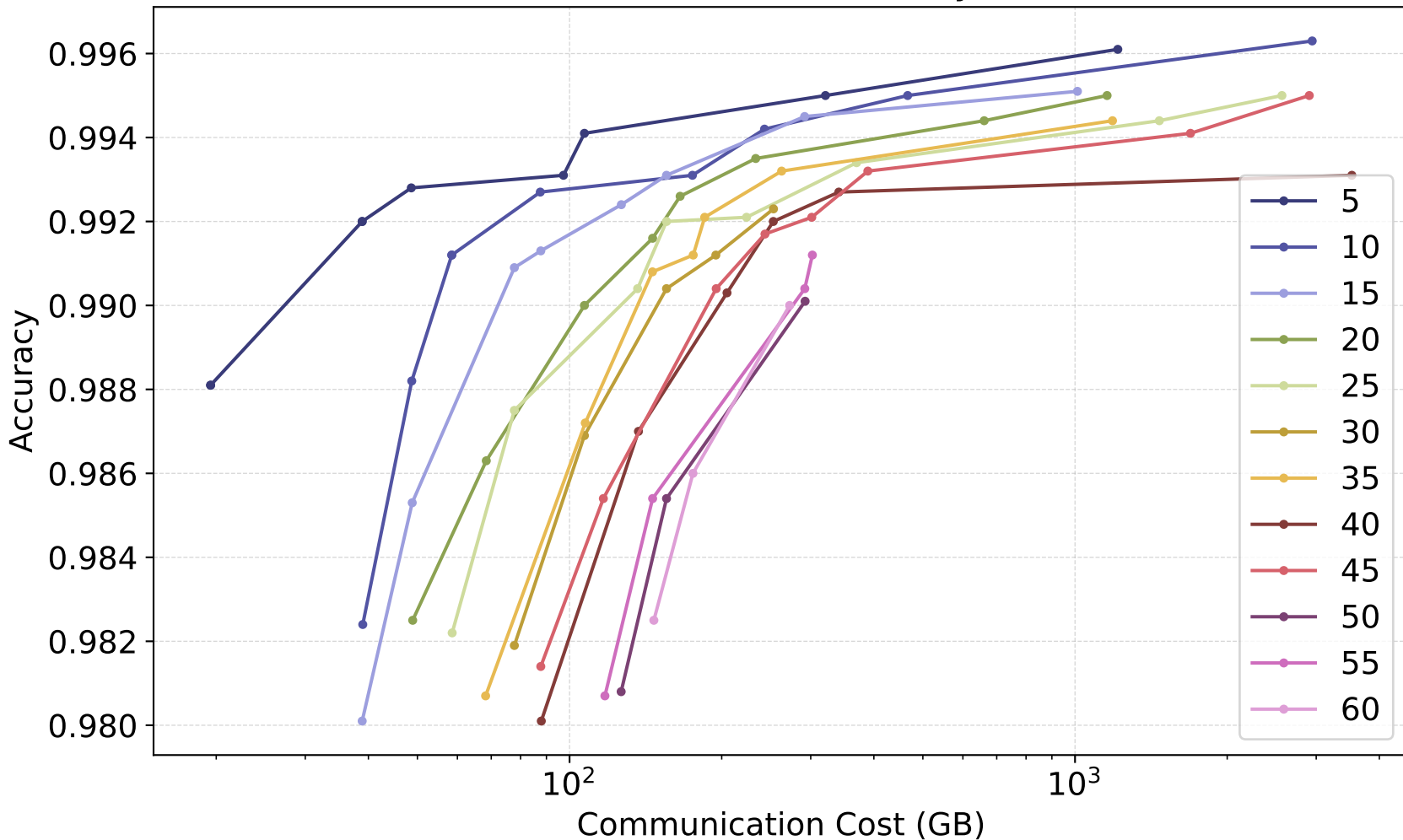
synchronous  
Batch Size : 32 , Bias: only label 8



synchronous  
Batch Size : 64 , Bias: only label 8



synchronous  
Batch Size : 128 , Bias: only label 8



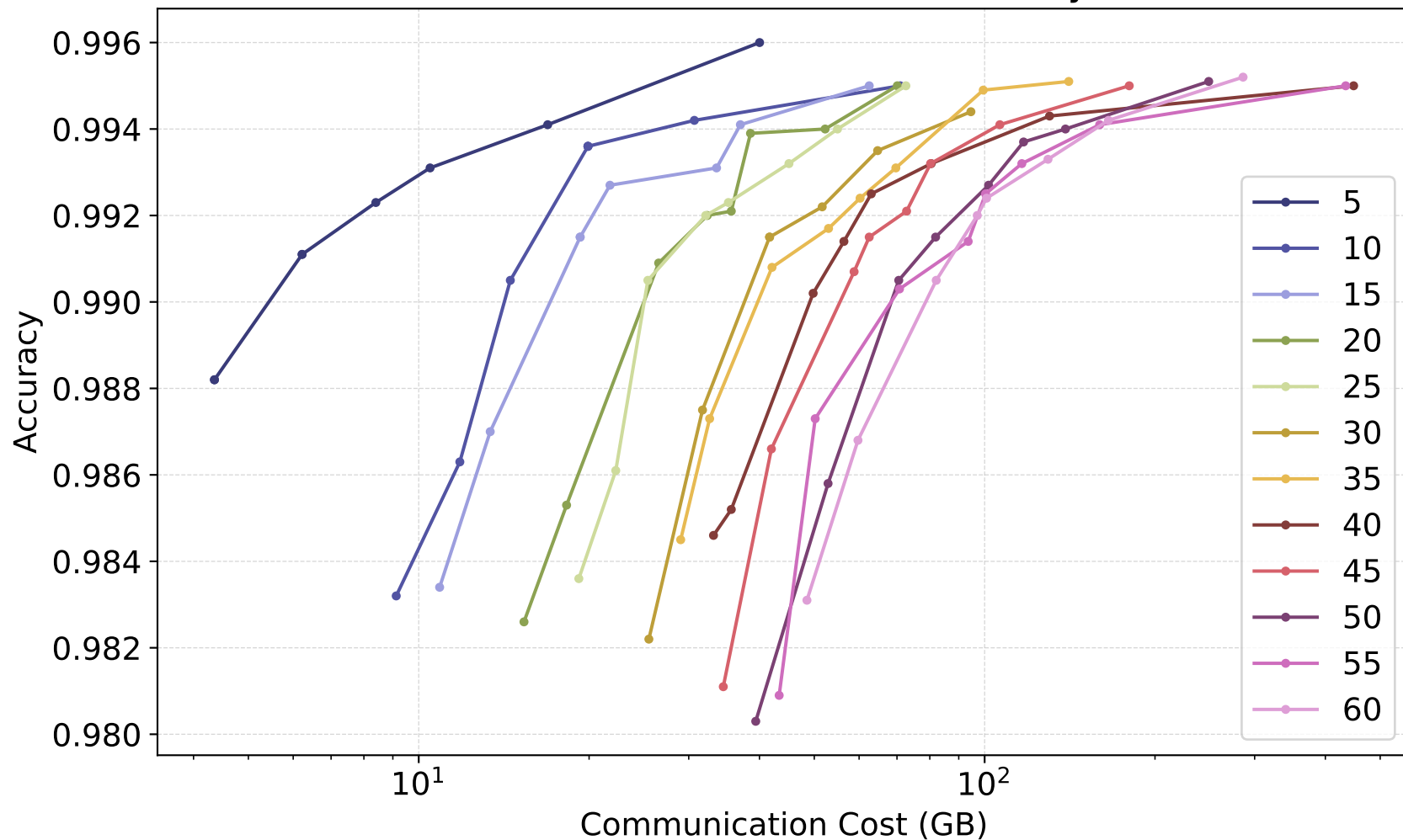


Batch Size : 256 , Bias: only label 8

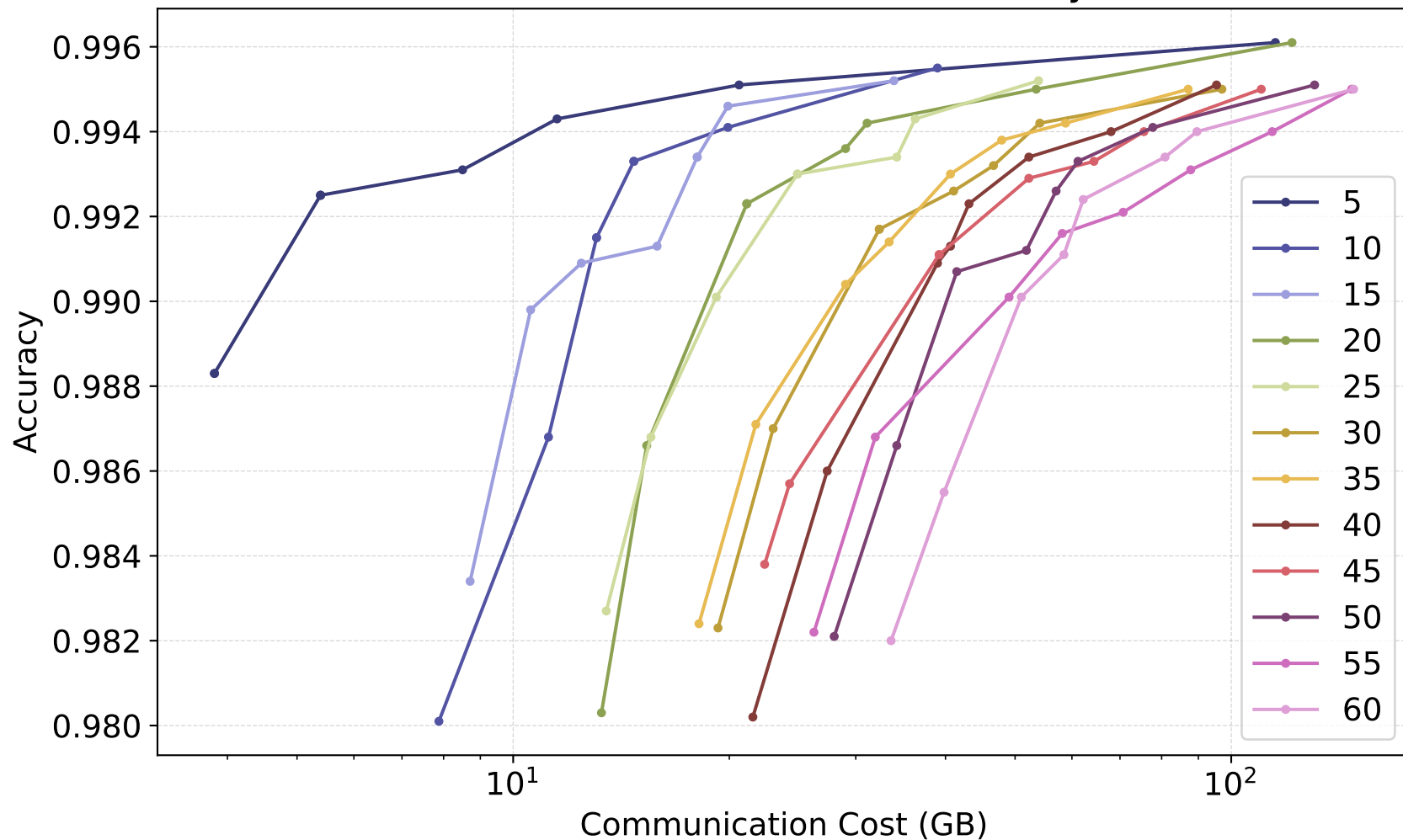


gm

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

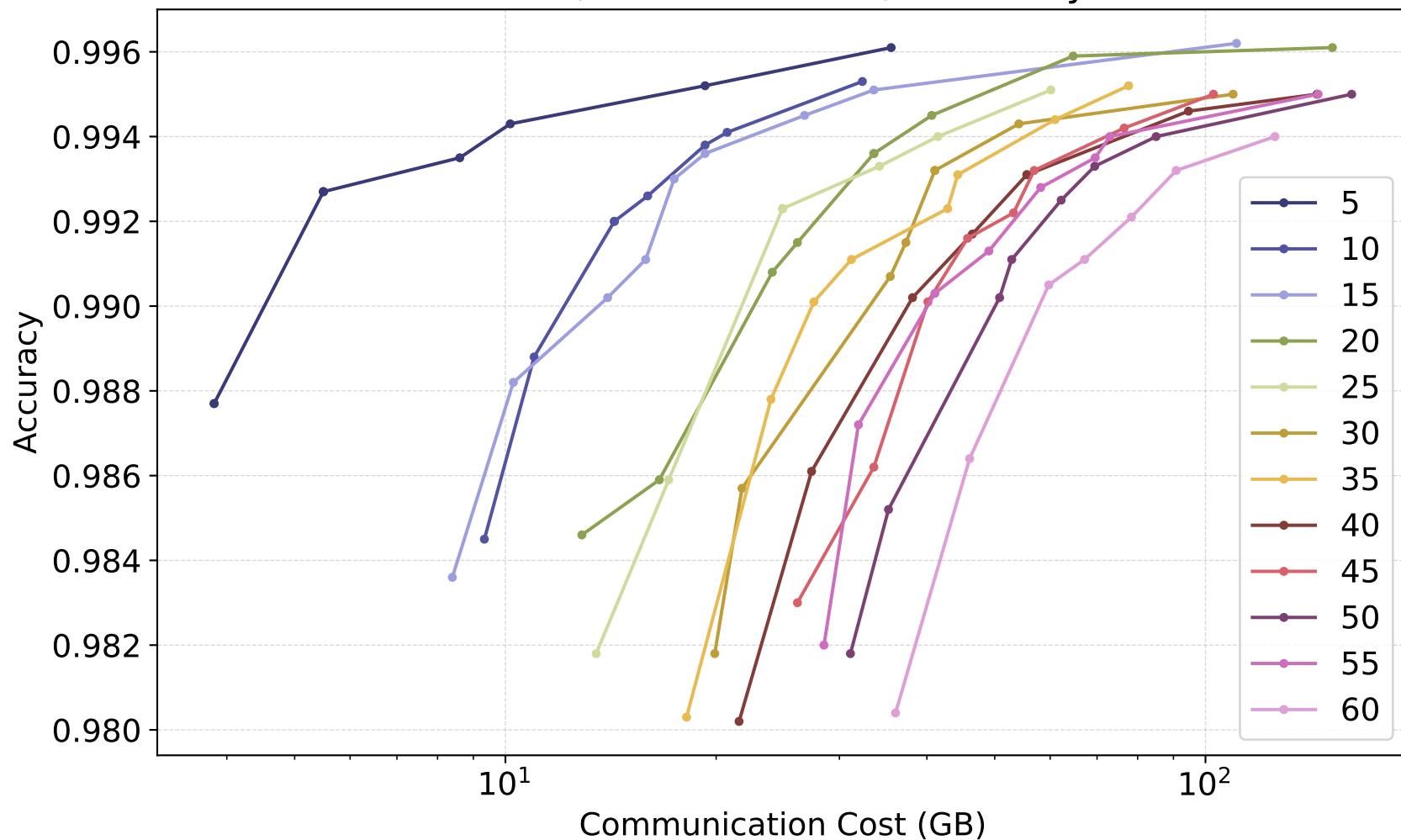


*Theta* : 15.0 , Batch Size: 32 , Bias: only label 8

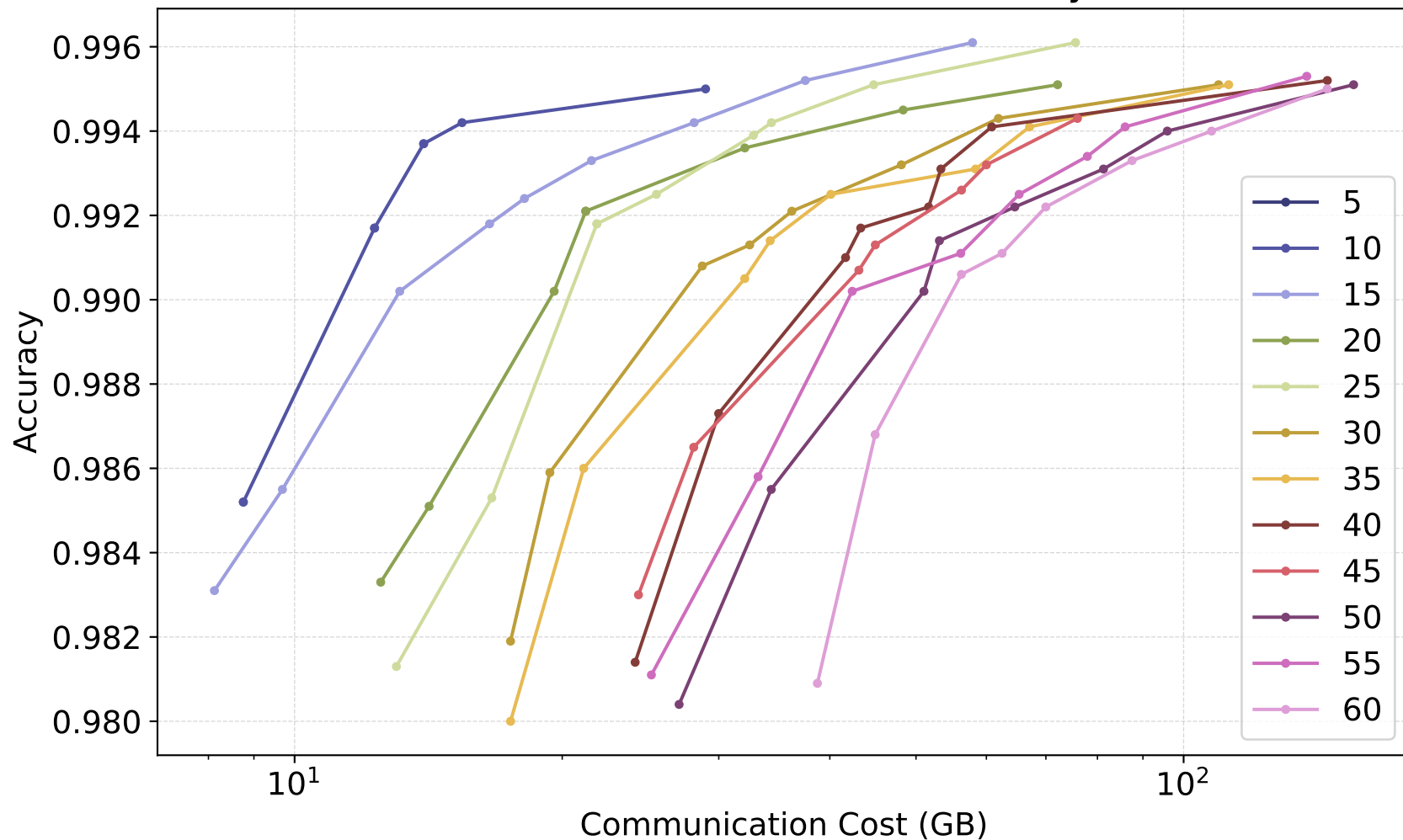


linear

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

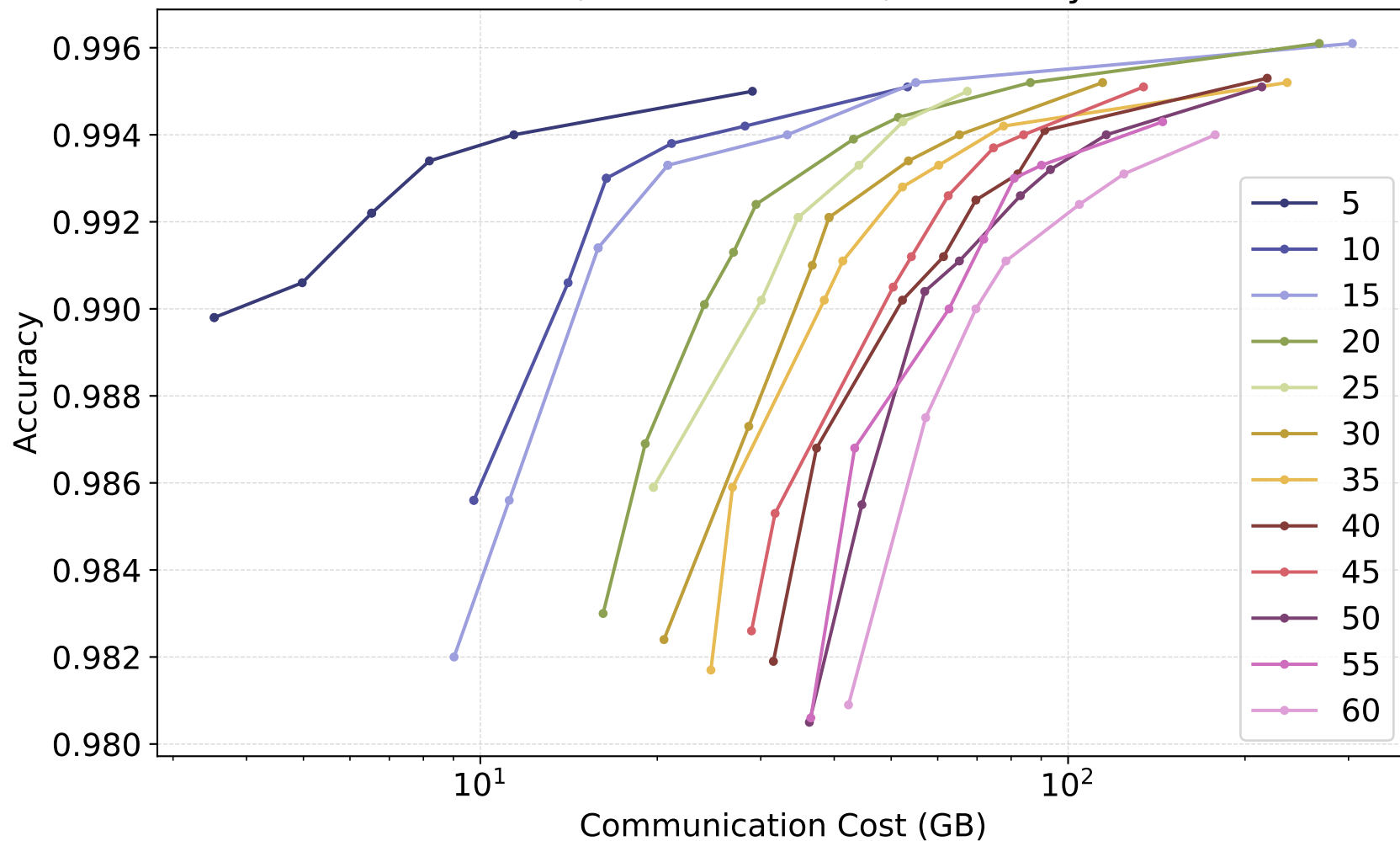


*Theta* : 15.0 , Batch Size: 32 , Bias: only label 8

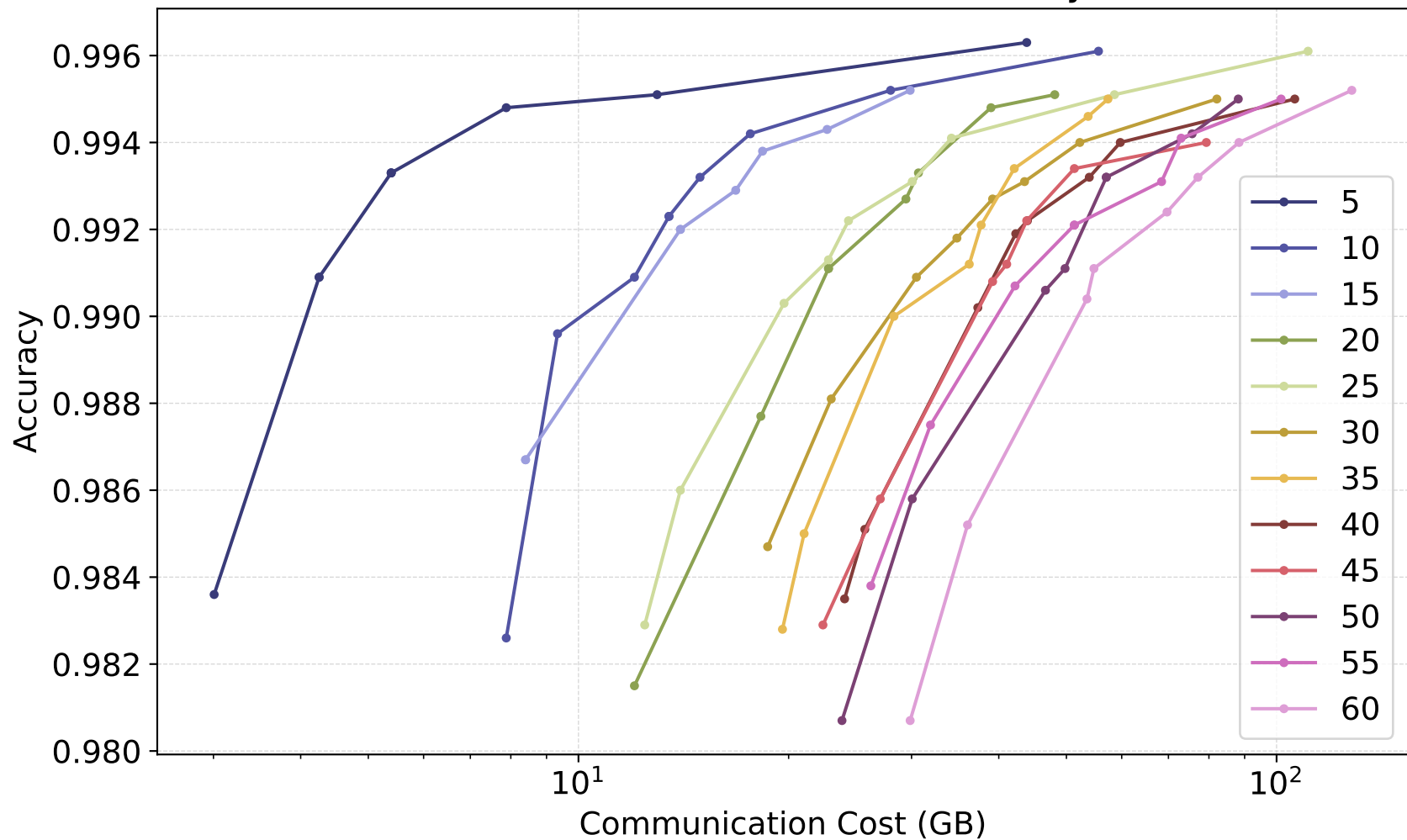


gm

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

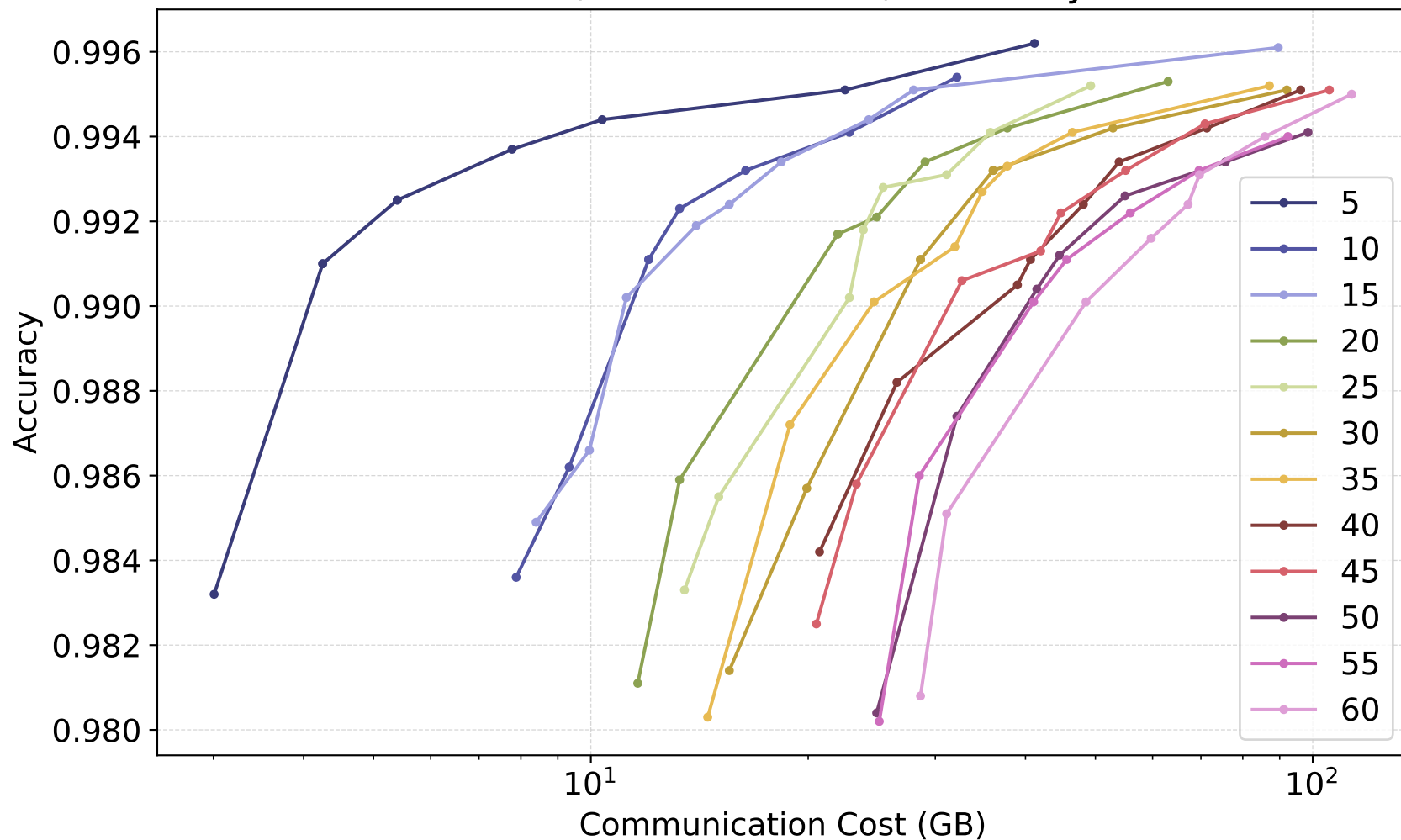


*Theta* : 20.0 , Batch Size: 32 , Bias: only label 8



linear

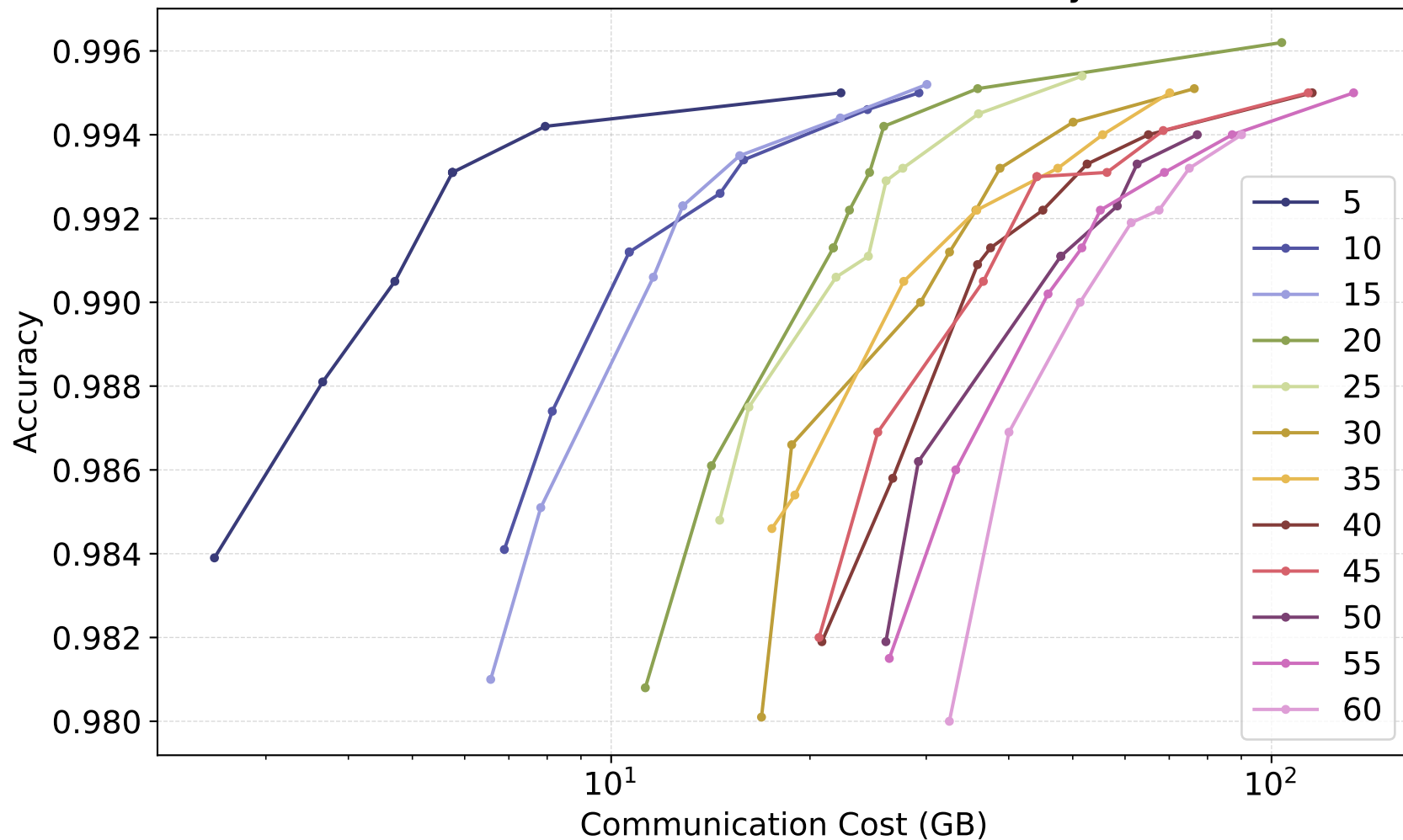
*Theta* : 20.0 , Batch Size: 32 , Bias: only label 8





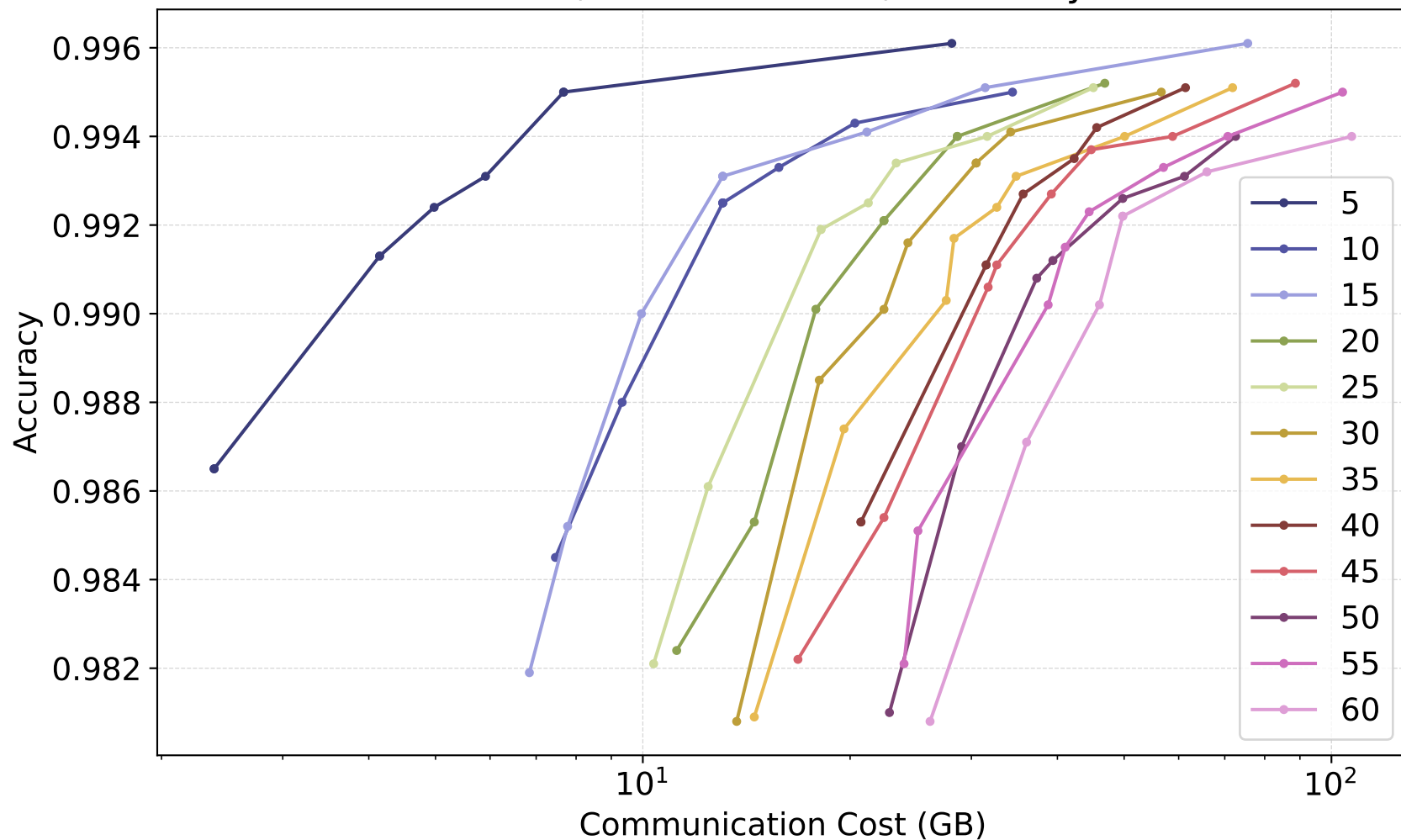
sketch

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



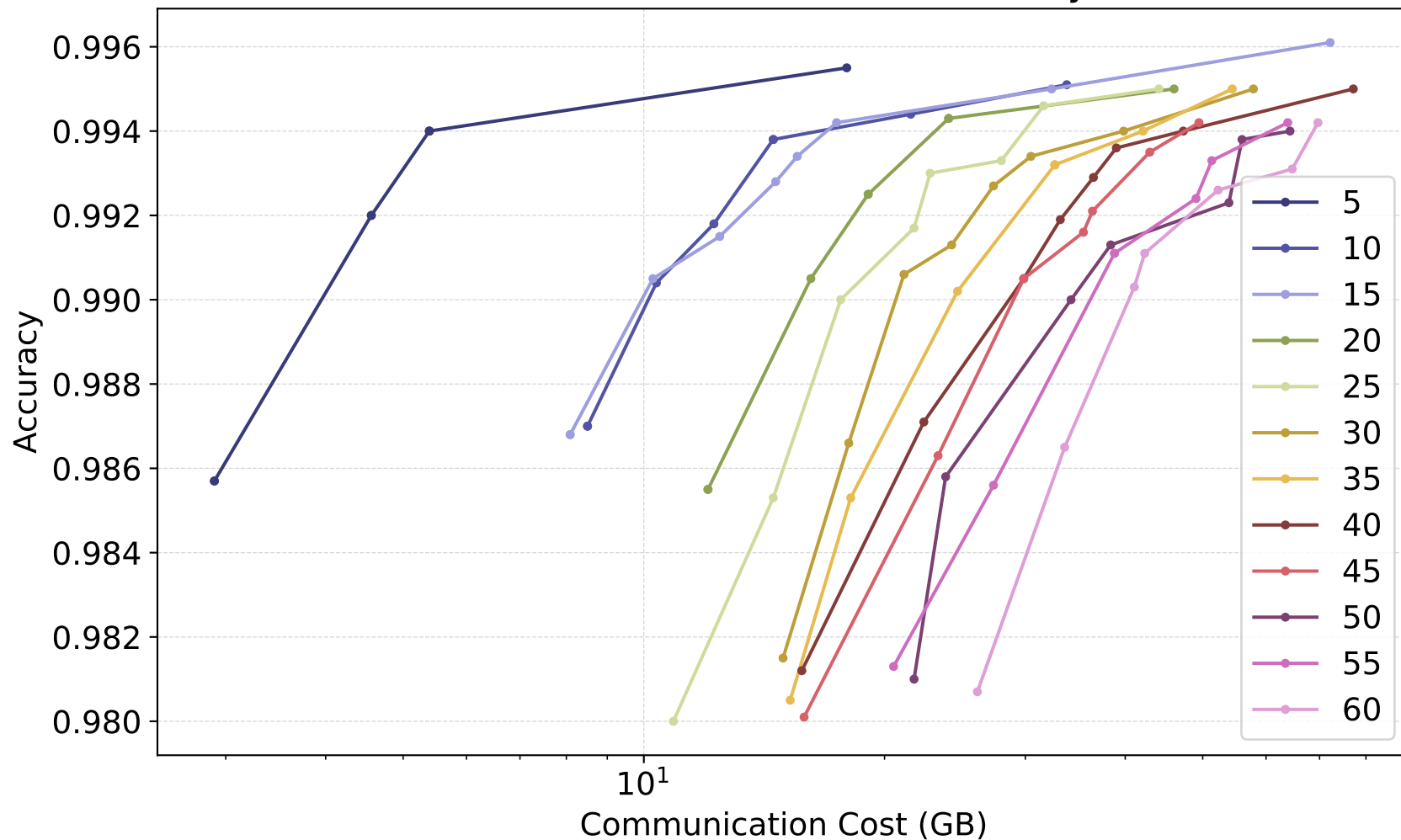


*Theta* : 30.0 , Batch Size: 32 , Bias: only label 8

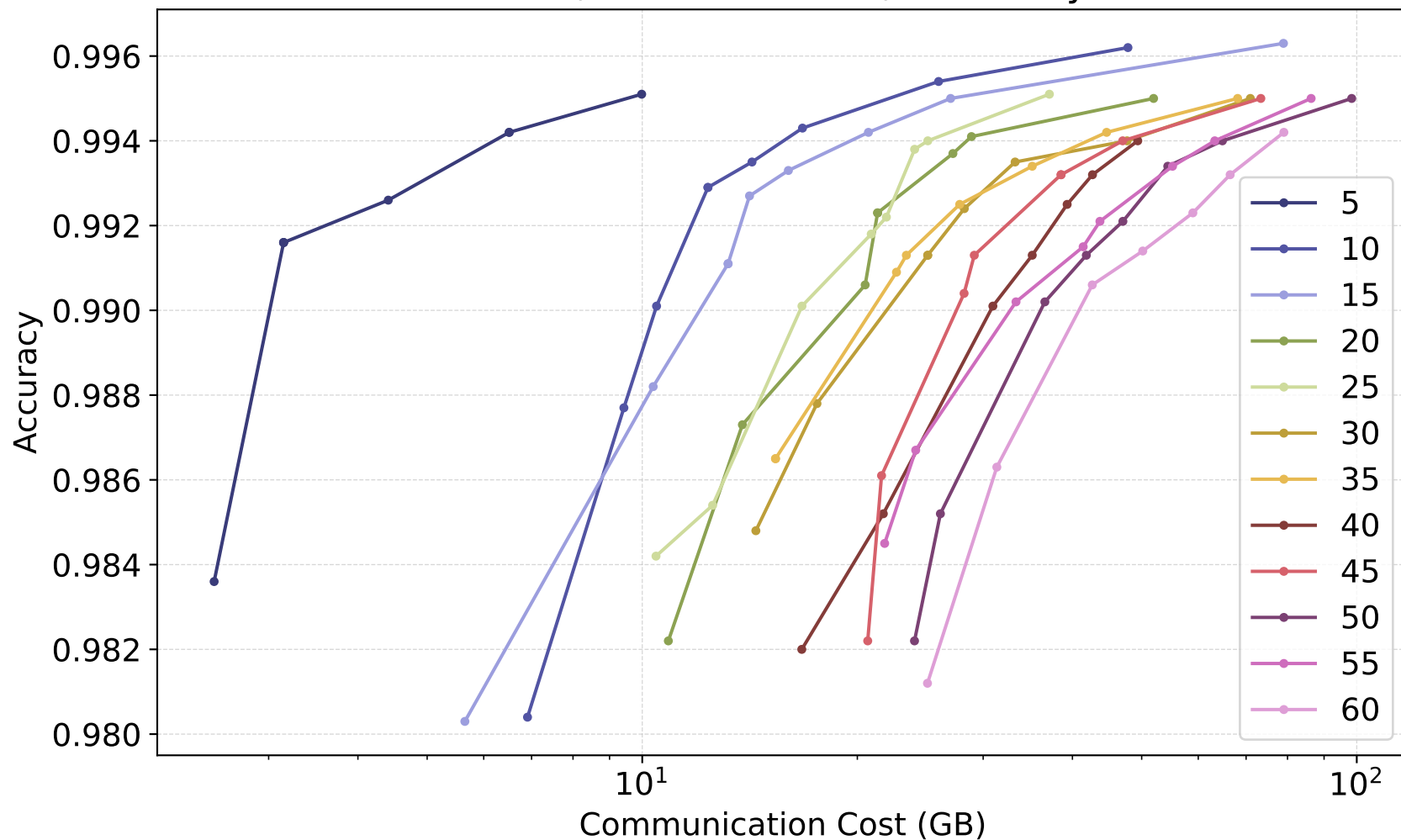


linear

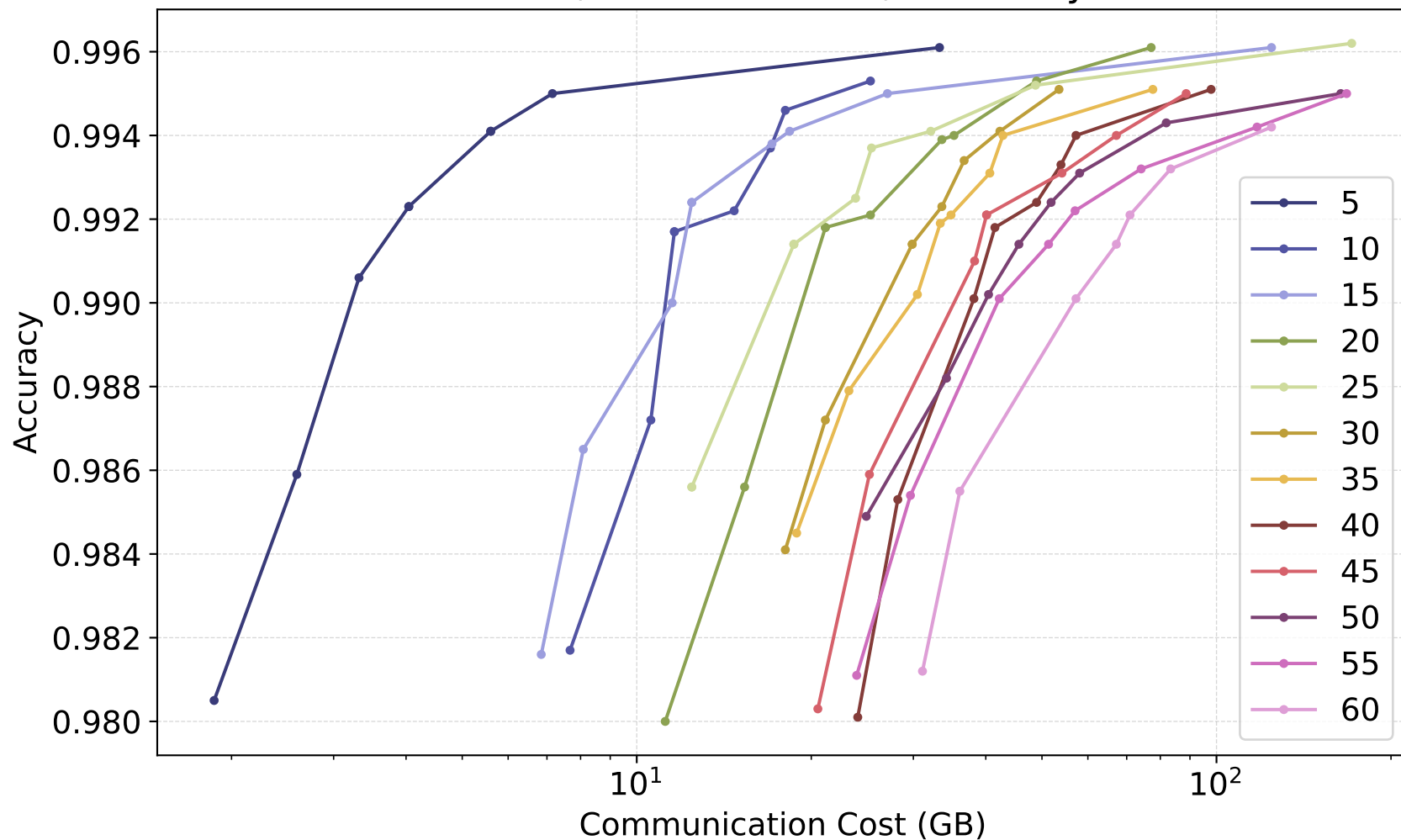
*Theta* : 30.0 , Batch Size: 32 , Bias: only label 8



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

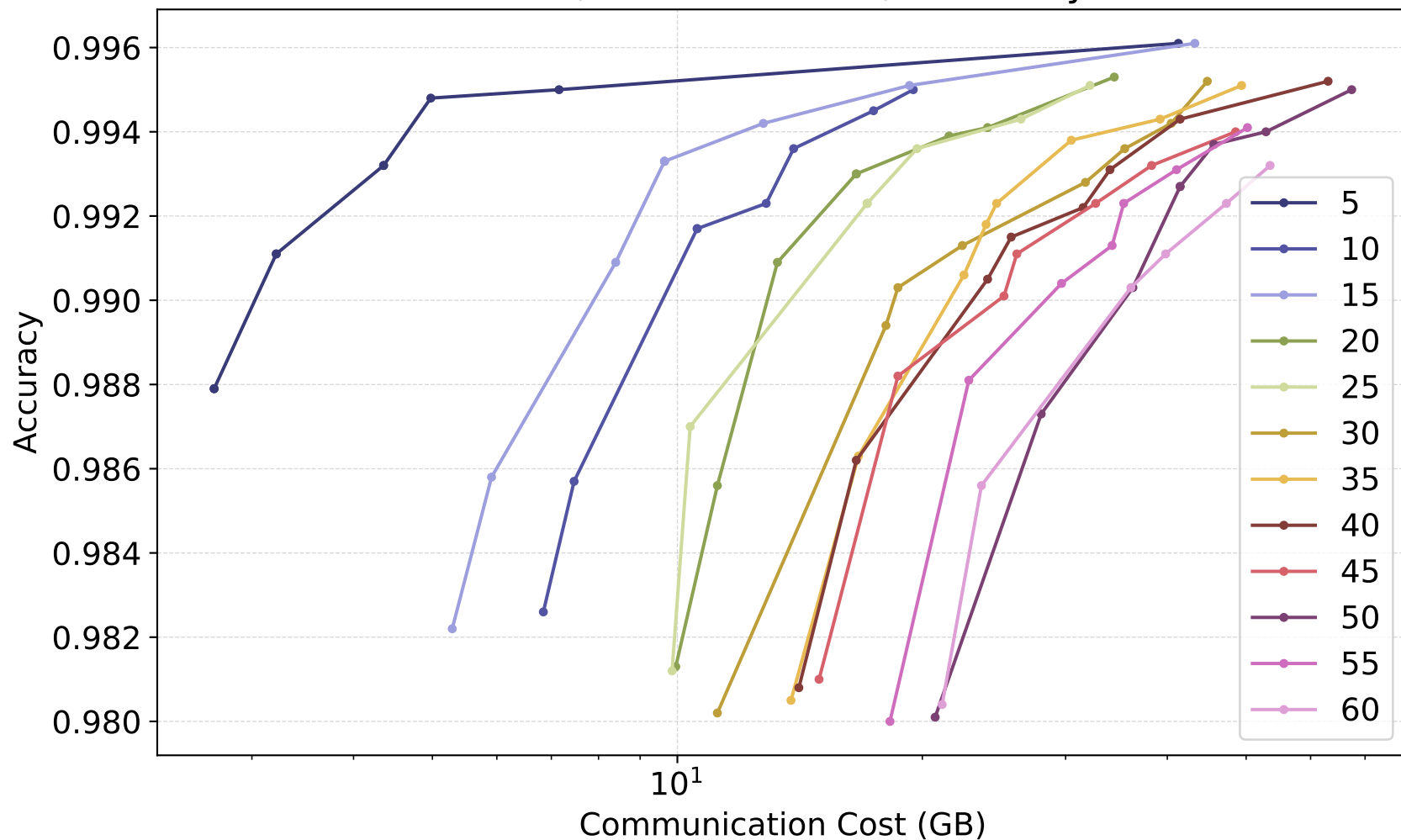


*Theta* : 50.0 , Batch Size: 32 , Bias: only label 8

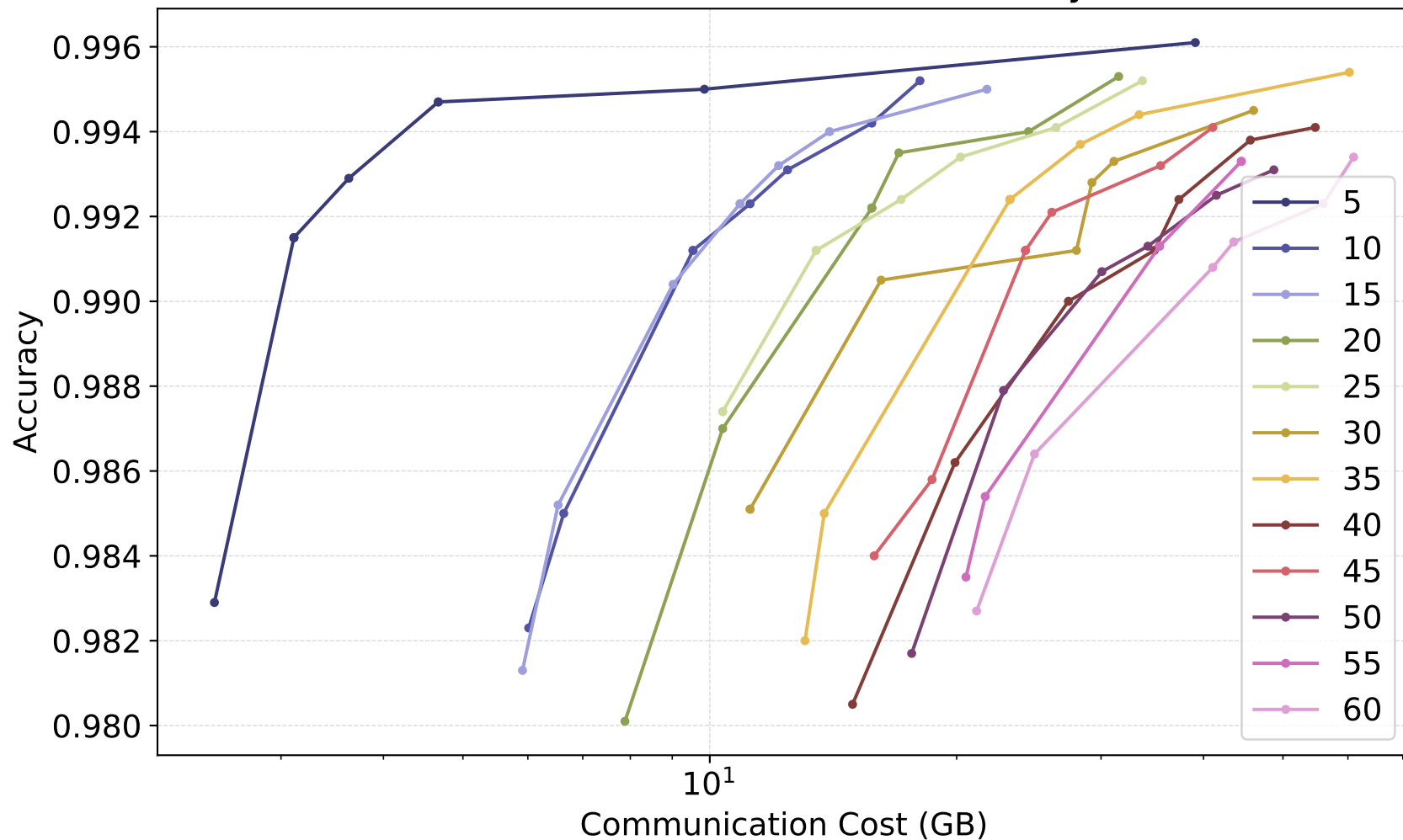


naive

*Theta* : 50.0 , Batch Size: 32 , Bias: only label 8



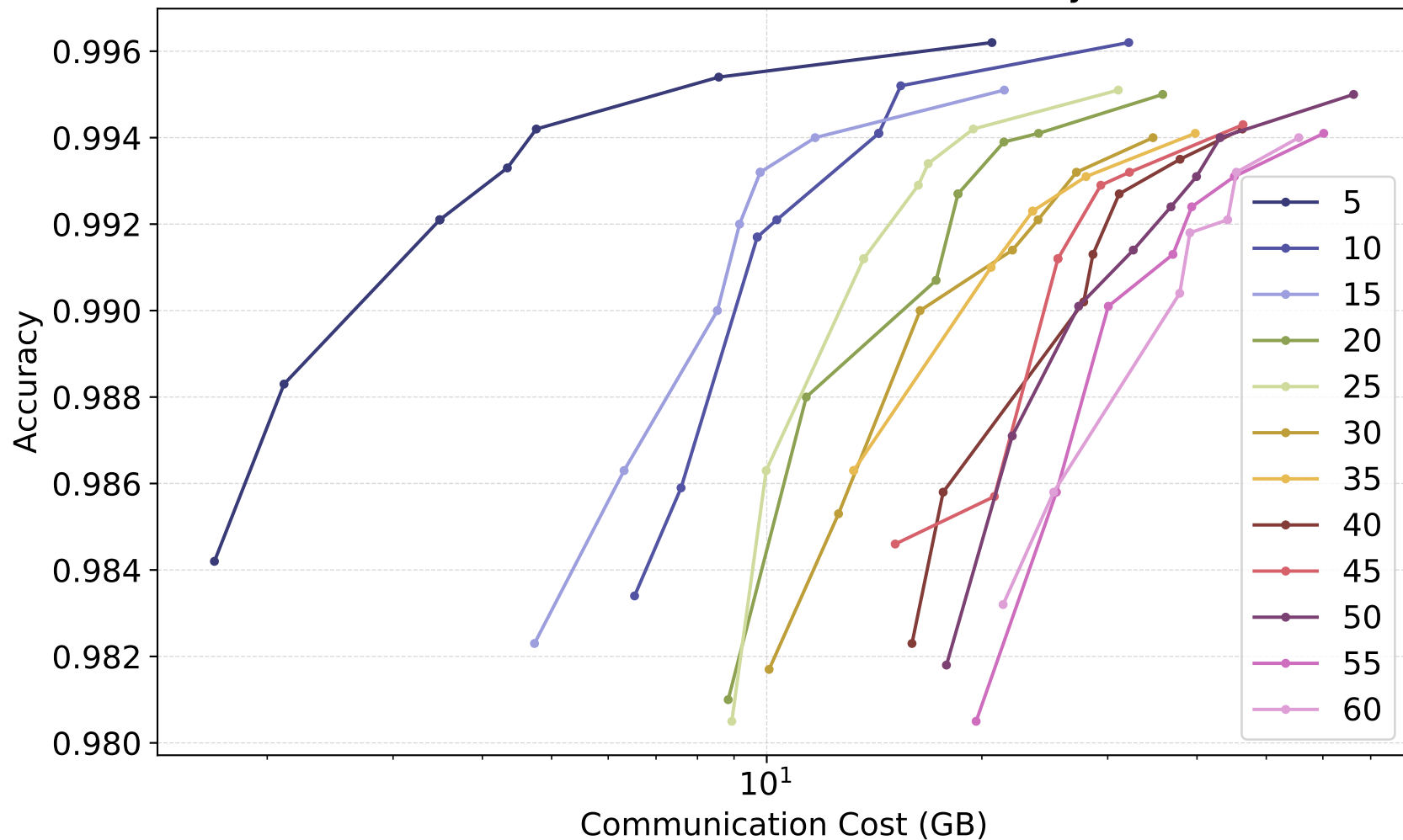
*Theta* : 50.0 , Batch Size: 32 , Bias: only label 8





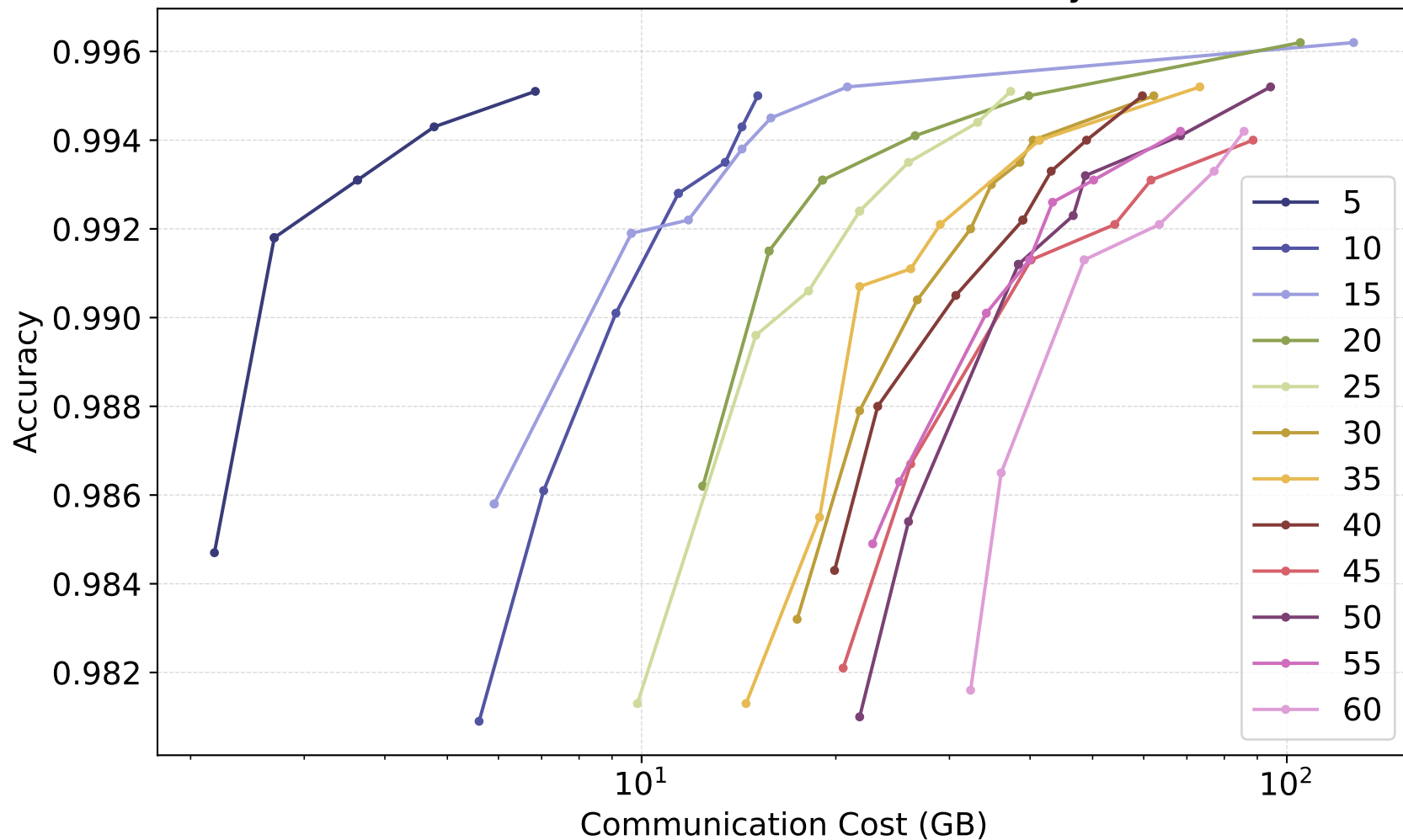
sketch

*Theta* : 50.0 , Batch Size: 32 , Bias: only label 8

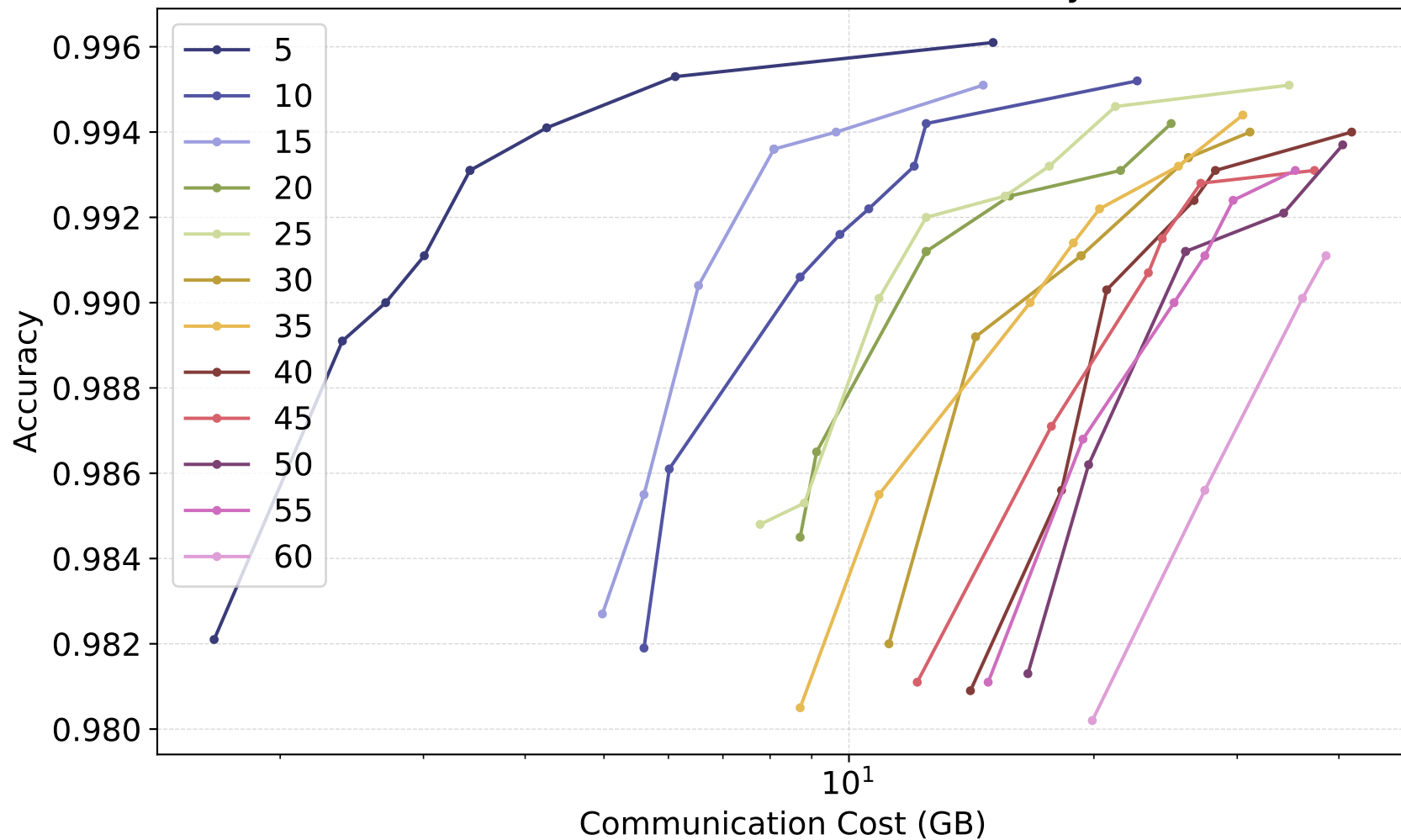


gm

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

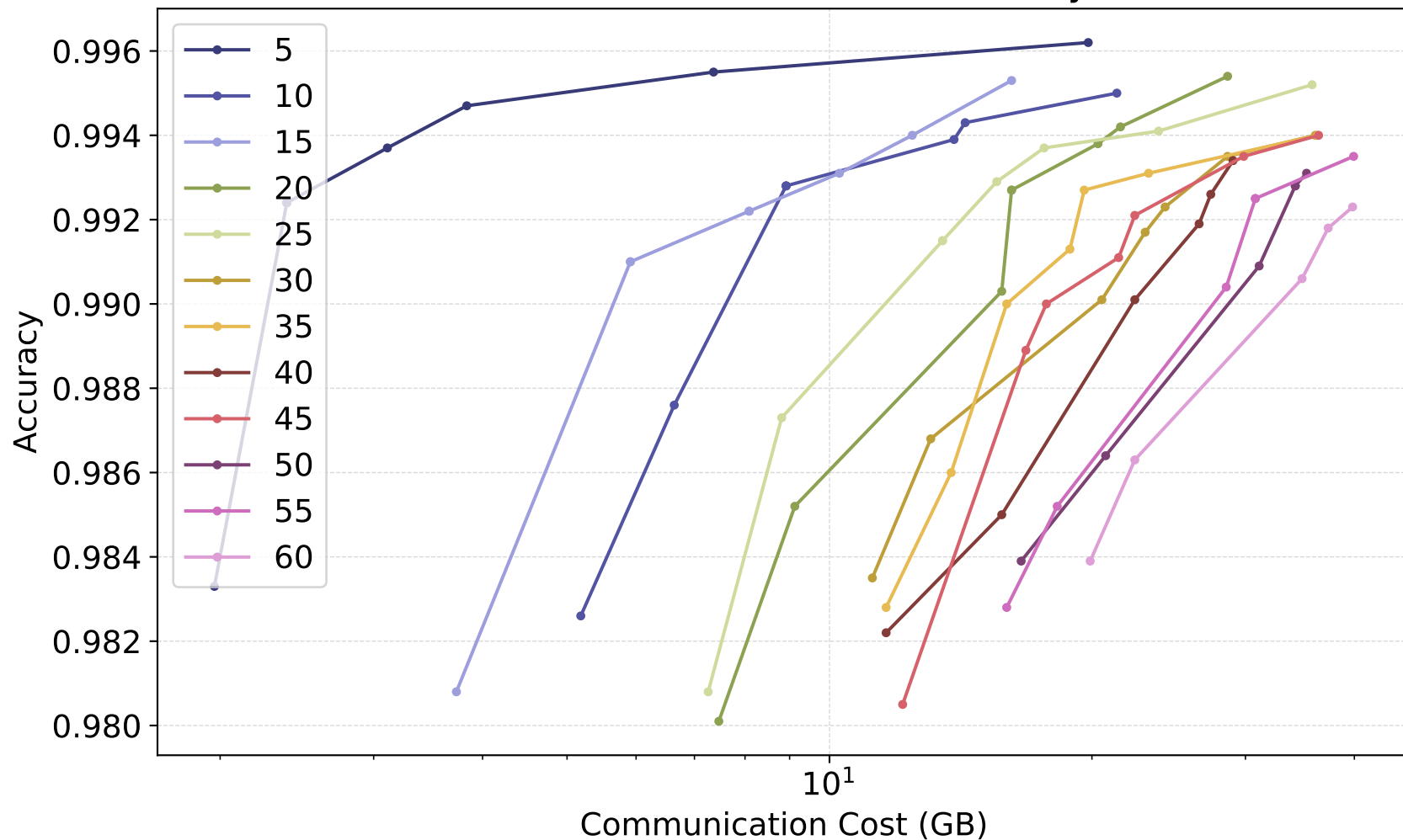


*Theta* : 75.0 , Batch Size: 32 , Bias: only label 8



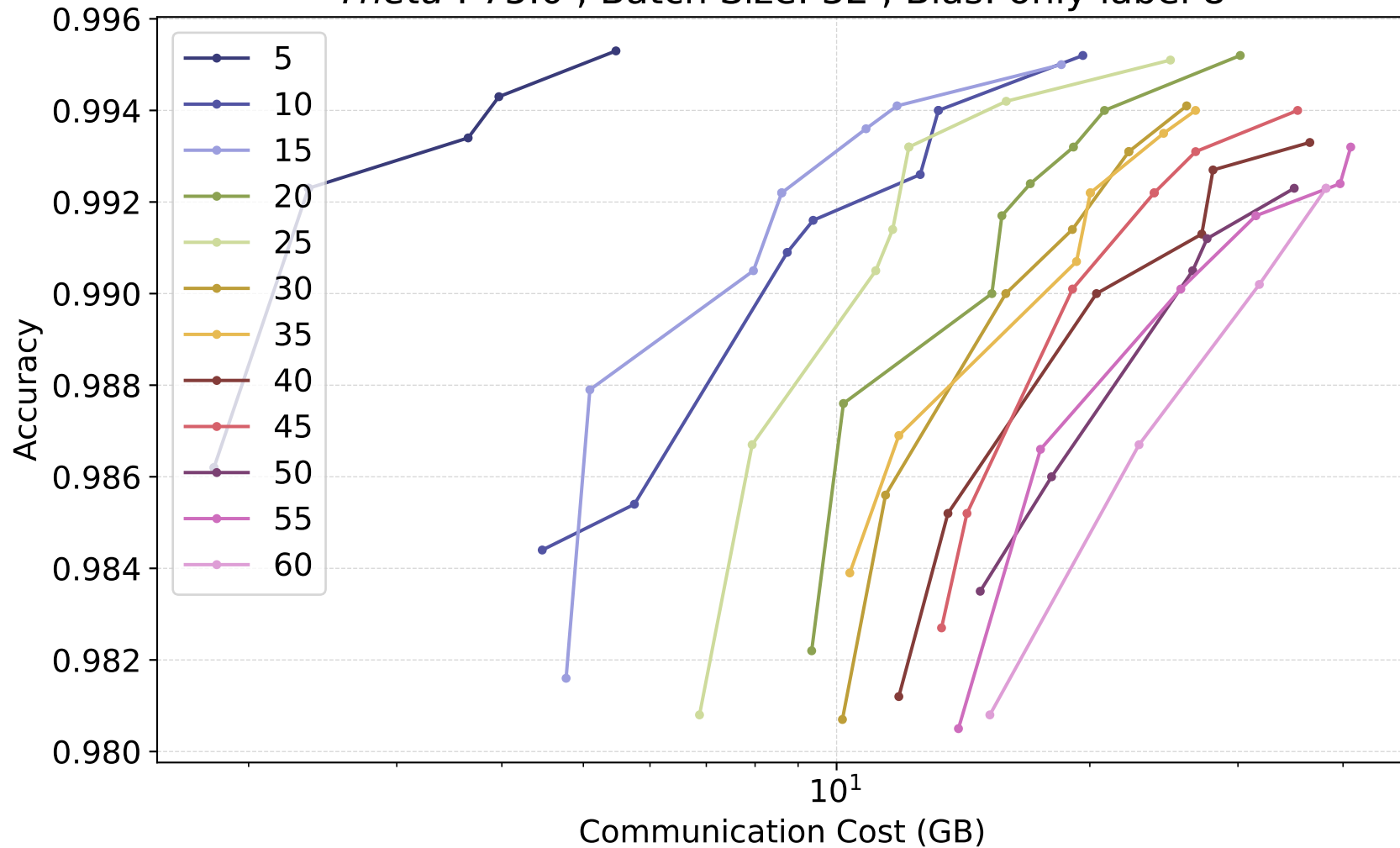
linear

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



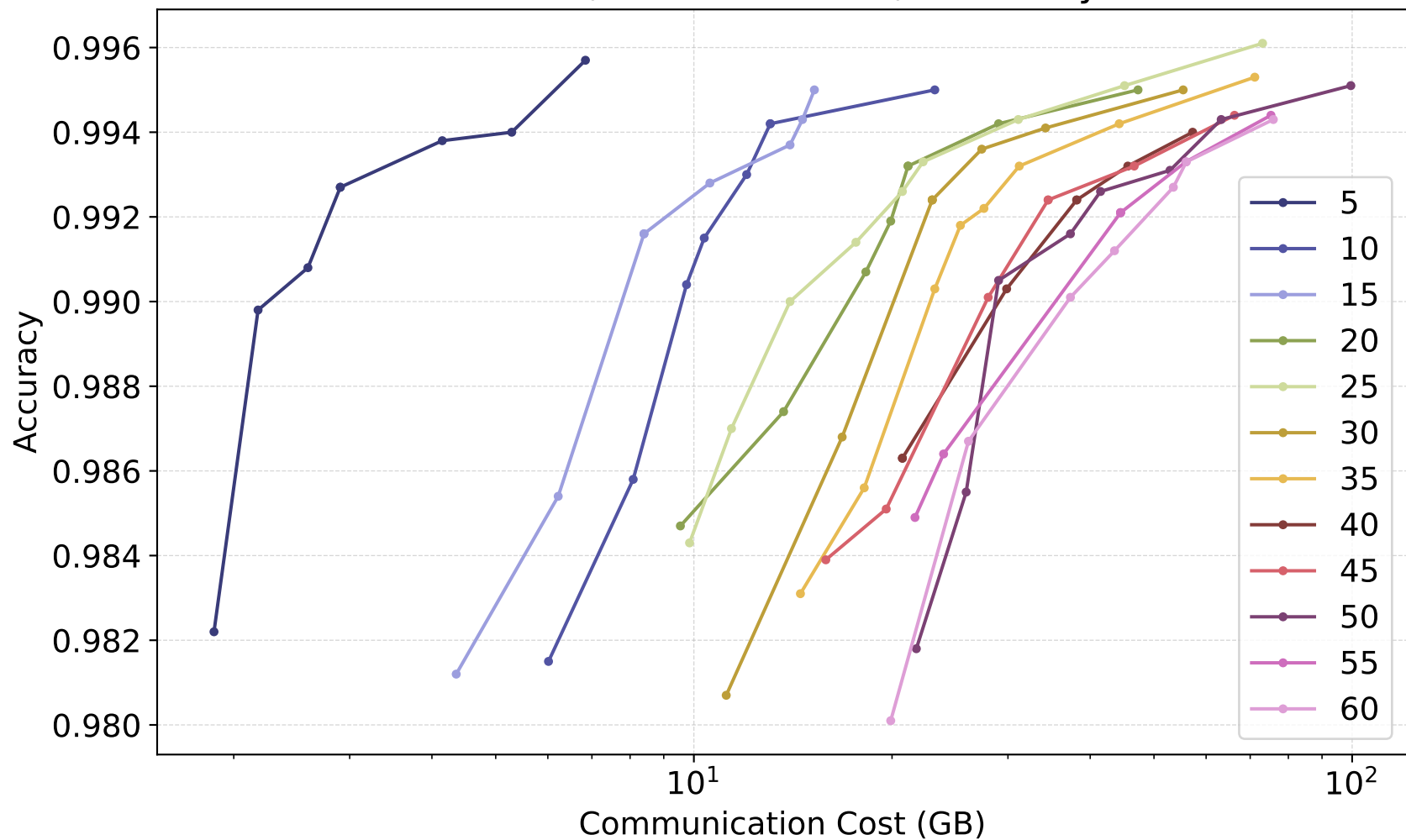
sketch

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

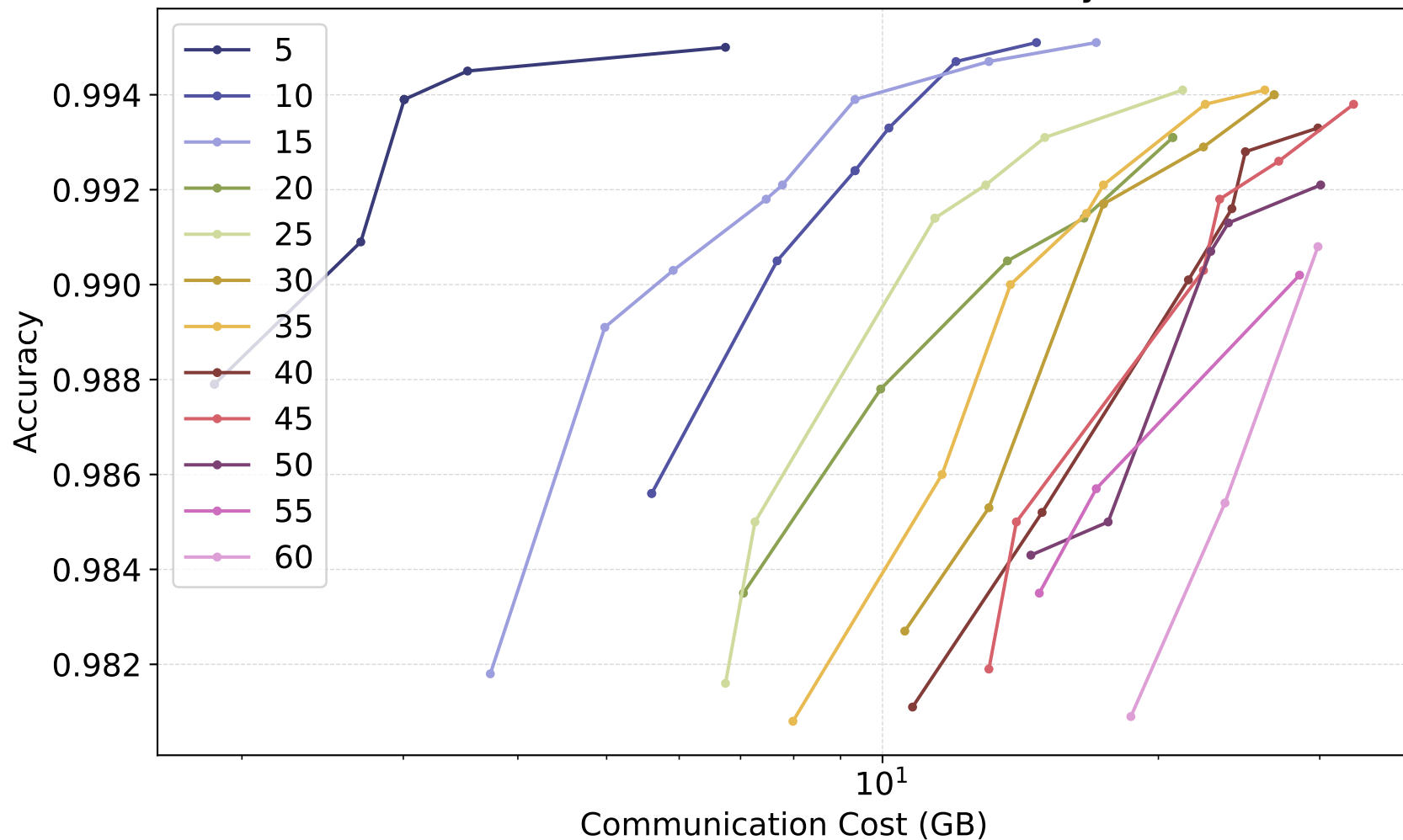


gm

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

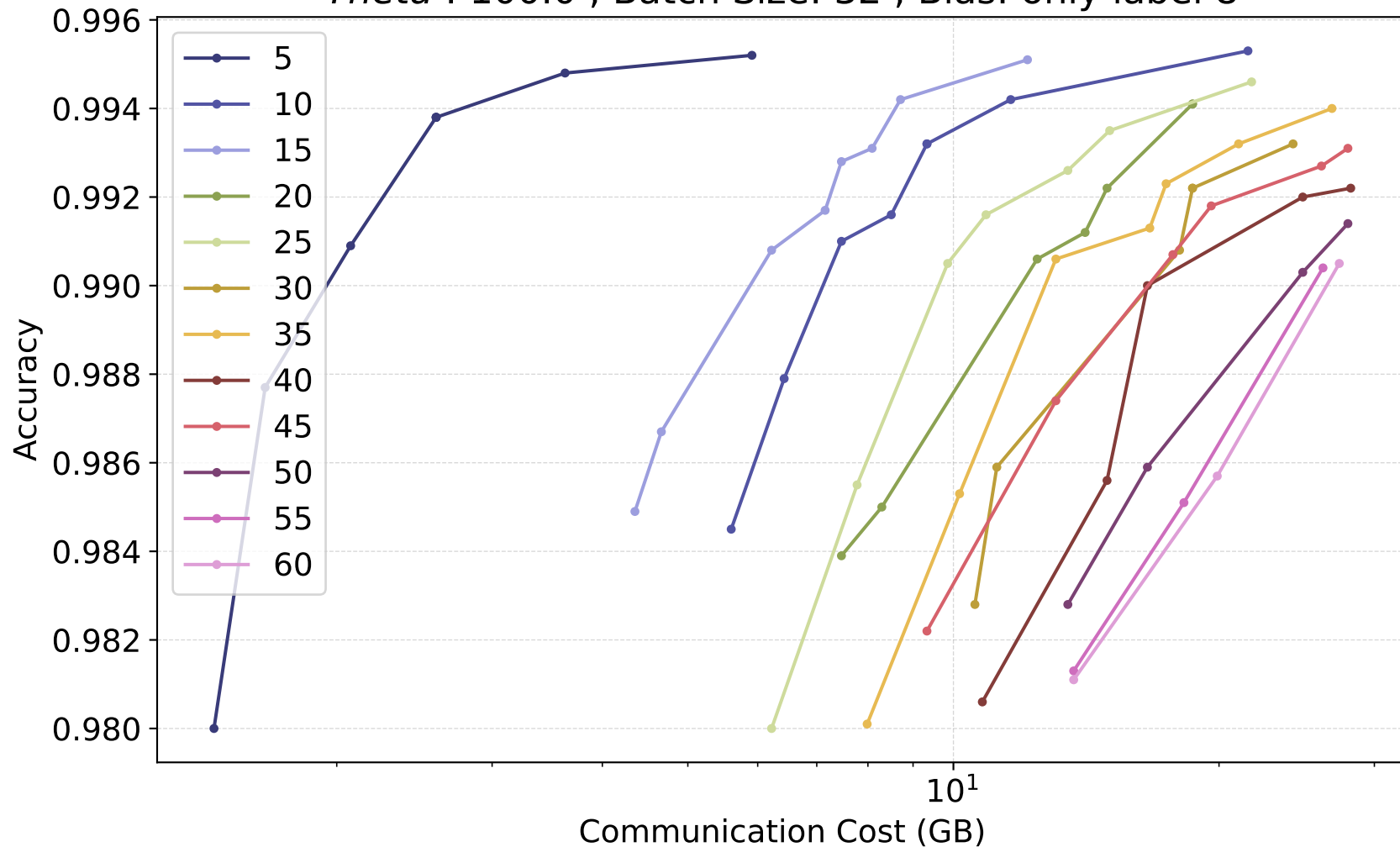


*Theta* : 100.0 , Batch Size: 32 , Bias: only label 8



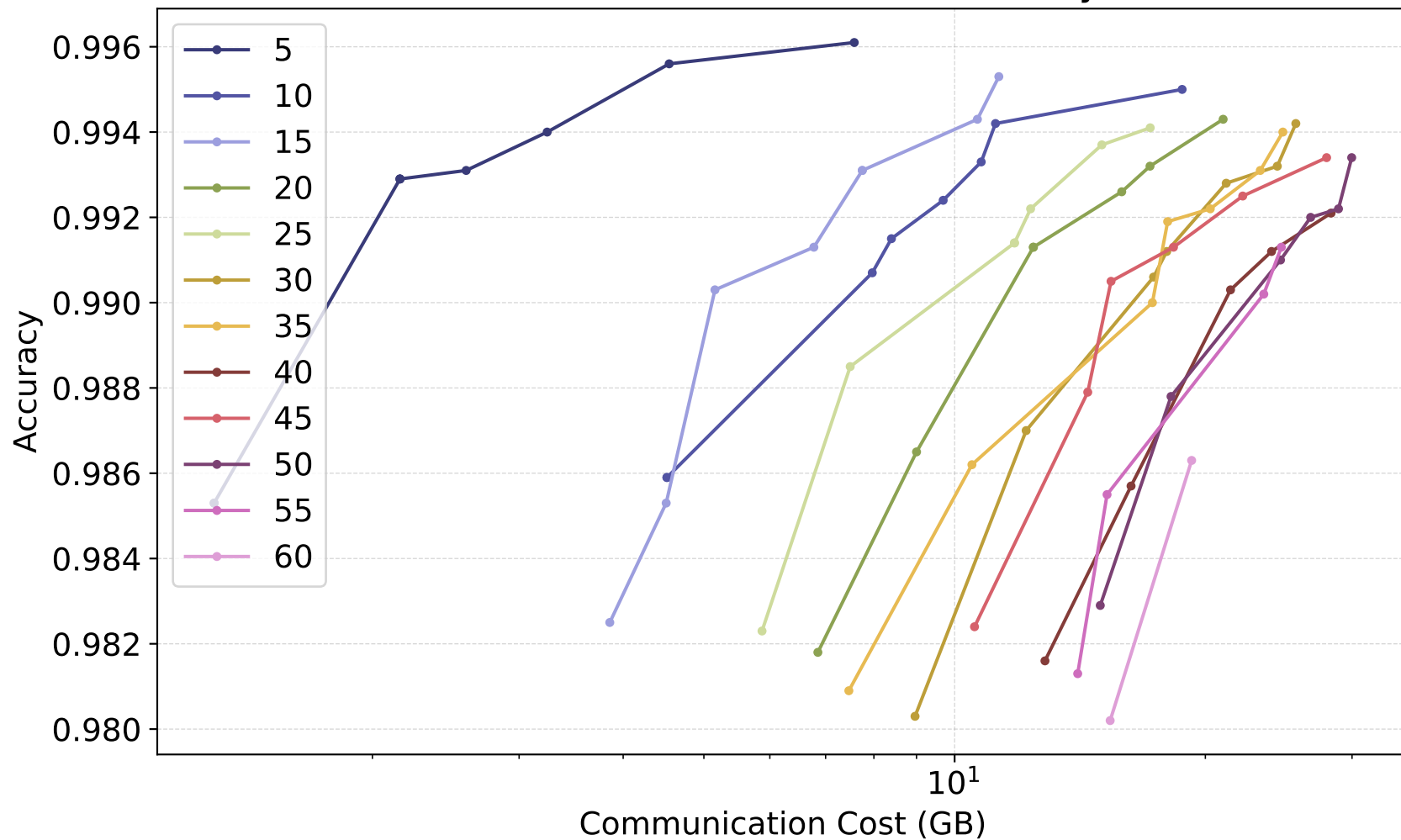
linear

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

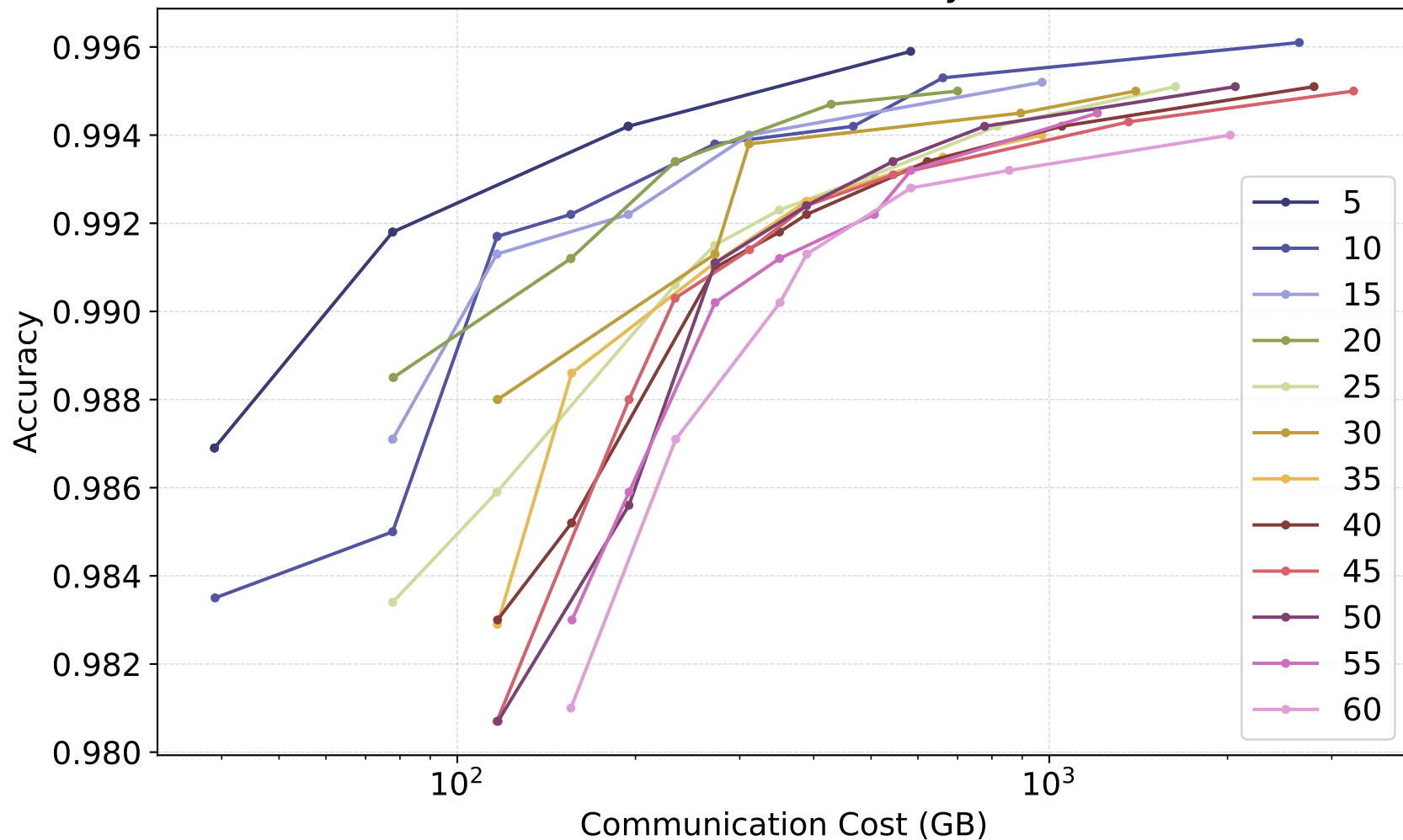




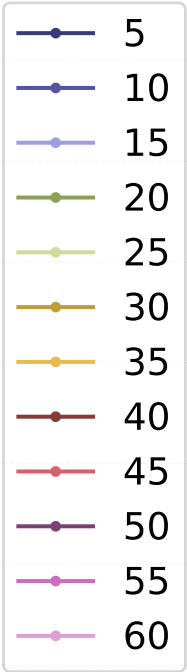
*Theta* : 100.0 , Batch Size: 32 , Bias: only label 8



synchronous  
Batch Size : 32 , Bias: only label 0



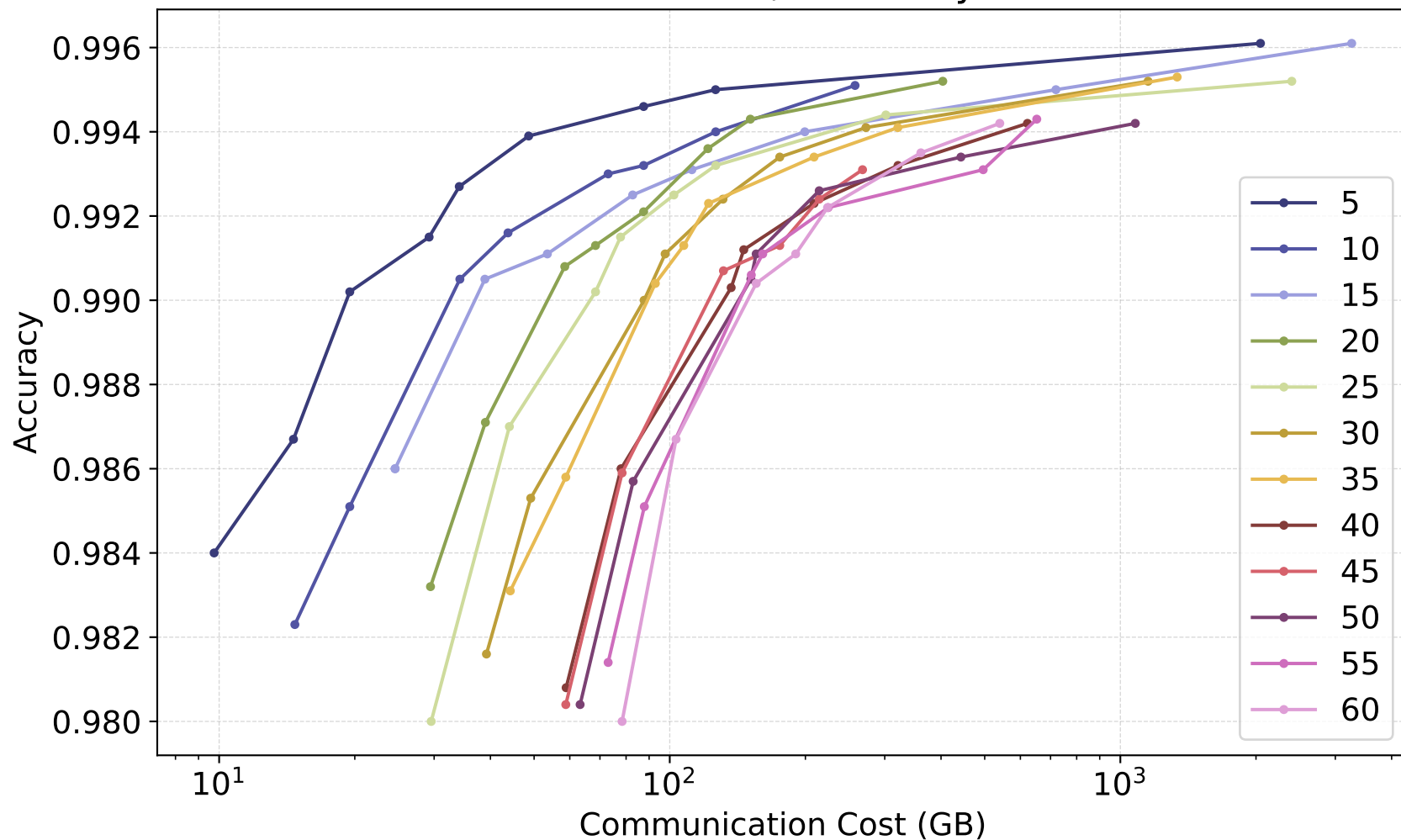
Batch Size : 64 , Bias: only label 0



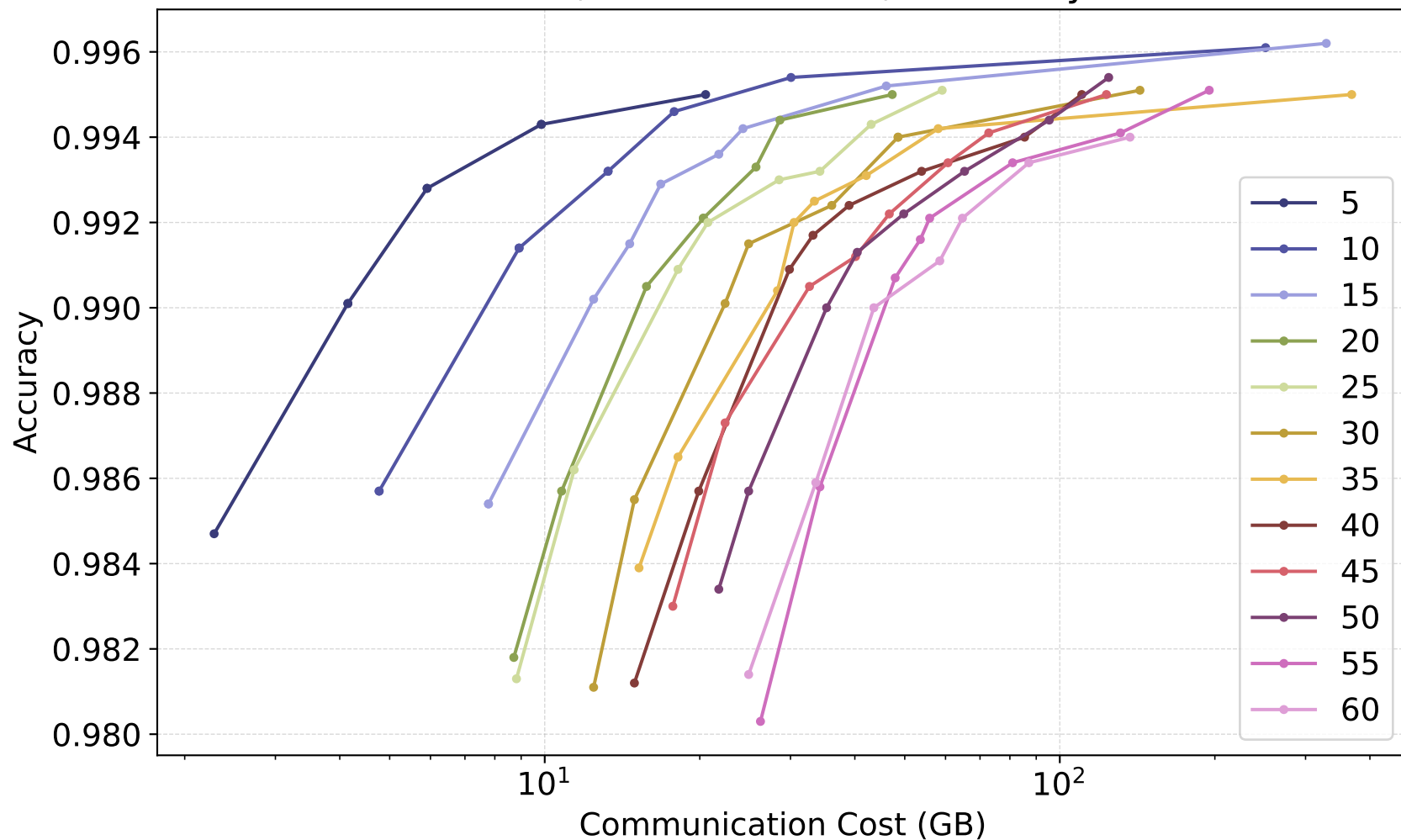
Batch Size : 128 , Bias: only label 0



synchronous  
Batch Size : 256 , Bias: only label 0

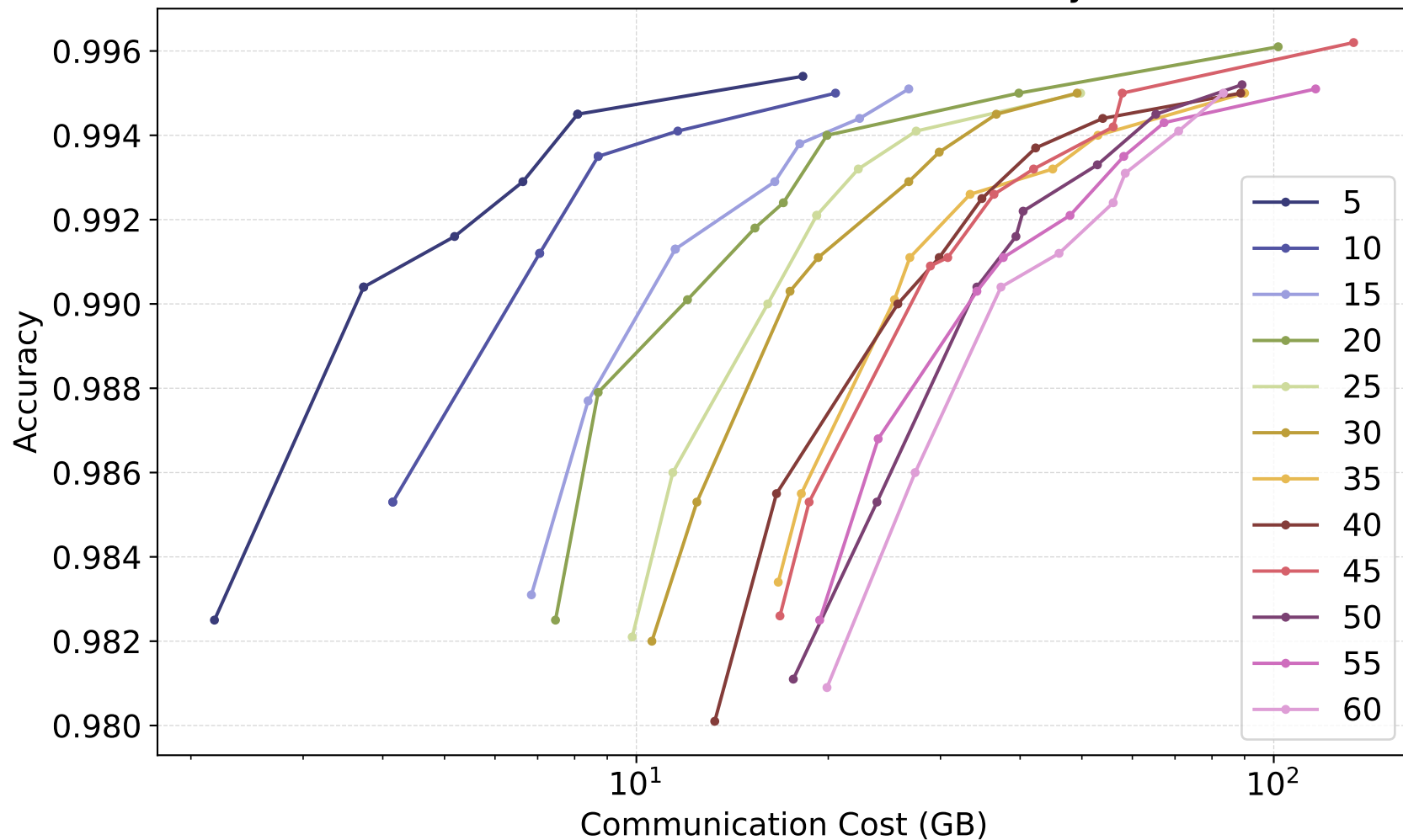


*Theta* : 15.0 , Batch Size: 32 , Bias: only label 0

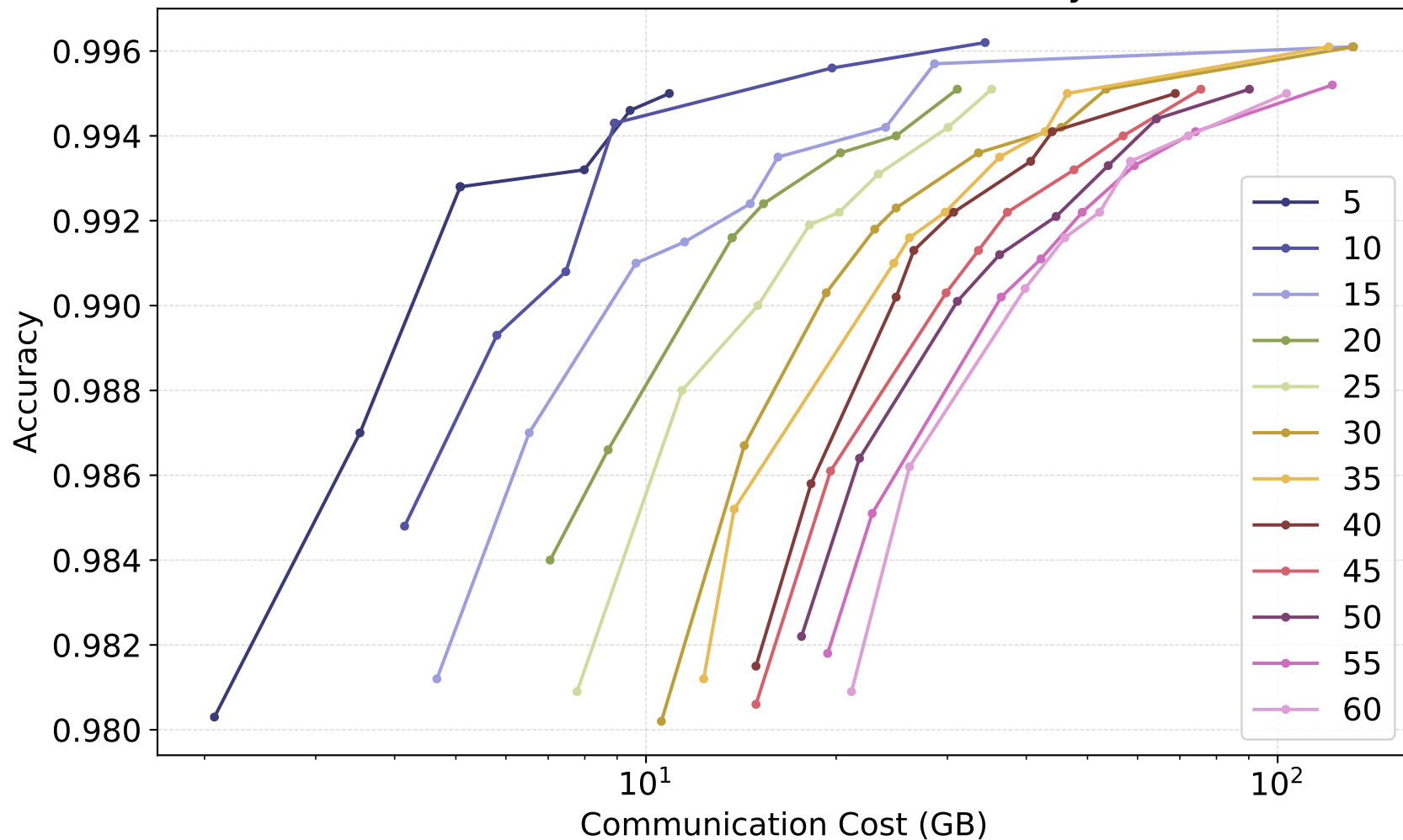


naive

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

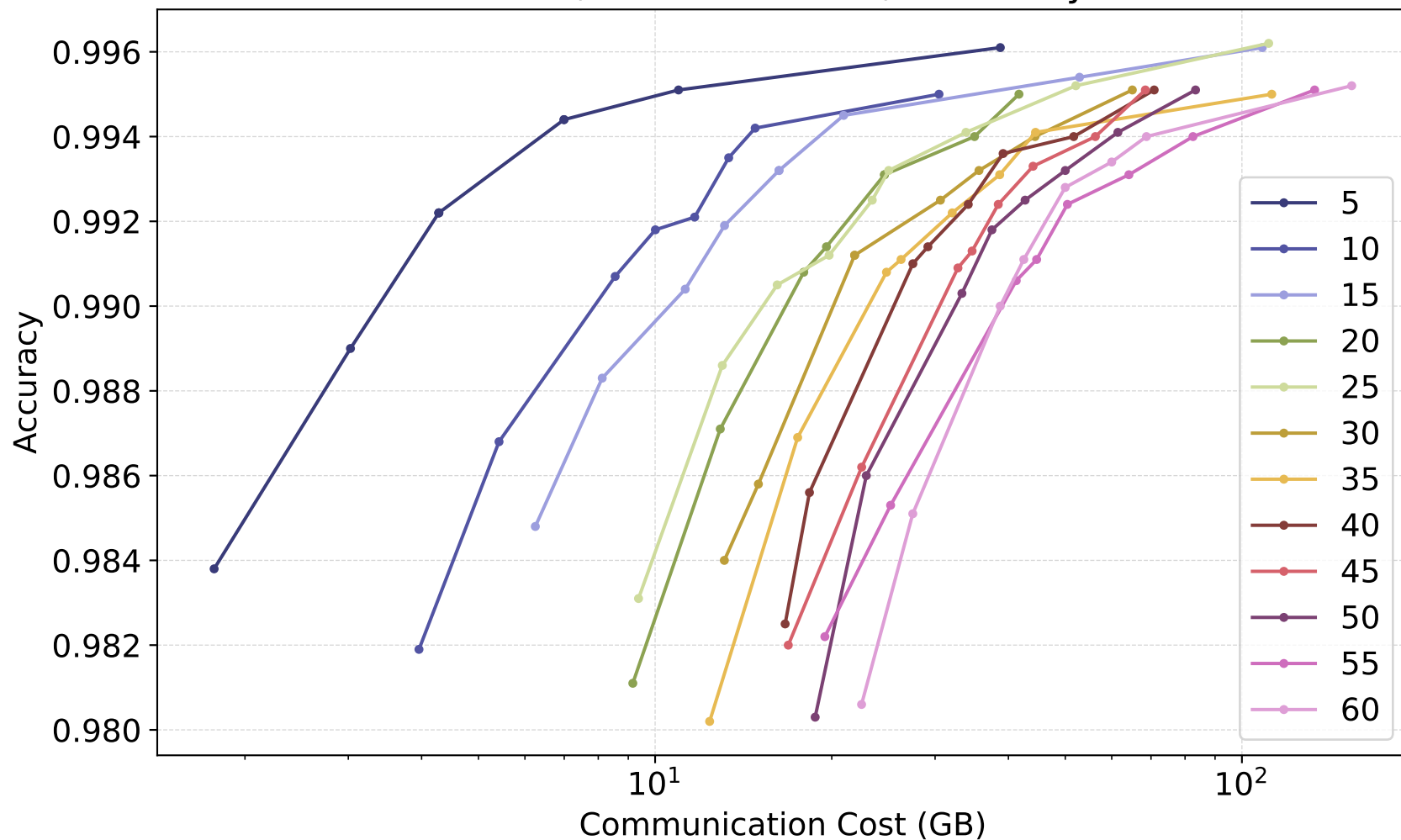


*Theta* : 15.0 , Batch Size: 32 , Bias: only label 0



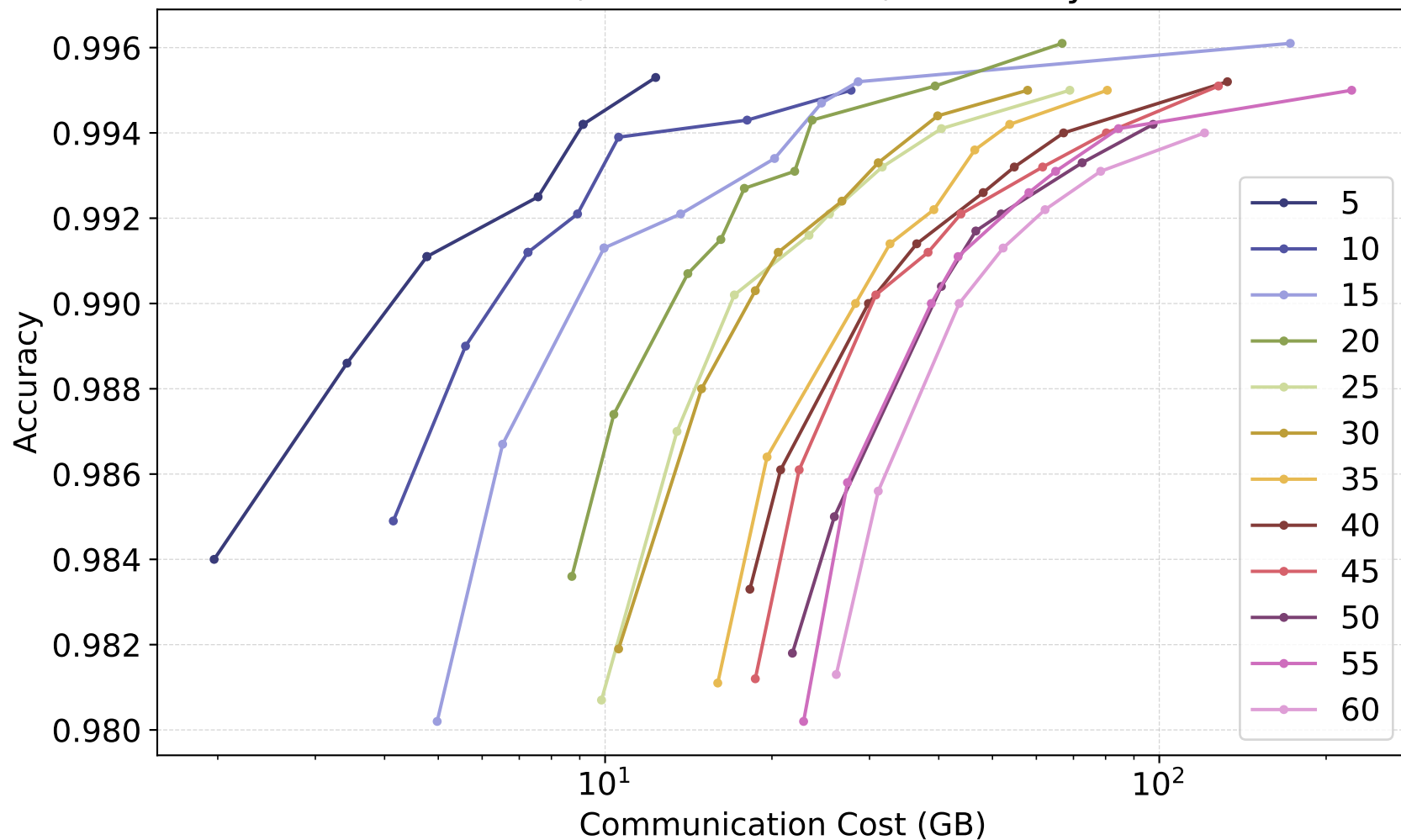


*Theta* : 15.0 , Batch Size: 32 , Bias: only label 0

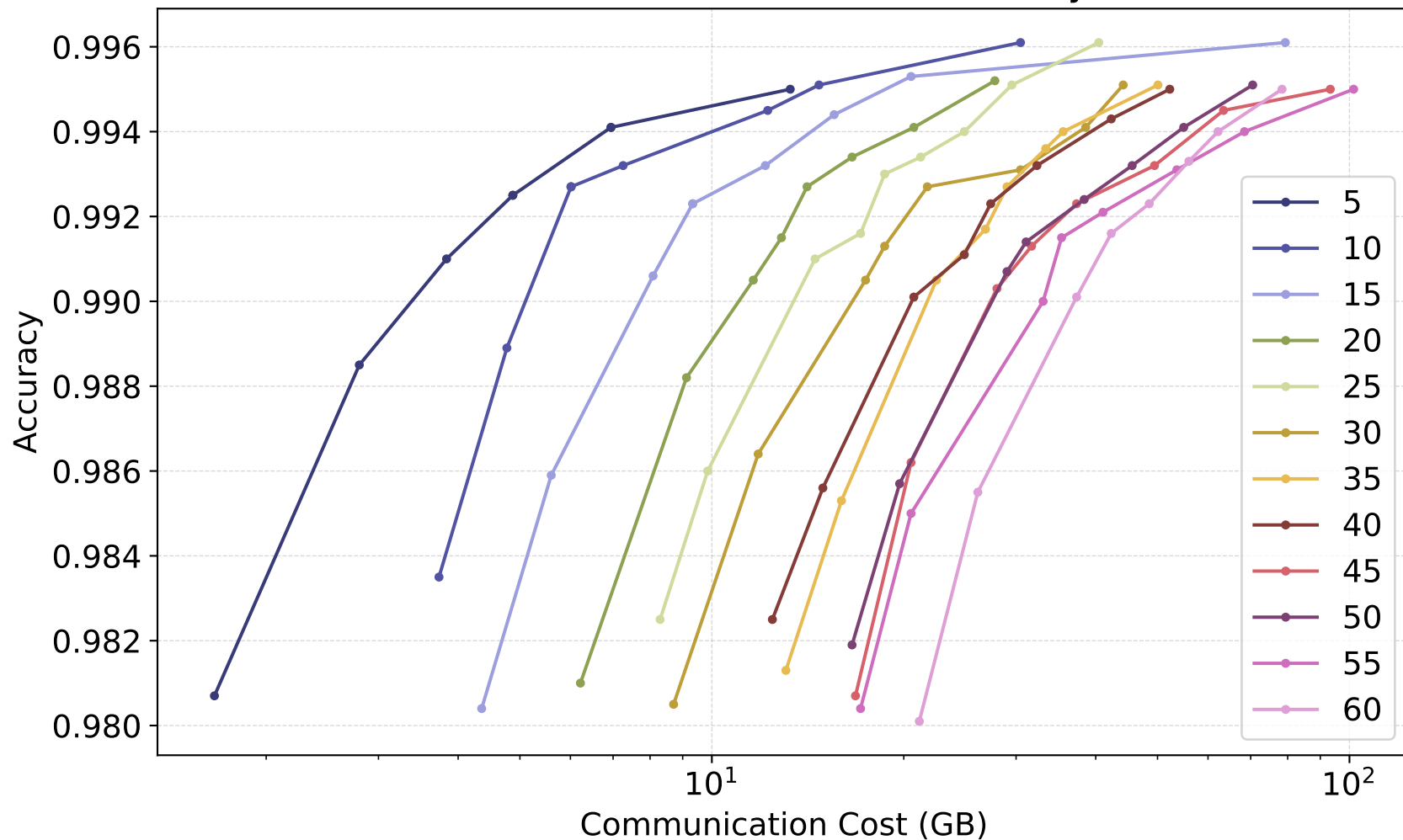


gm

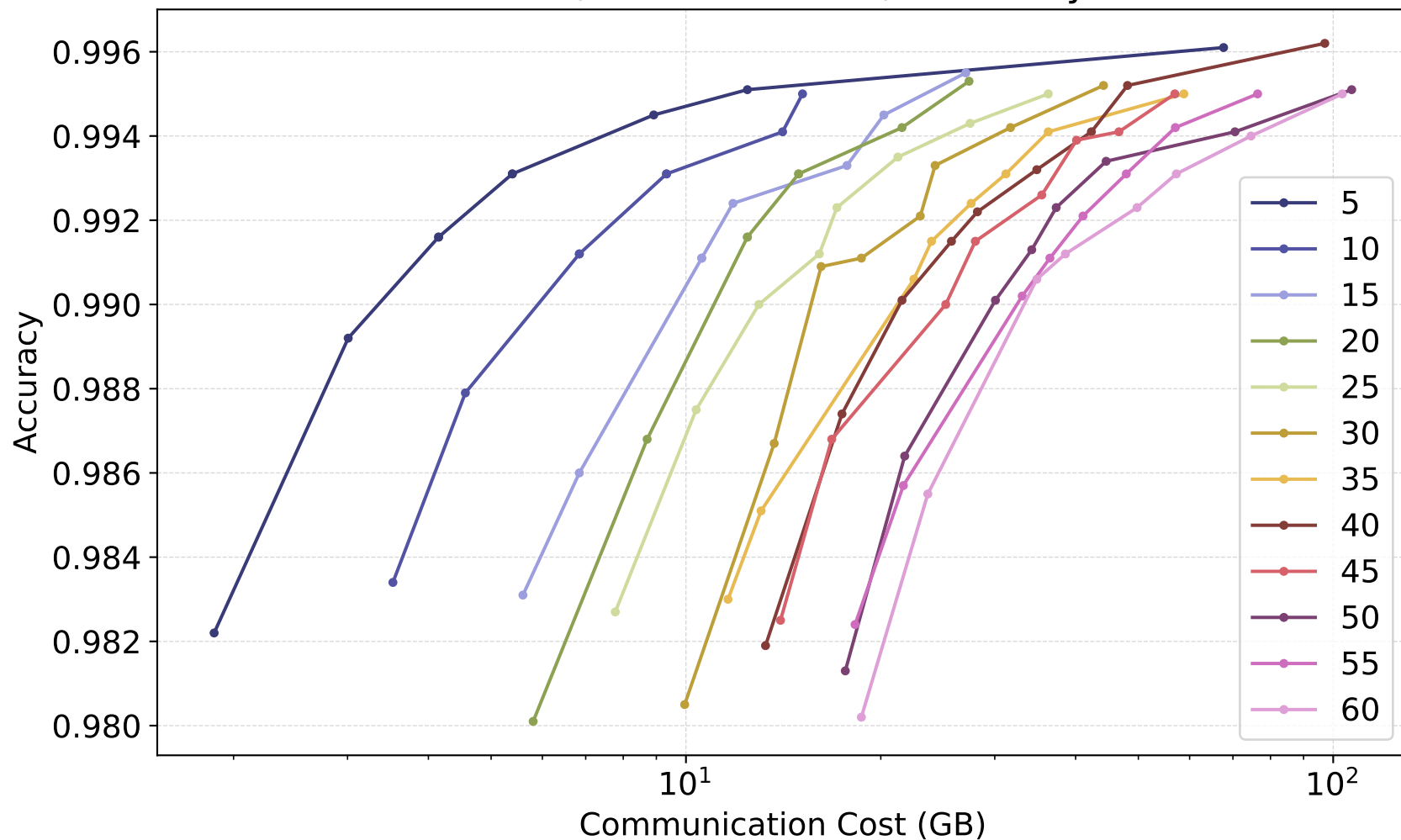
*Theta* : 20.0 , Batch Size: 32 , Bias: only label 0



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

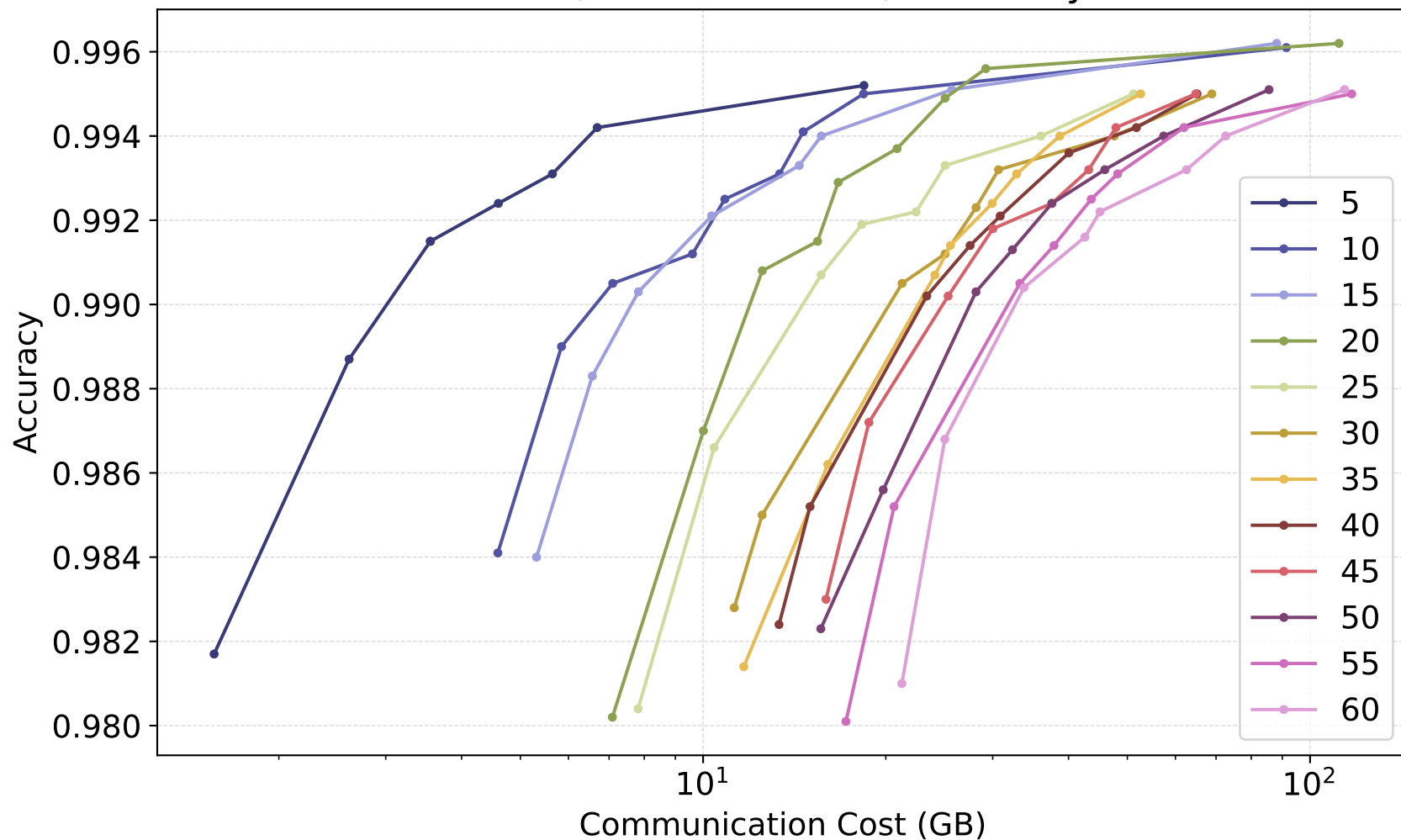


*Theta* : 20.0 , Batch Size: 32 , Bias: only label 0



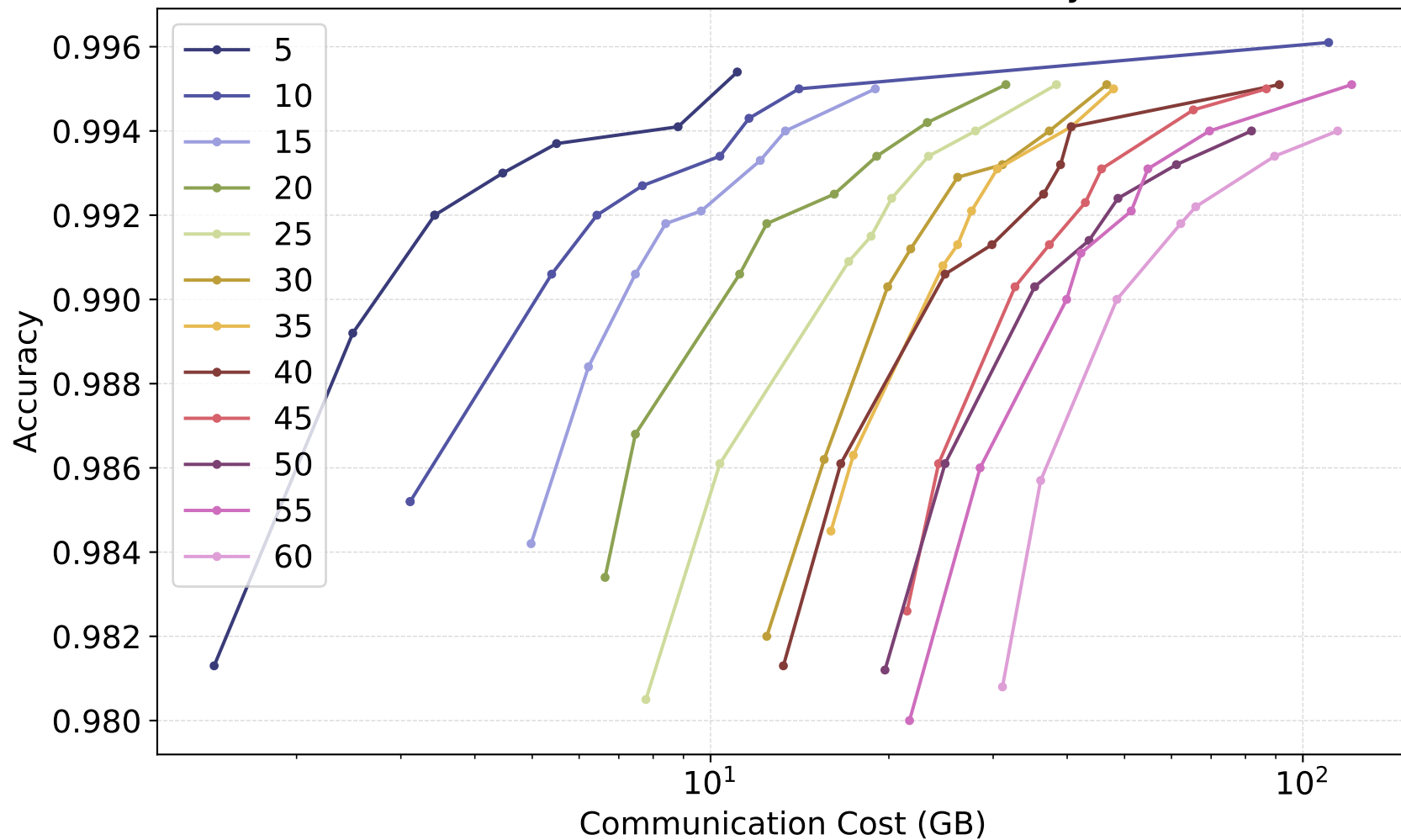
sketch

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

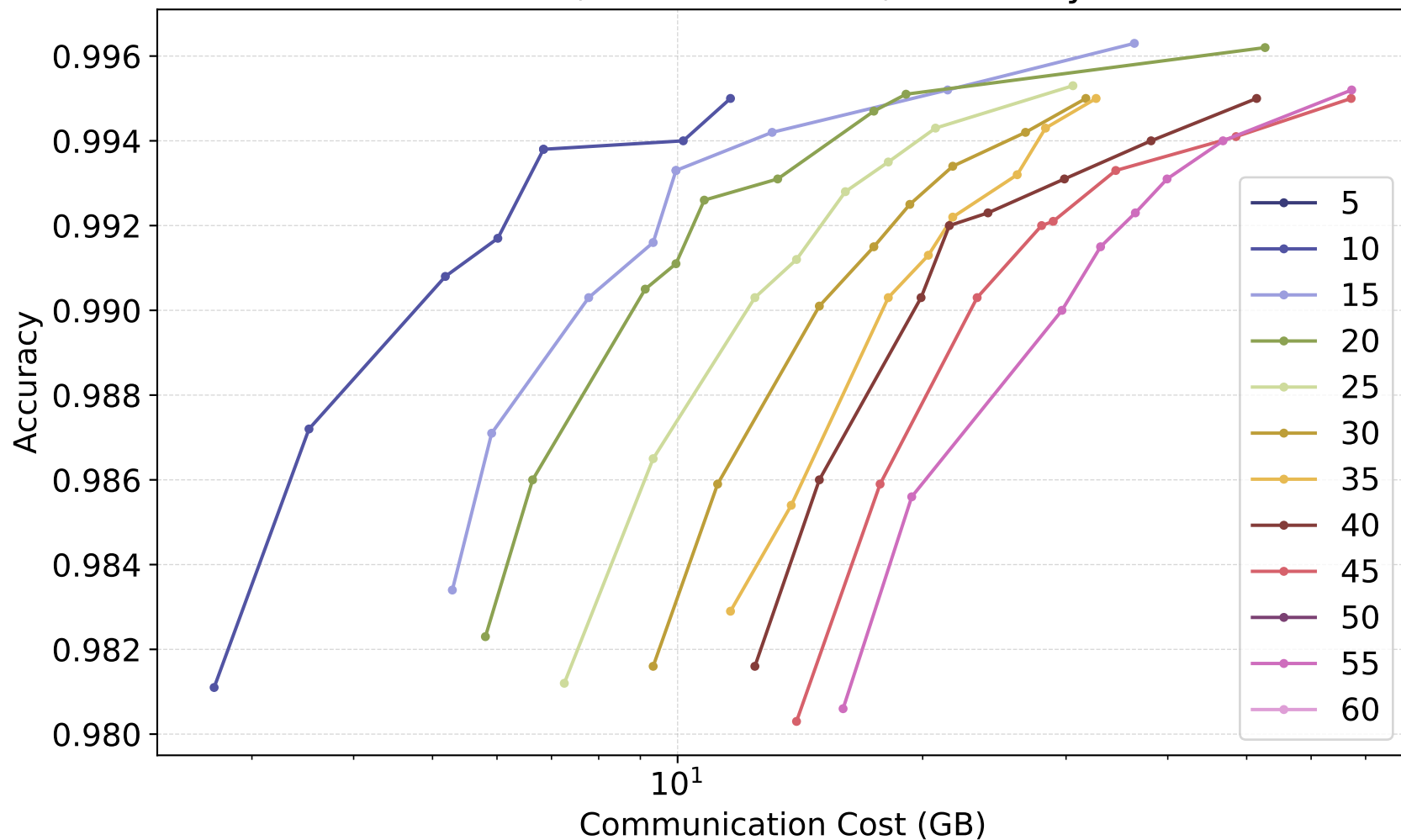


gm

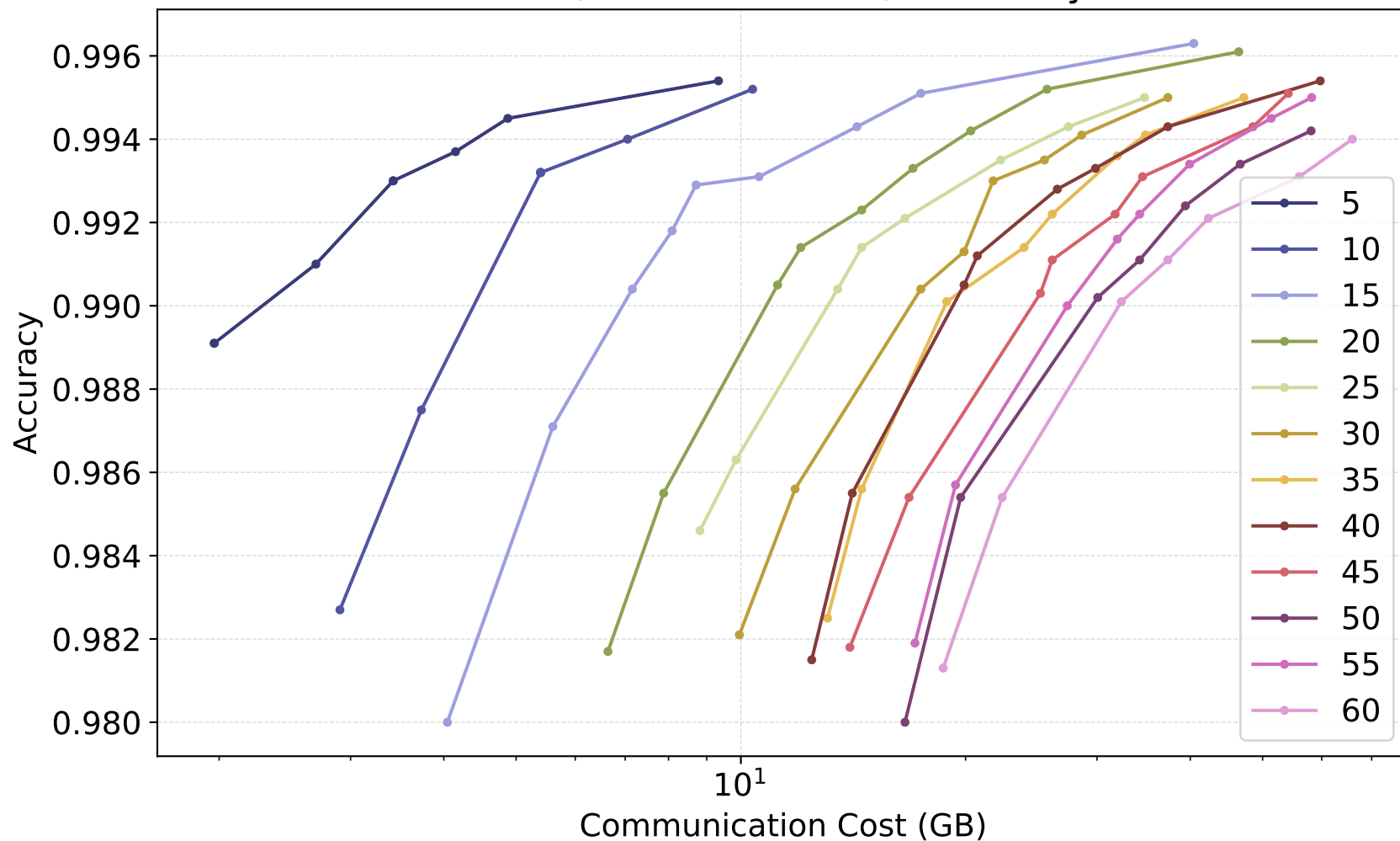
*Theta* : 30.0 , Batch Size: 32 , Bias: only label 0



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

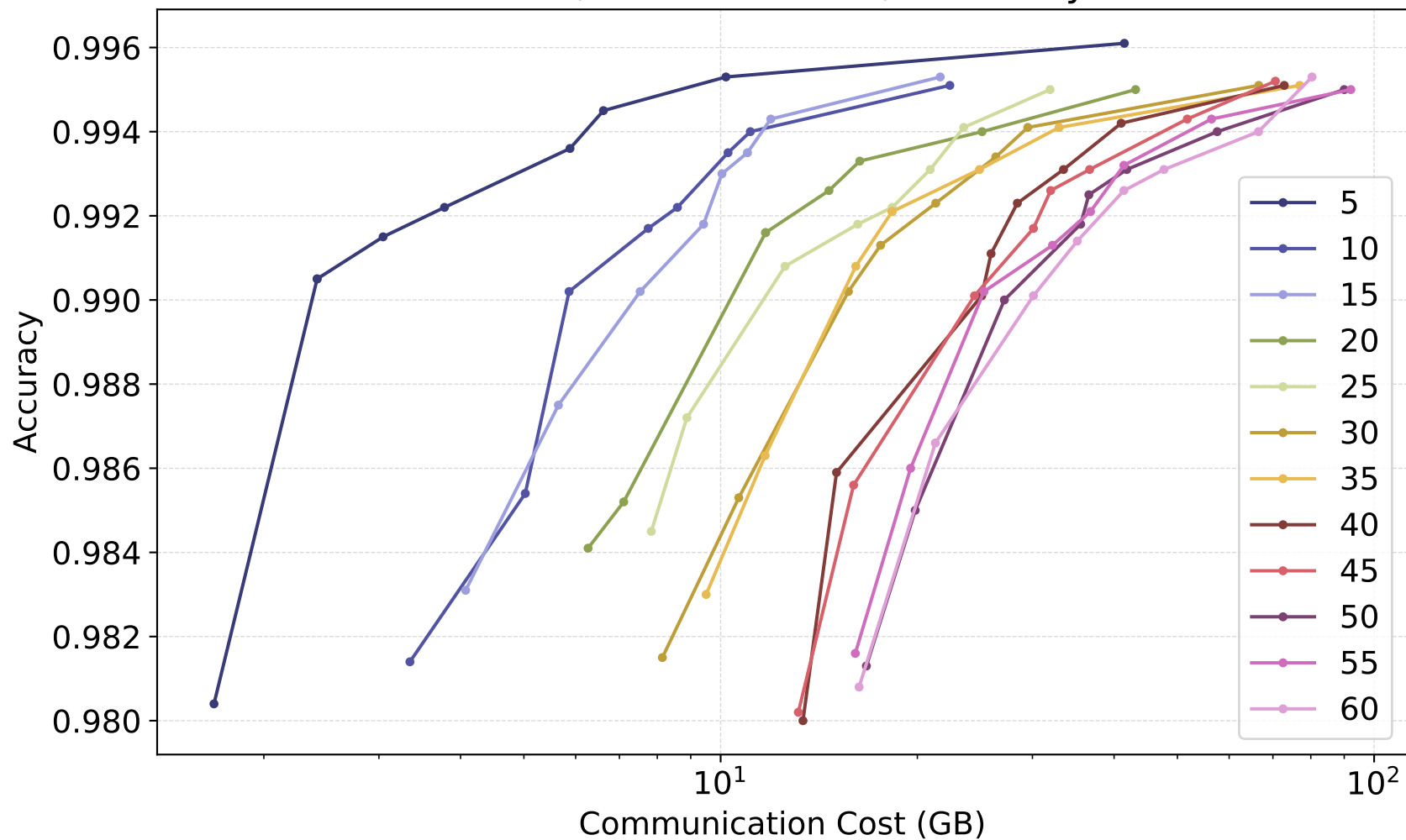


*Theta* : 30.0 , Batch Size: 32 , Bias: only label 0



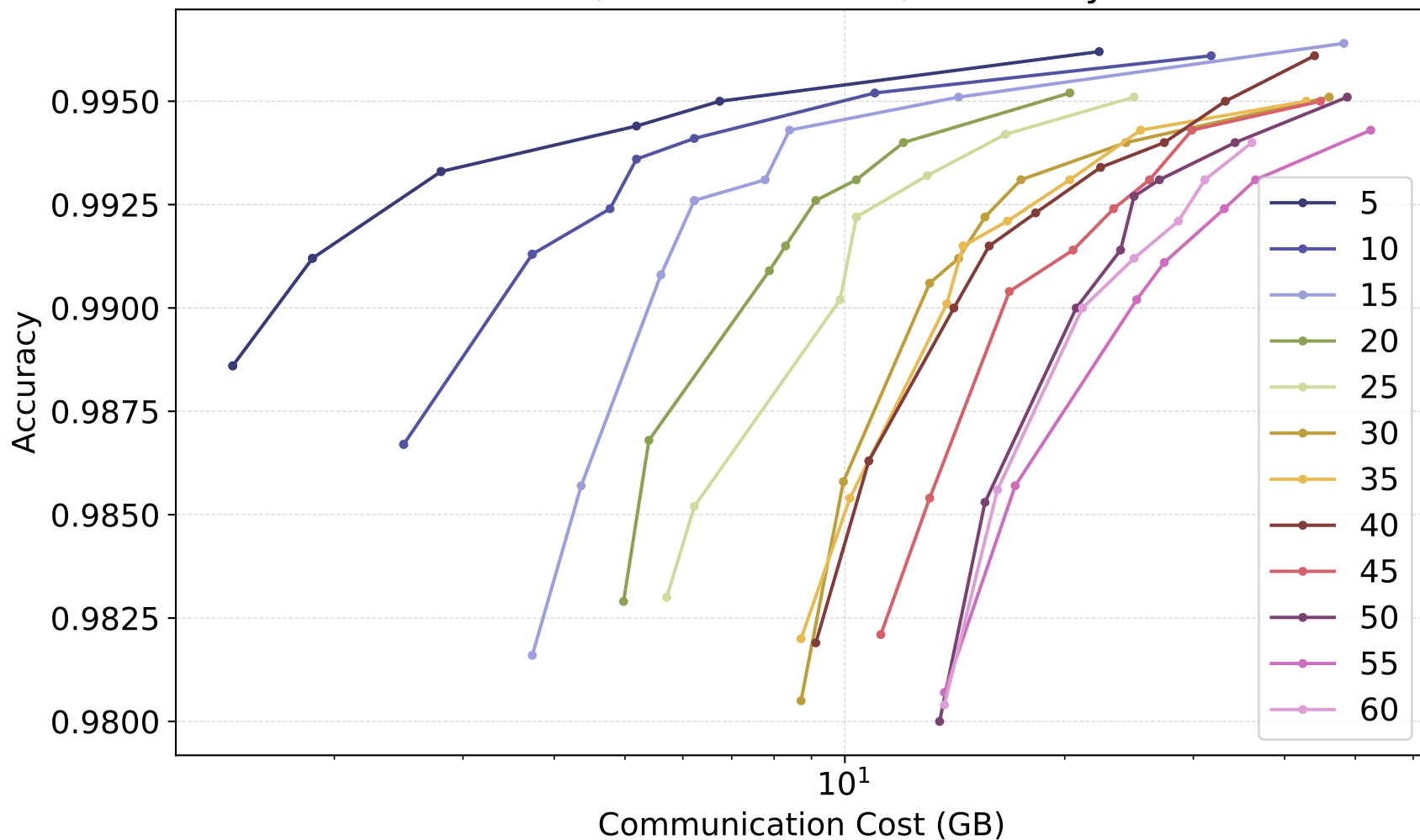


*Theta* : 30.0 , Batch Size: 32 , Bias: only label 0

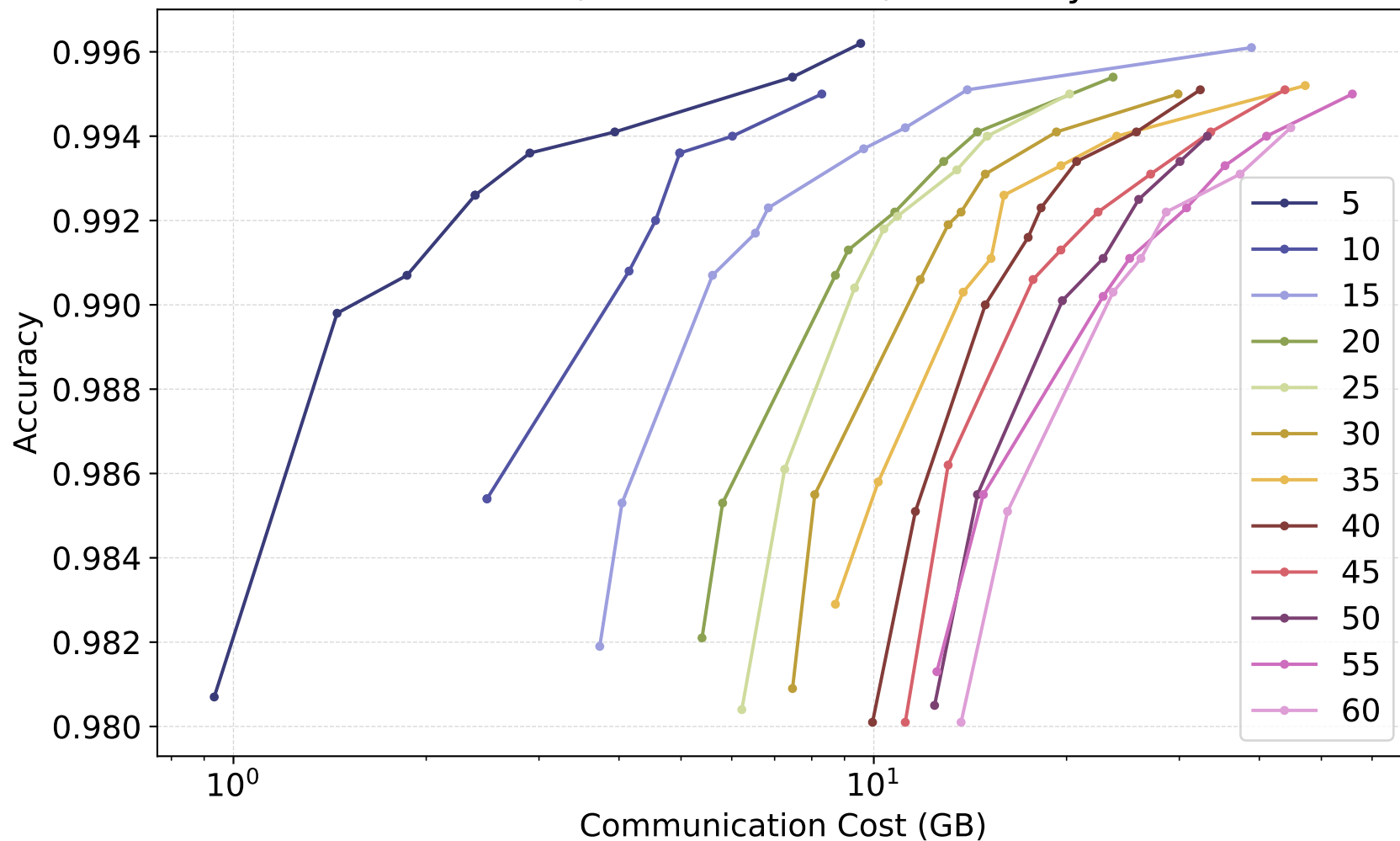




*Theta* : 50.0 , Batch Size: 32 , Bias: only label 0

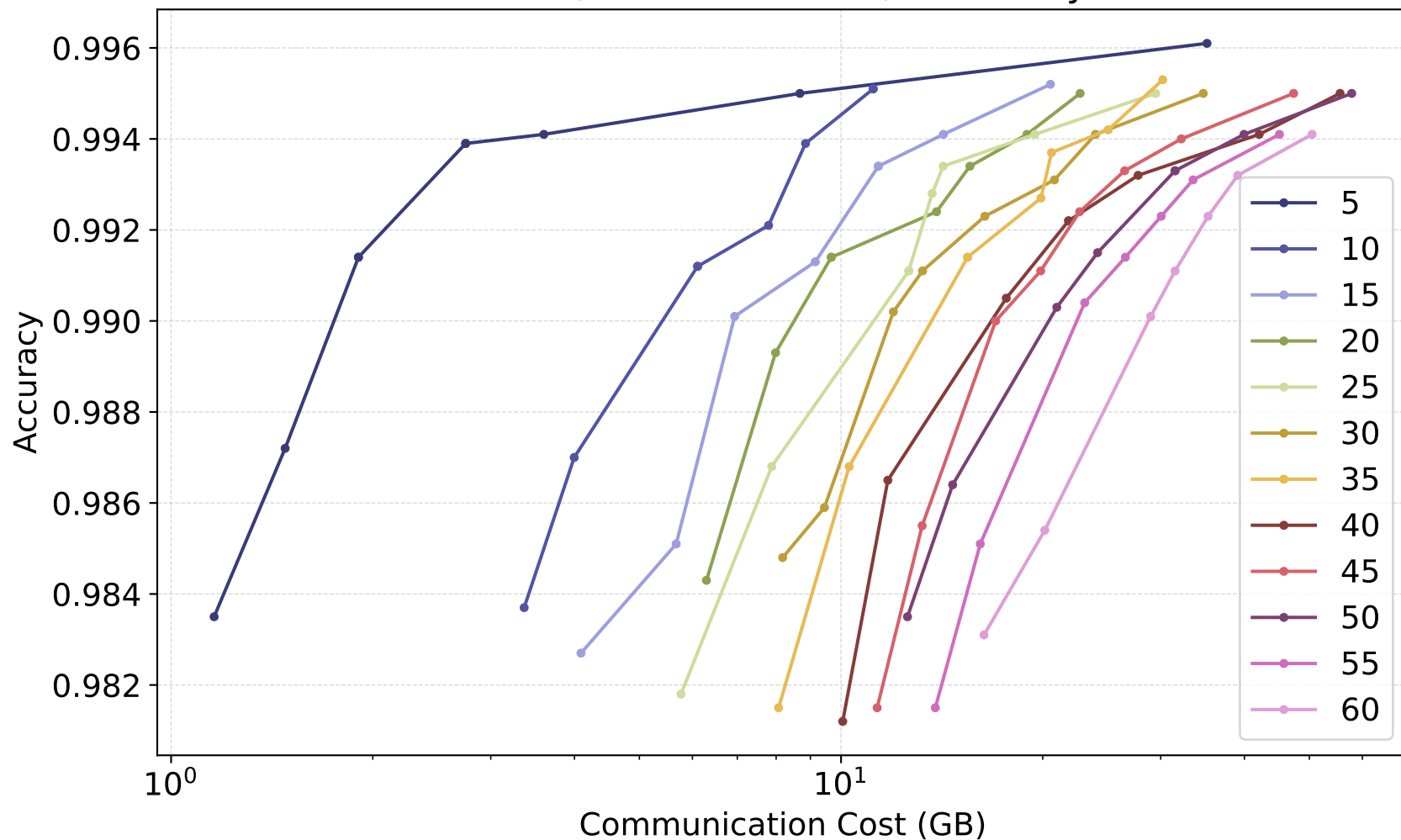


*Theta* : 50.0 , Batch Size: 32 , Bias: only label 0



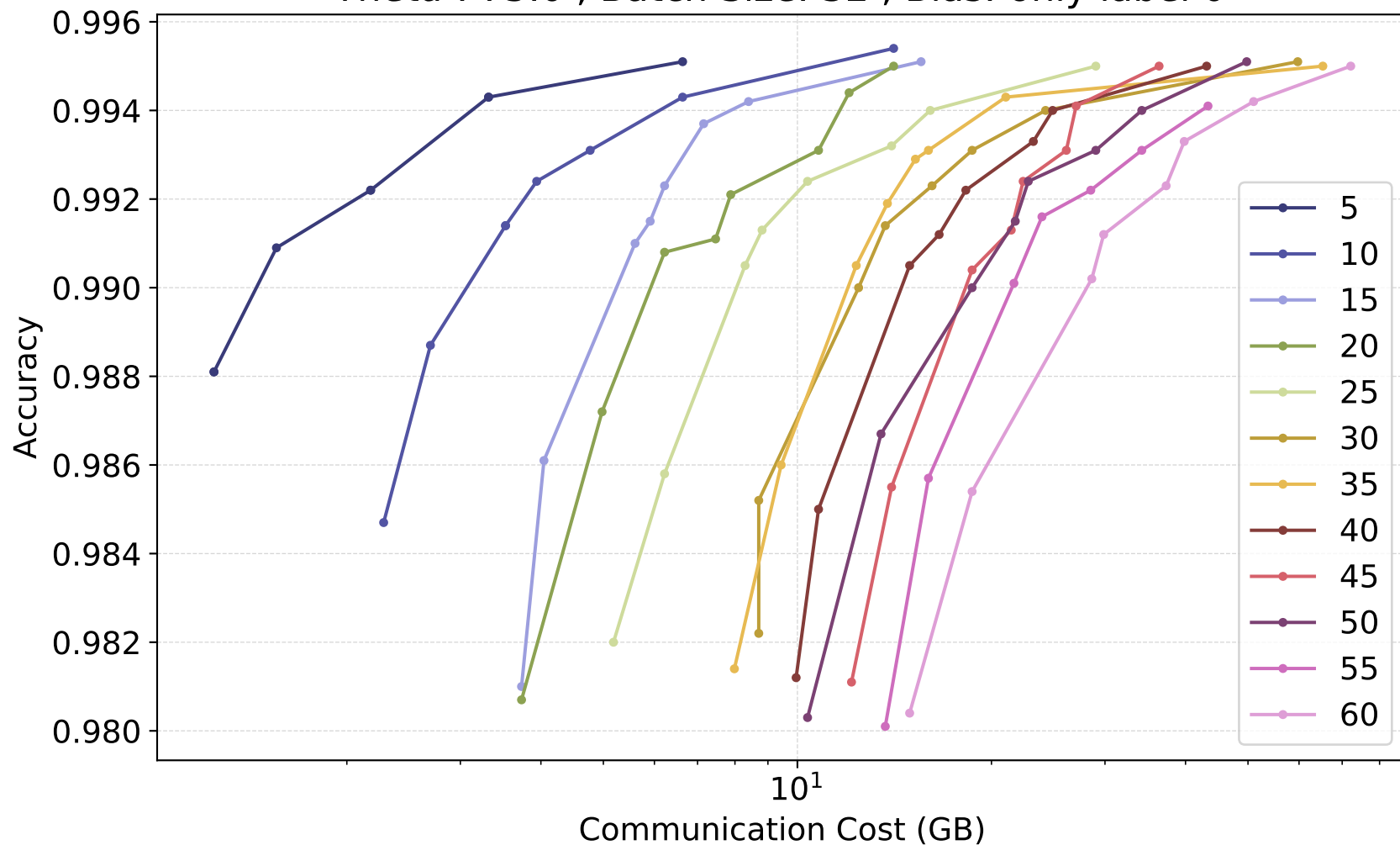
sketch

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

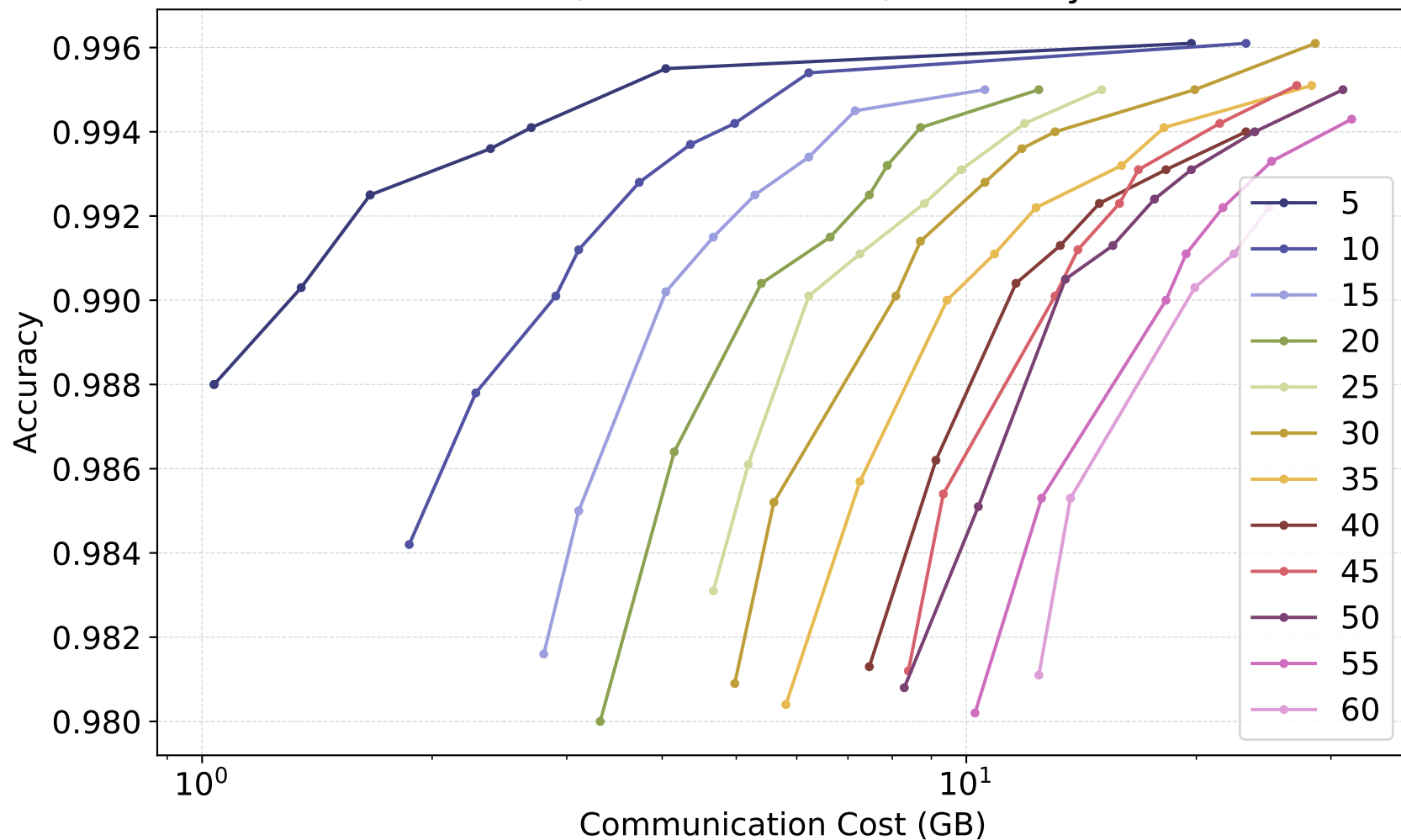


gm

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

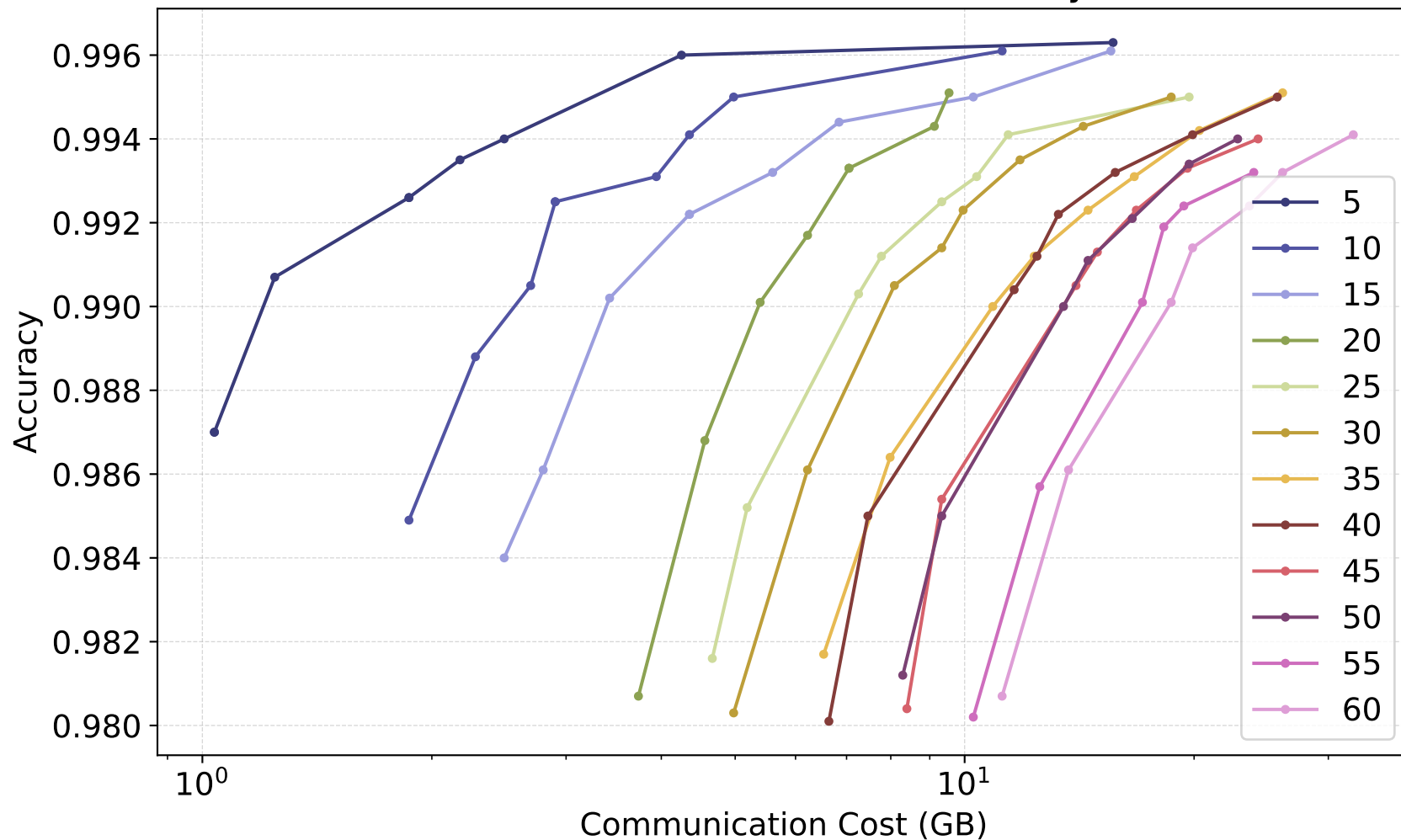


*Theta* : 75.0 , Batch Size: 32 , Bias: only label 0



linear

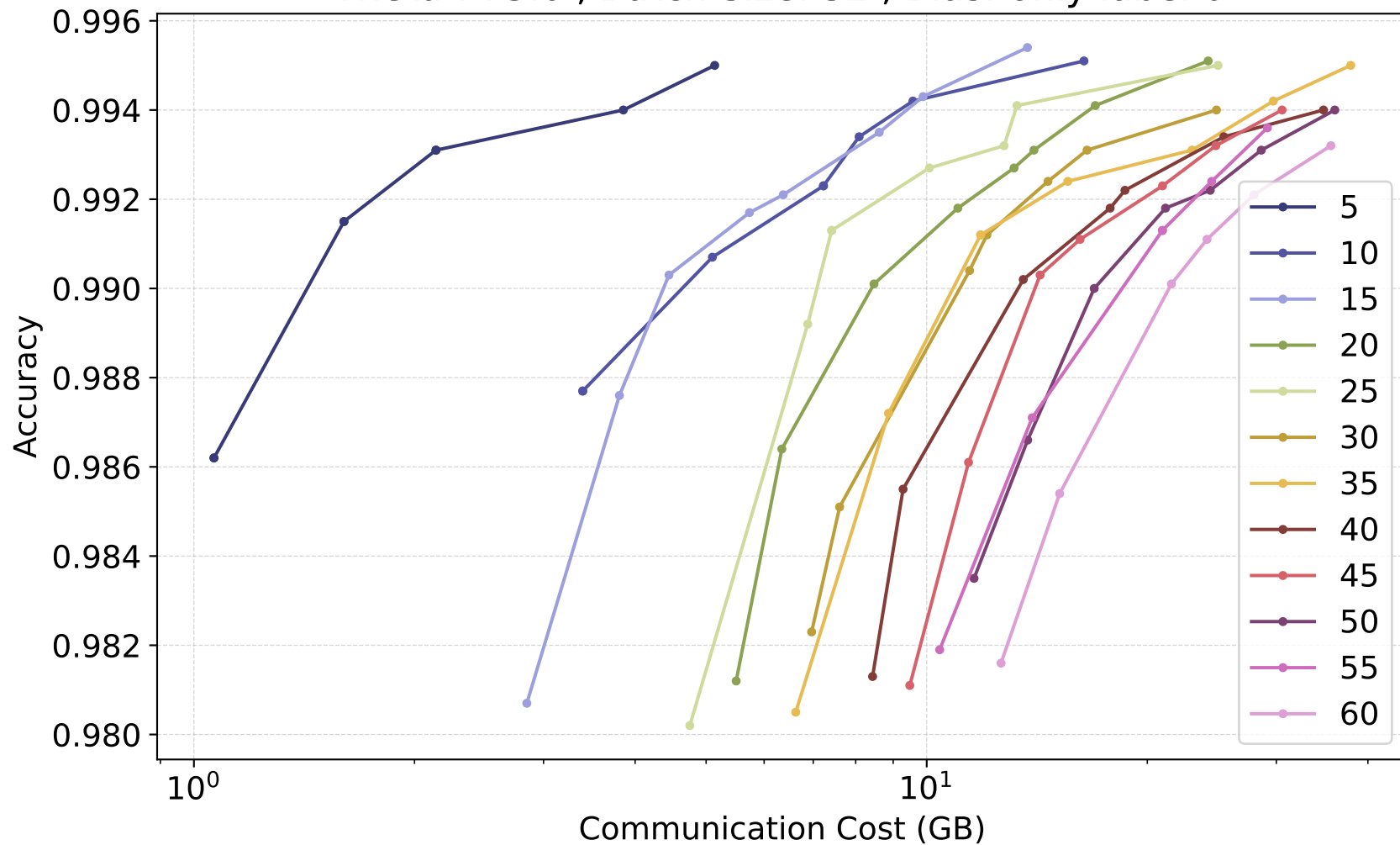
*Theta* : 75.0 , Batch Size: 32 , Bias: only label 0



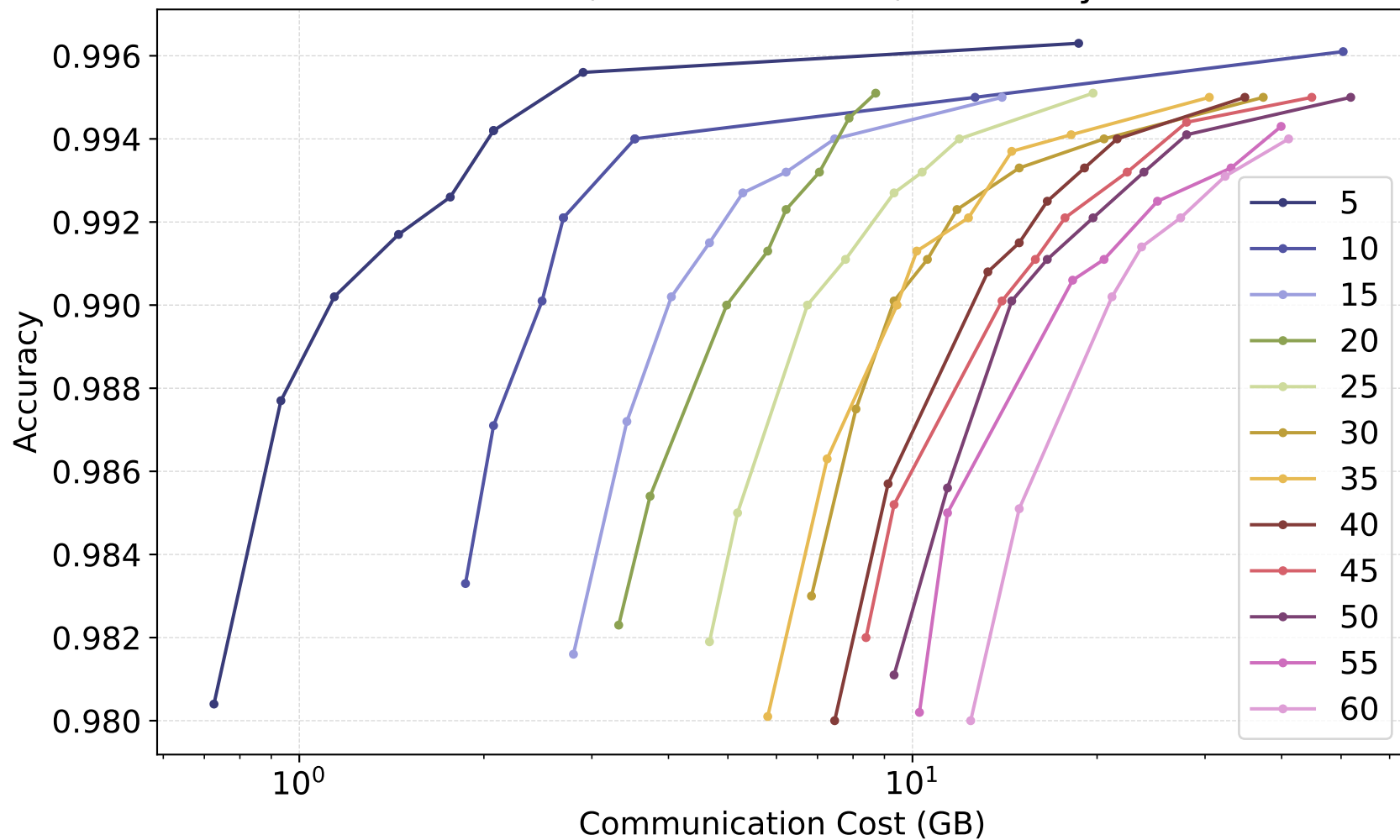


sketch

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

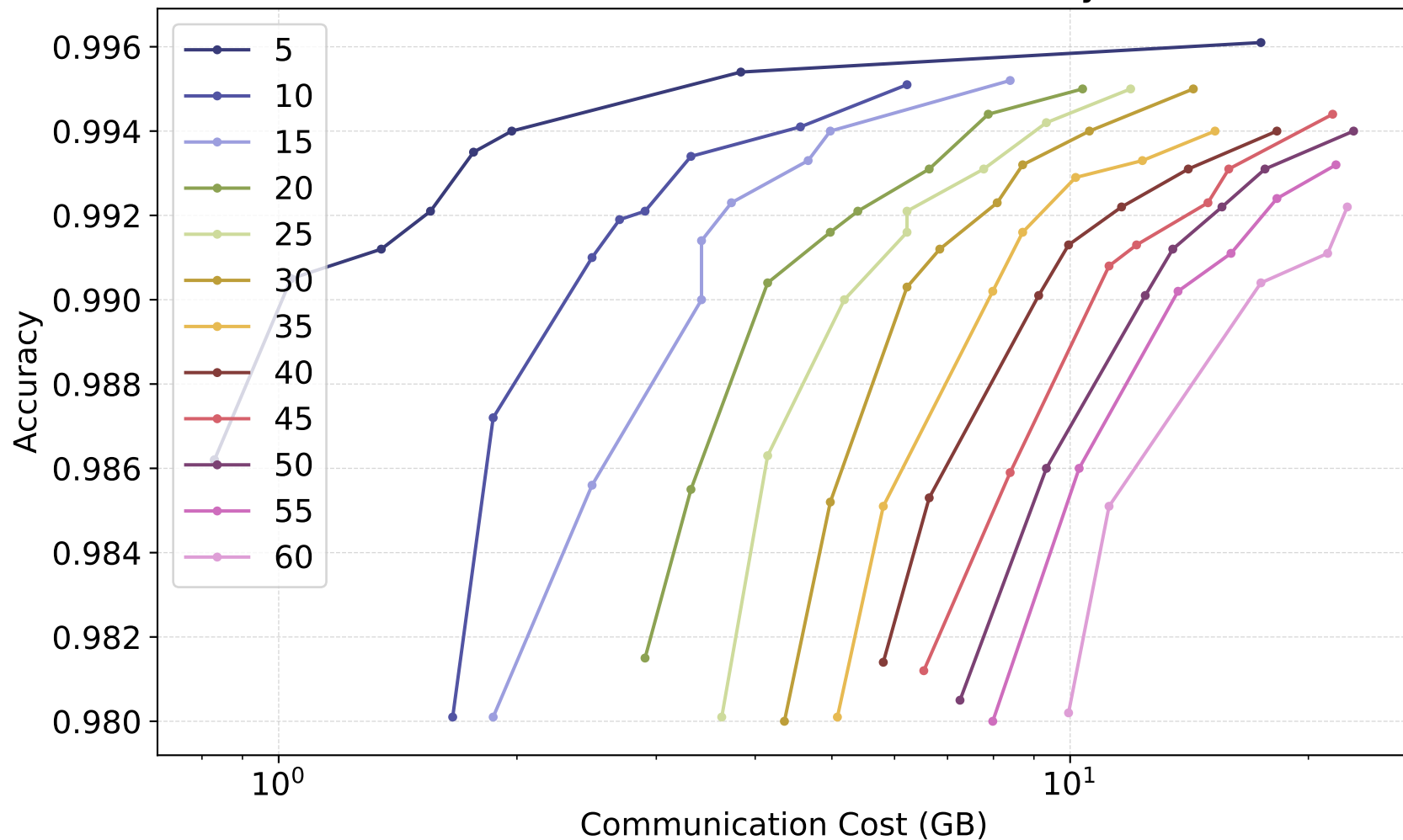


*Theta* : 100.0 , Batch Size: 32 , Bias: only label 0



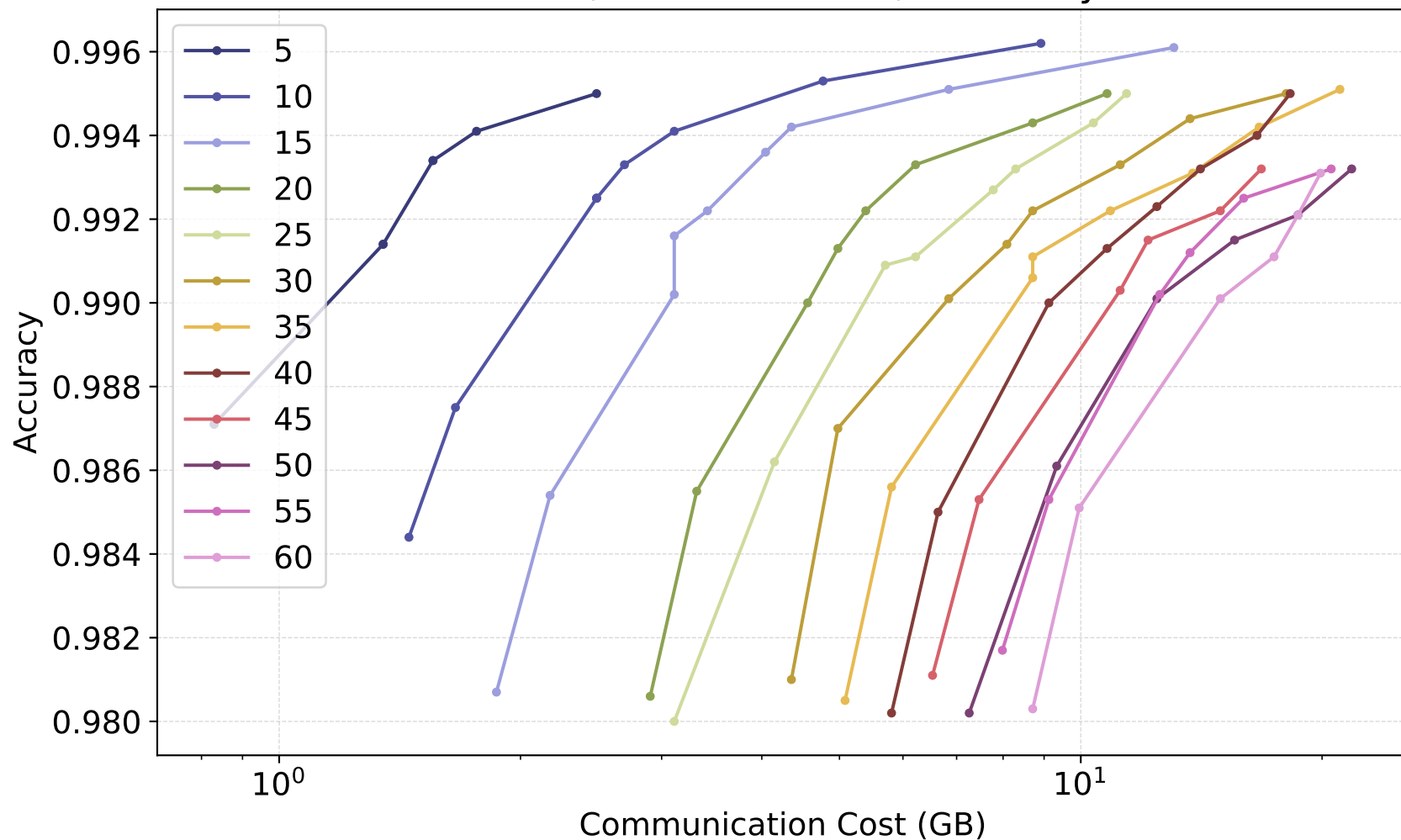
naive

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

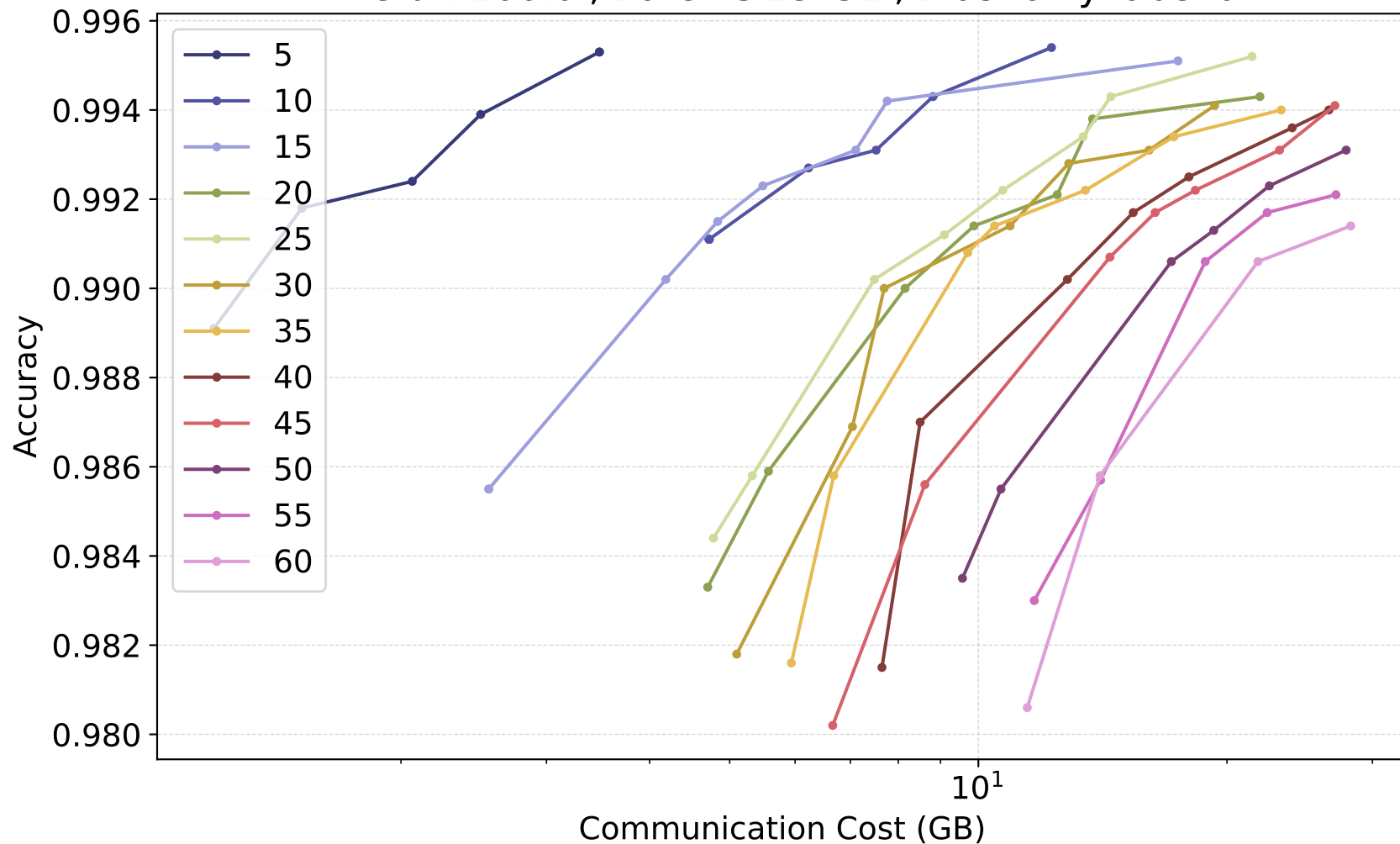


linear

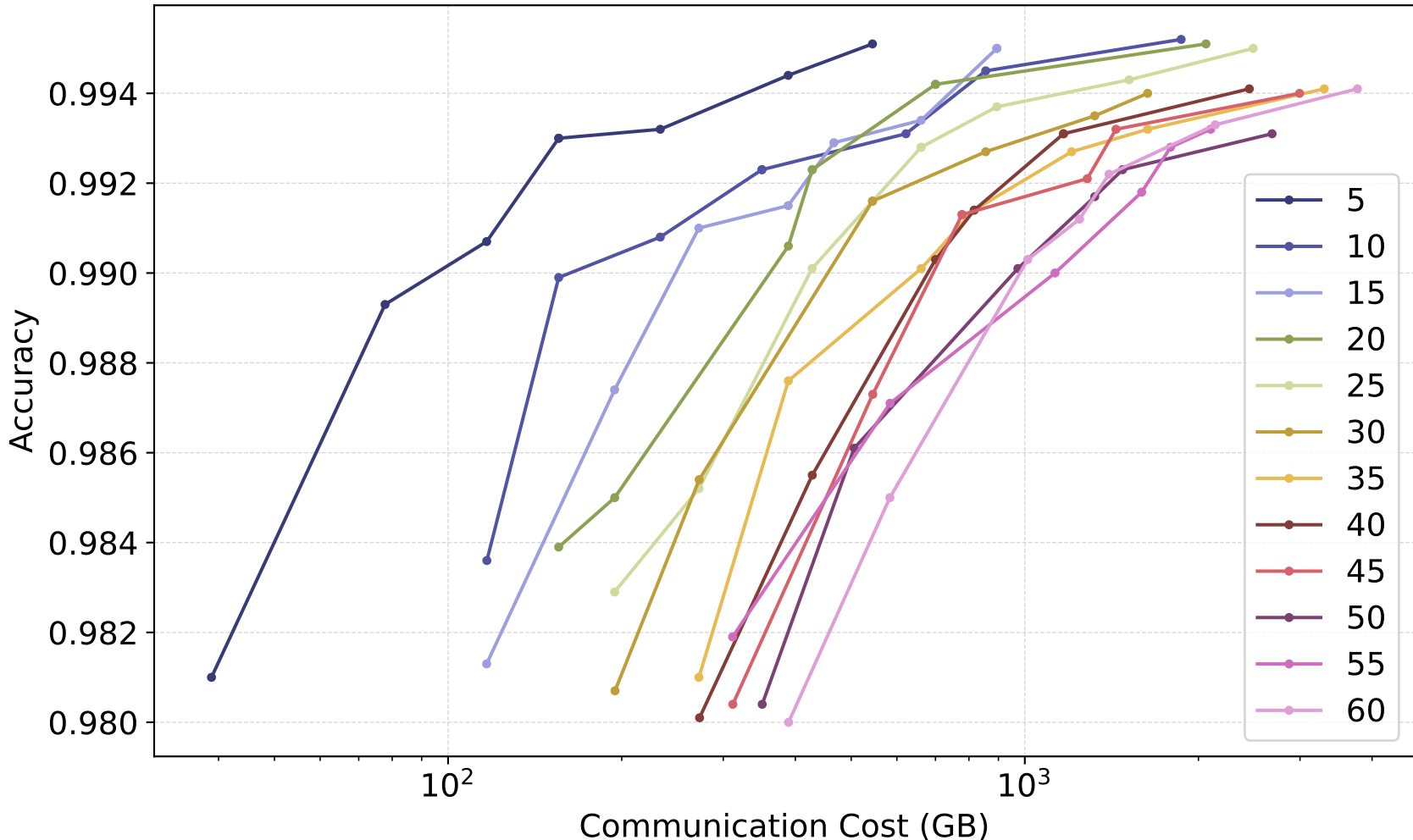
*Theta* : 100.0 , Batch Size: 32 , Bias: only label 0



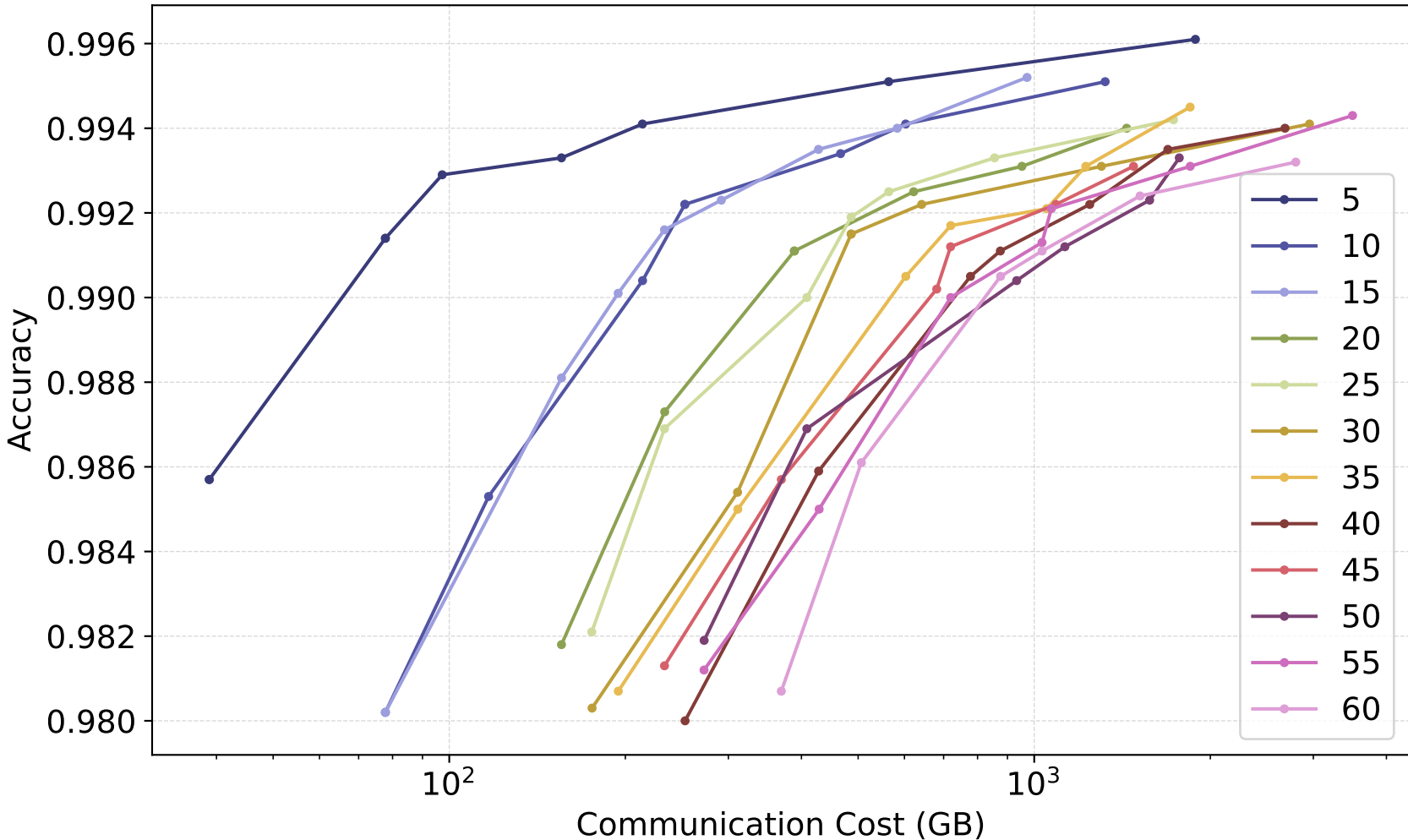
*Theta* : 100.0 , Batch Size: 32 , Bias: only label 0



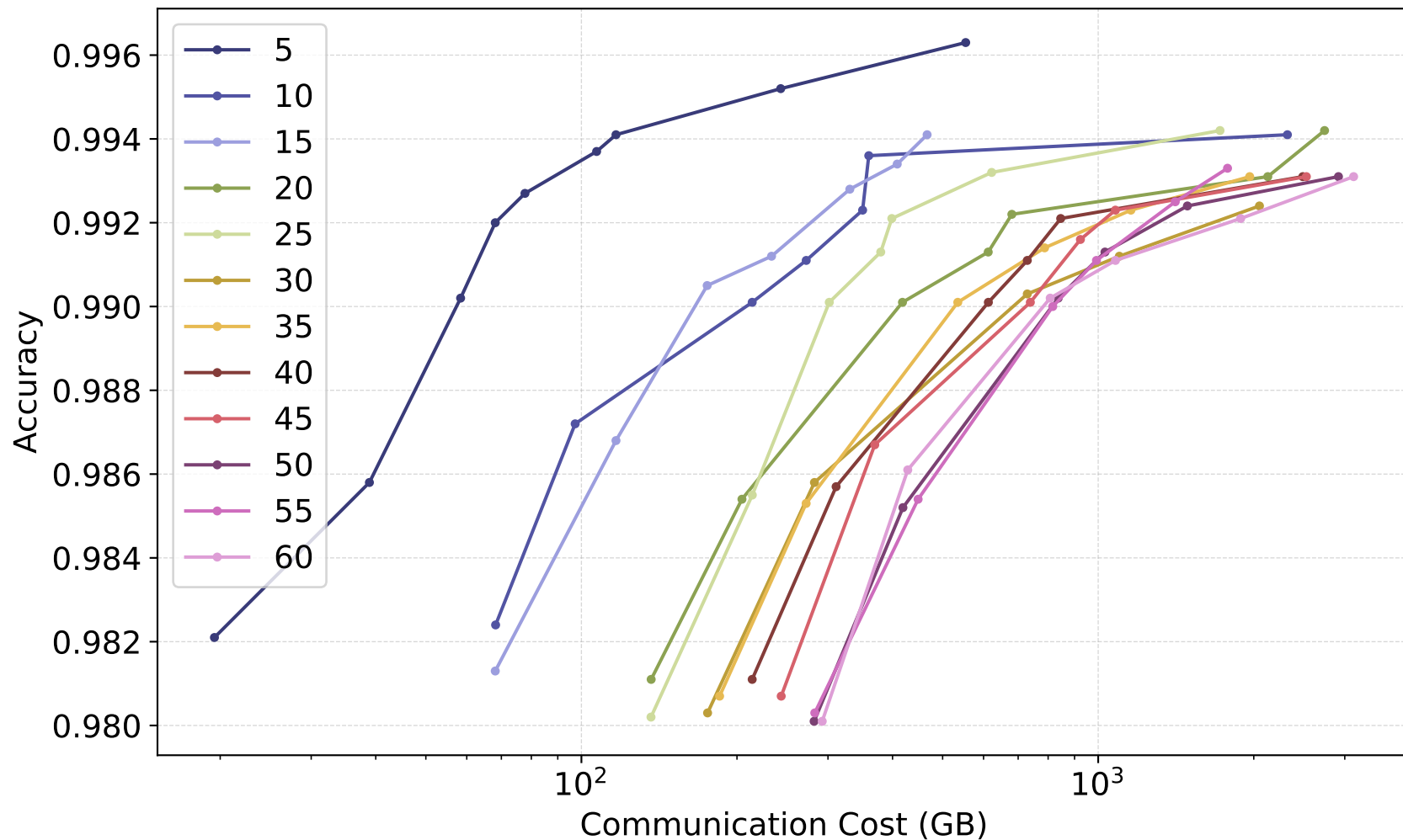
synchronous  
Batch Size : 32 , Bias: 0.9



synchronous  
Batch Size : 64 , Bias: 0.9

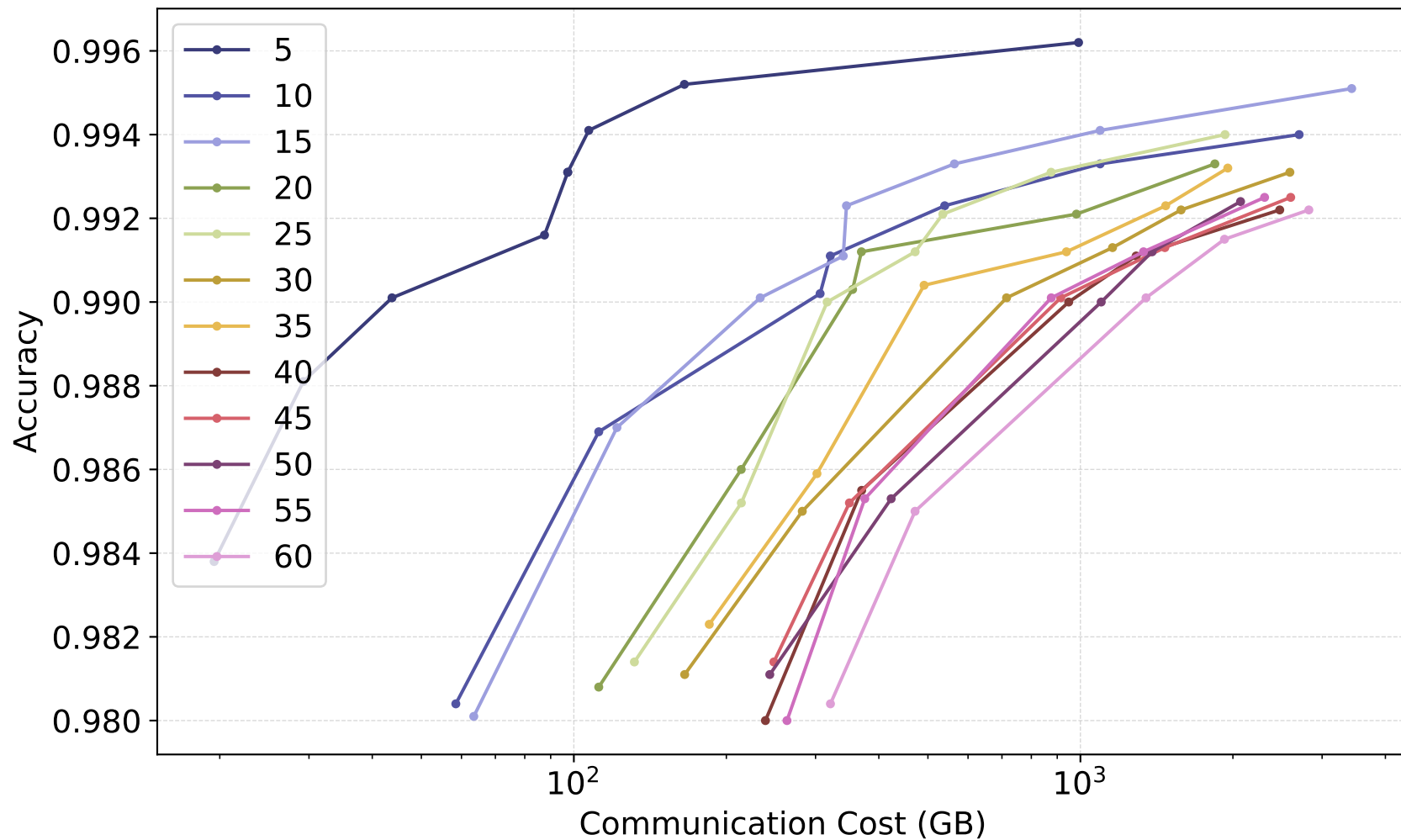


synchronous  
Batch Size : 128 , Bias: 0.9

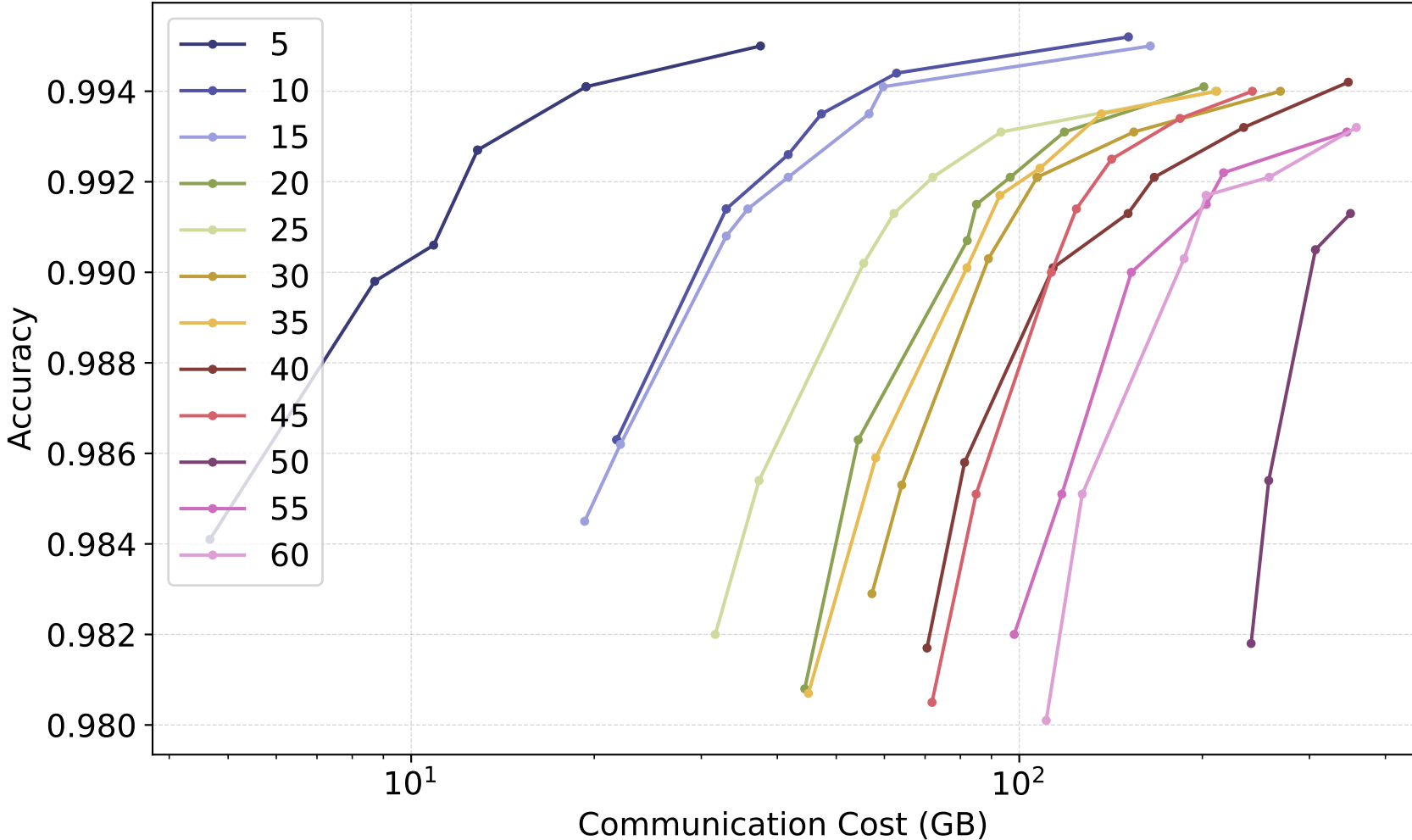




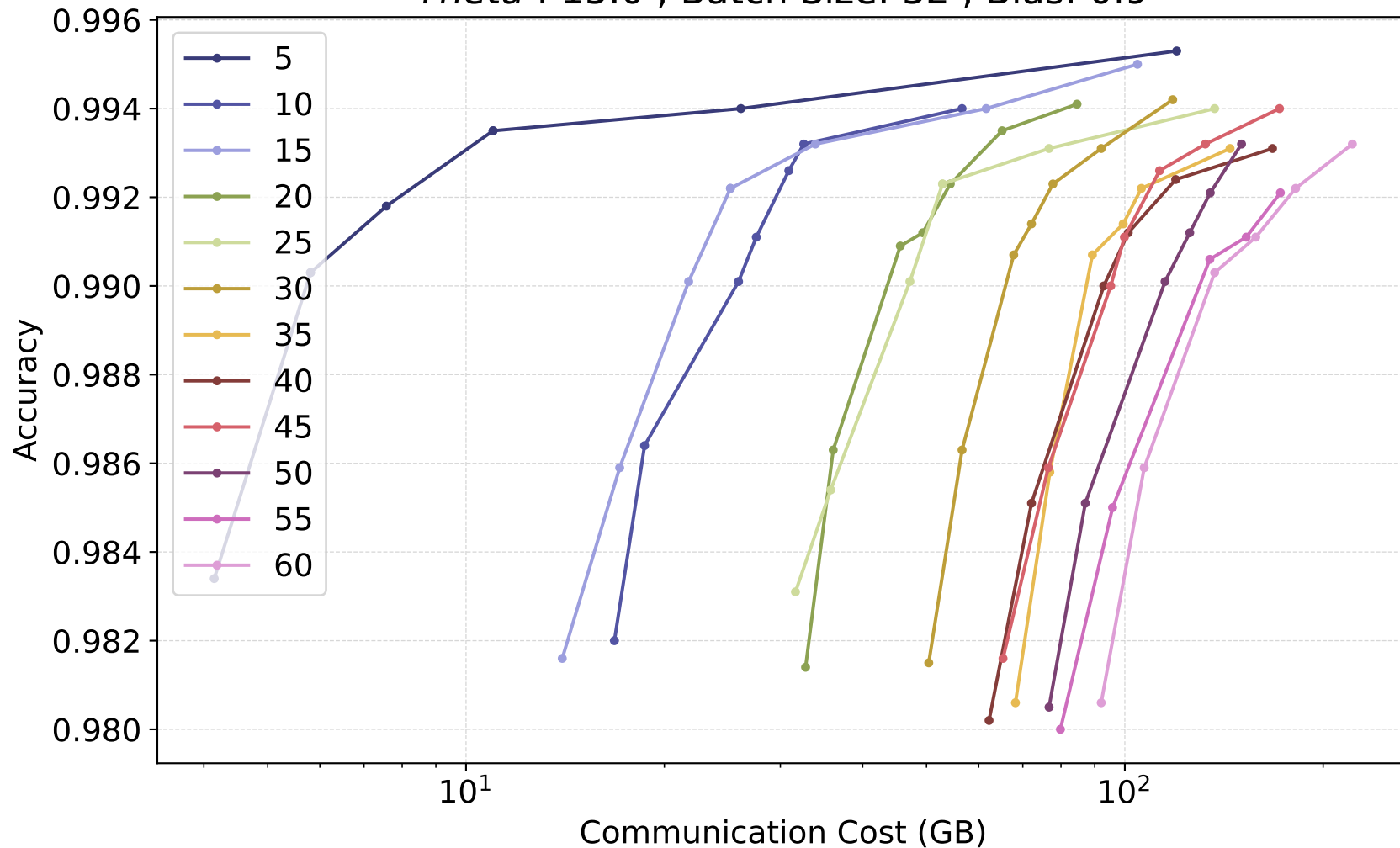
synchronous  
Batch Size : 256 , Bias: 0.9



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

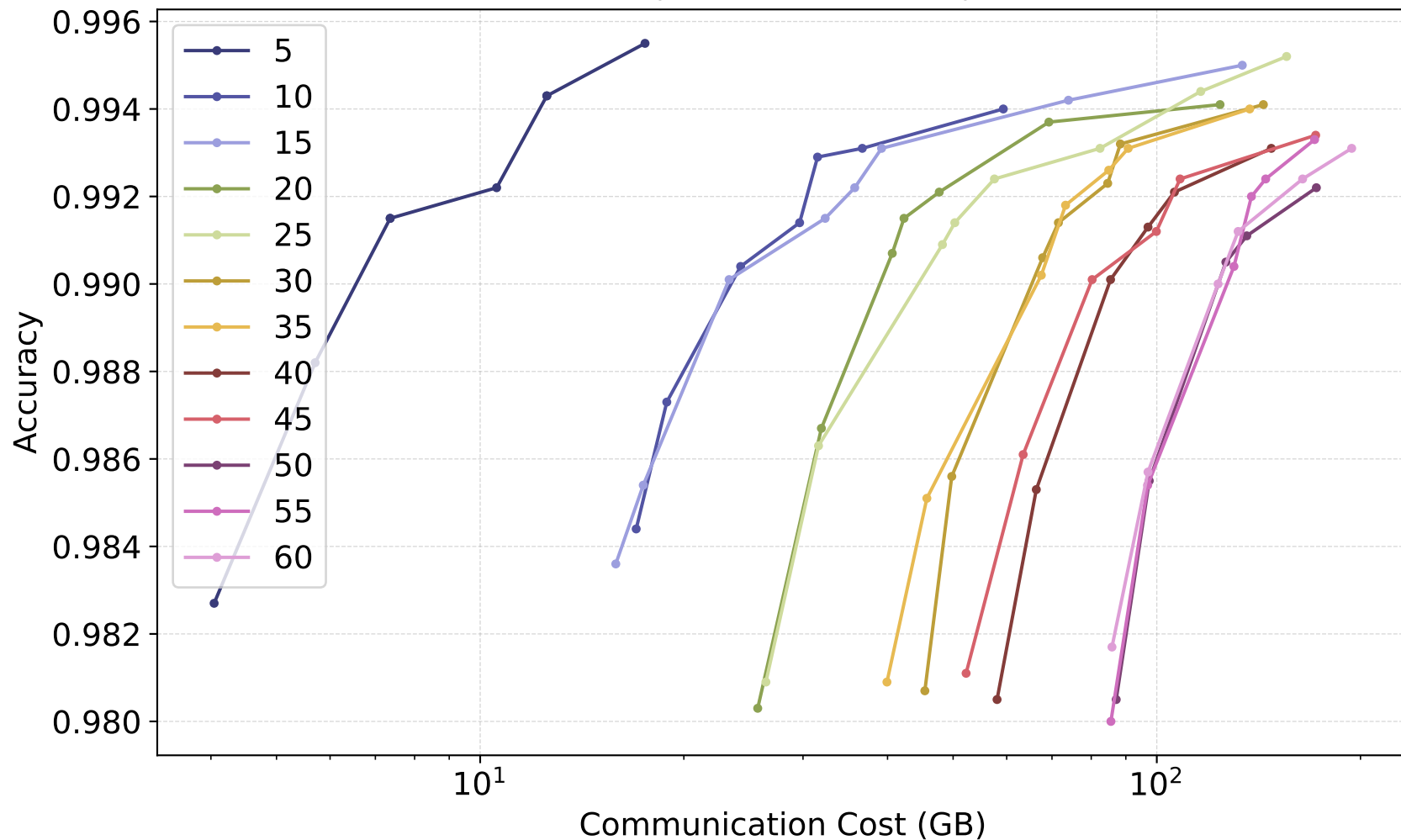


*Theta* : 15.0 , Batch Size: 32 , Bias: 0.9



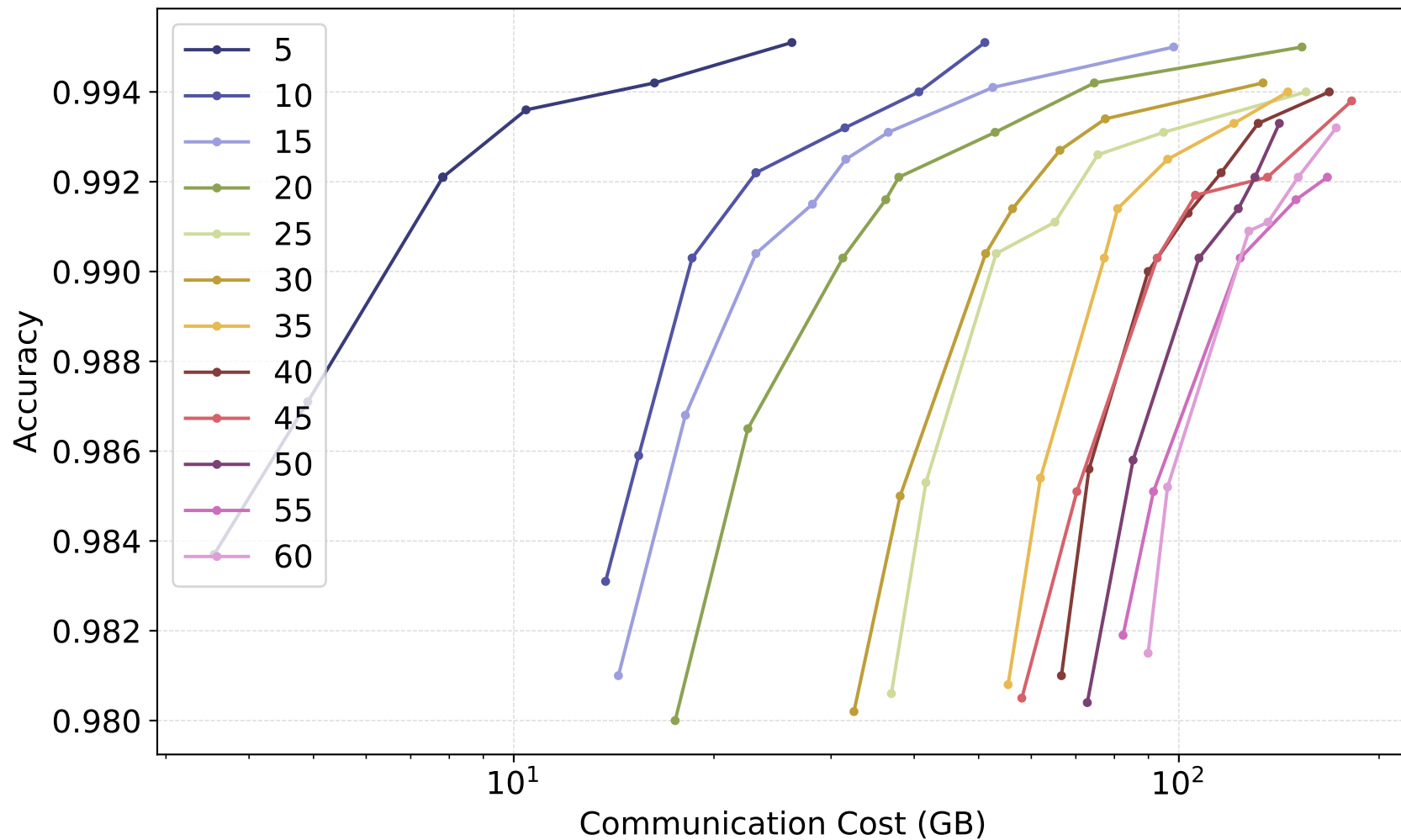
linear

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

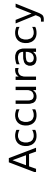


sketch

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

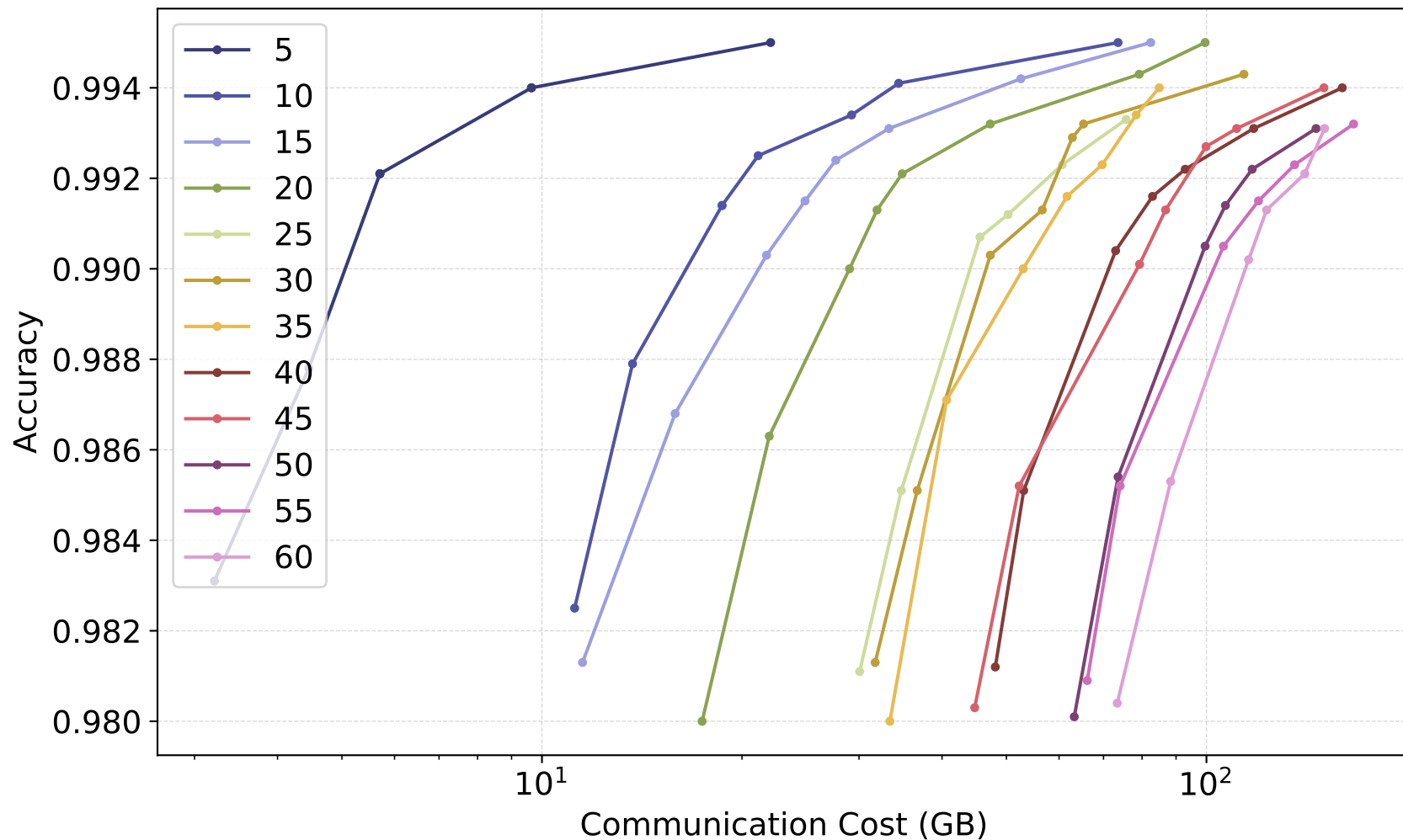


*Theta* : 20.0 , Batch Size: 32 , Bias: 0.9



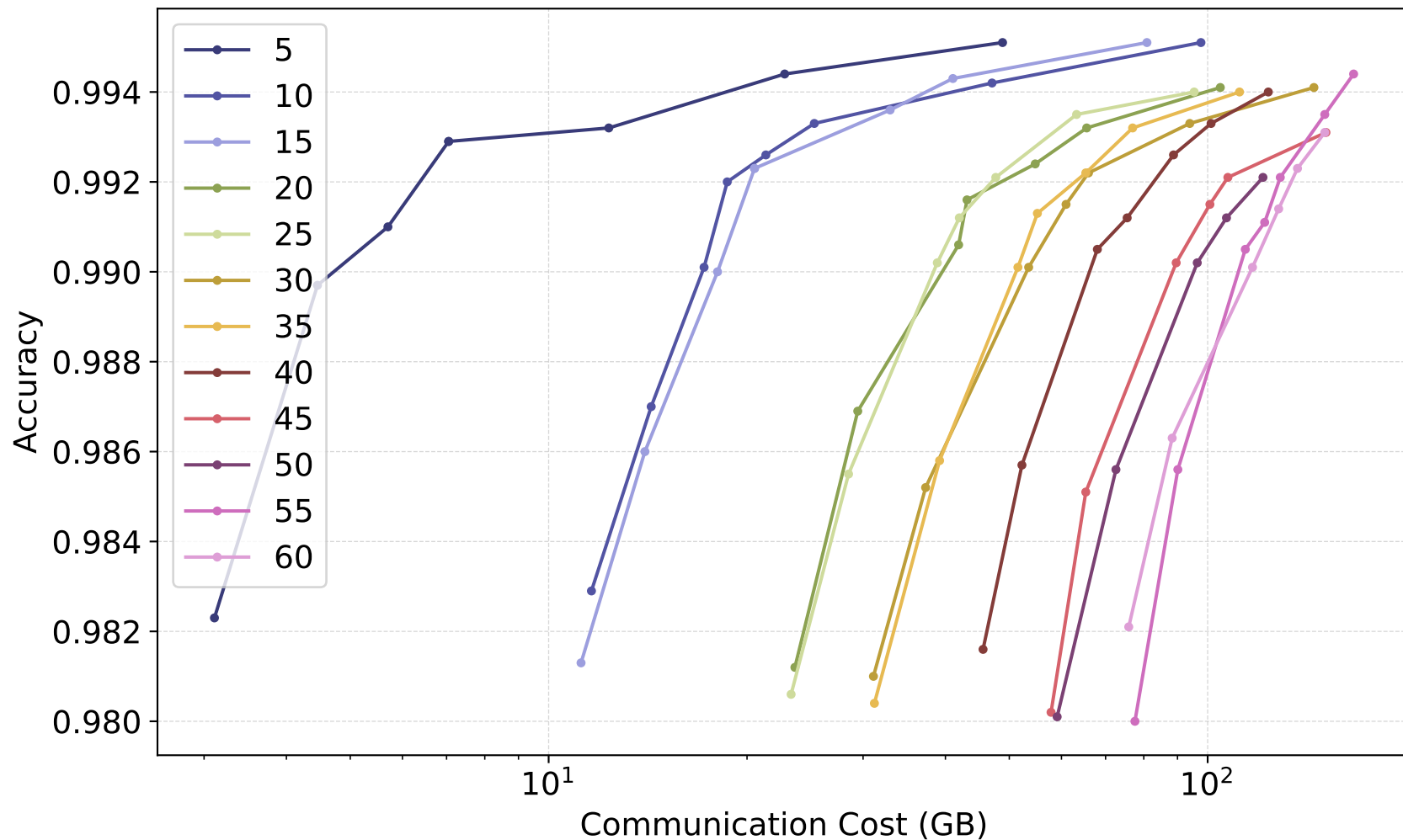
## Communication Cost (GB)

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



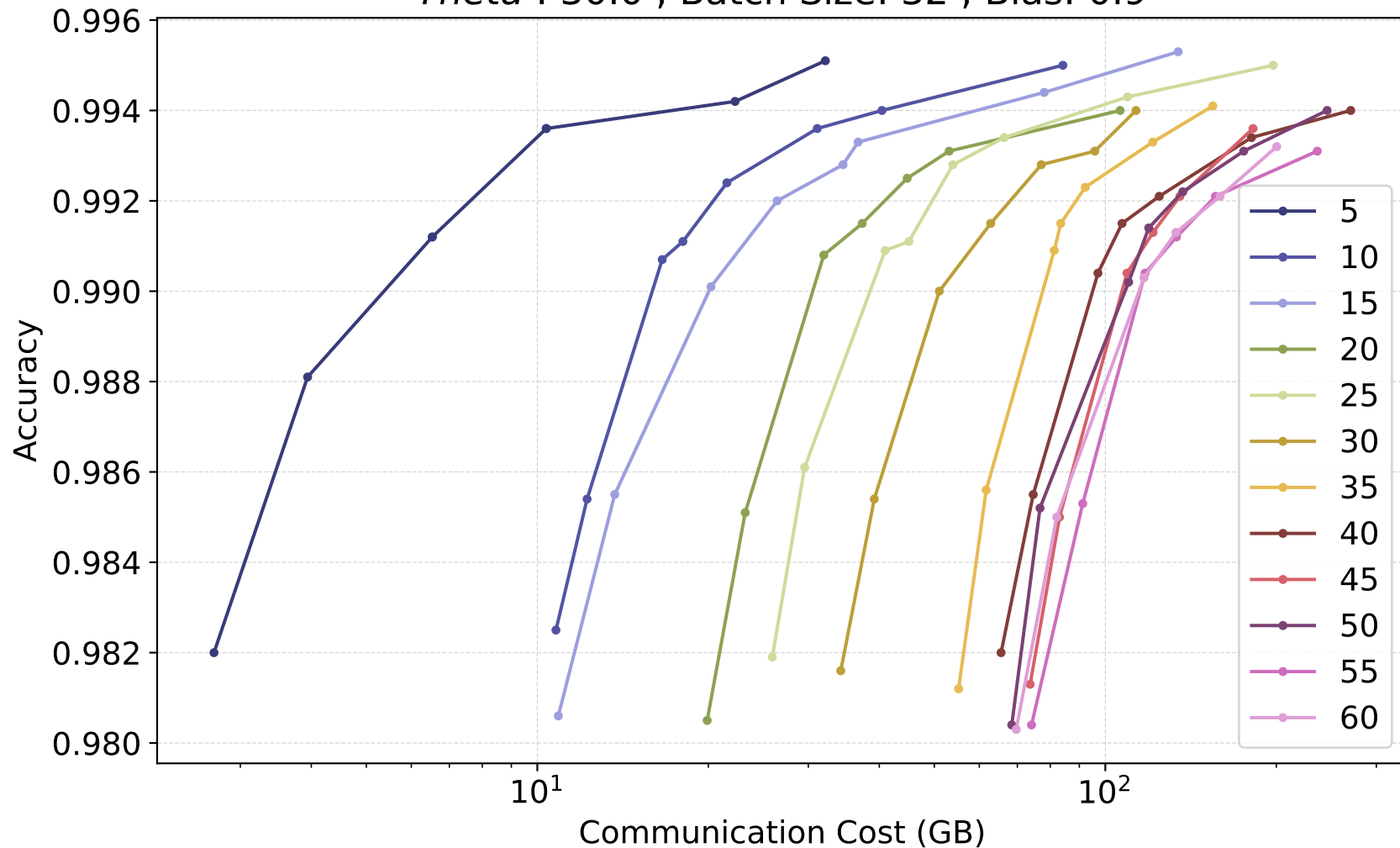
linear

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



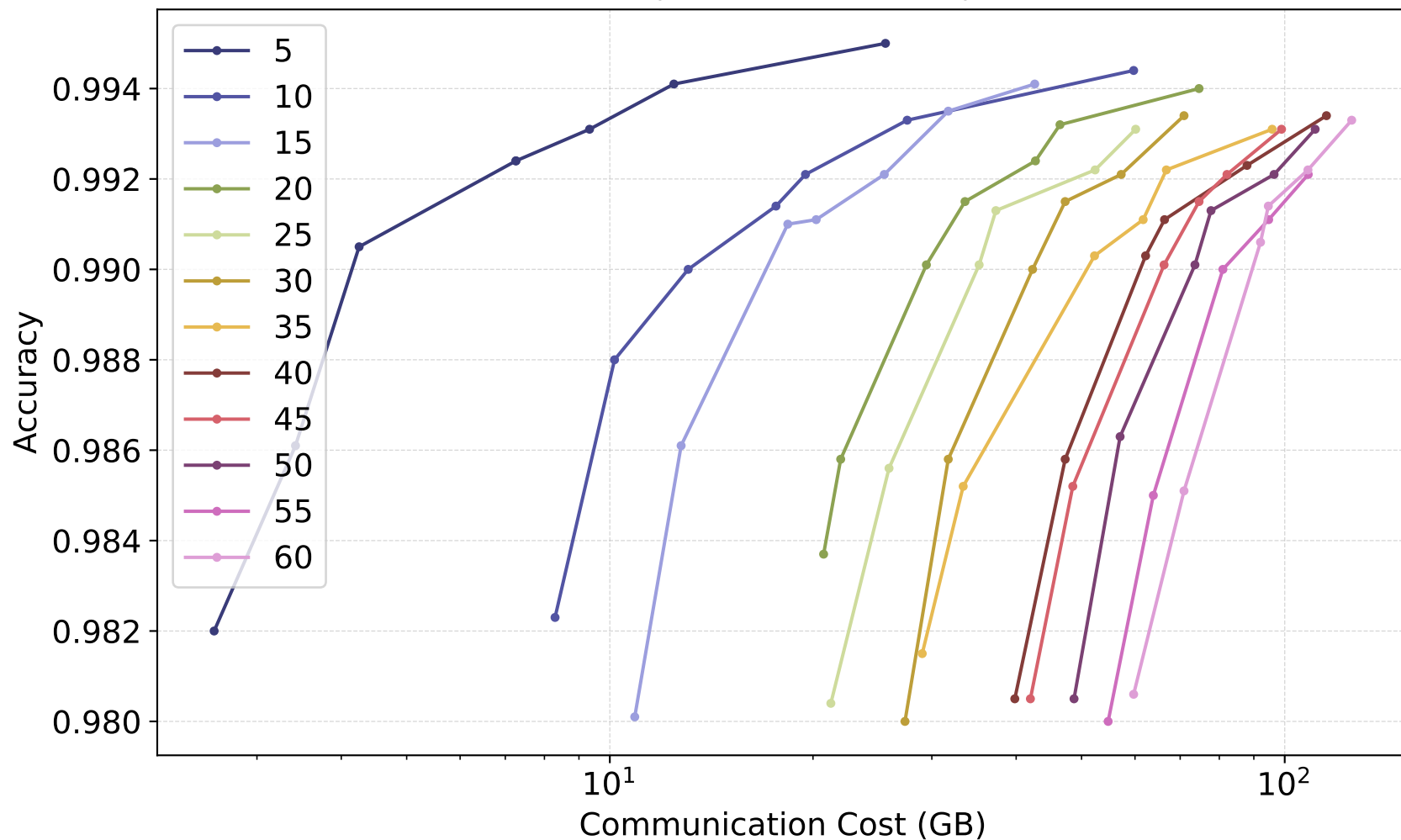


gm  
*Theta* : 30.0 , Batch Size: 32 , Bias: 0.9



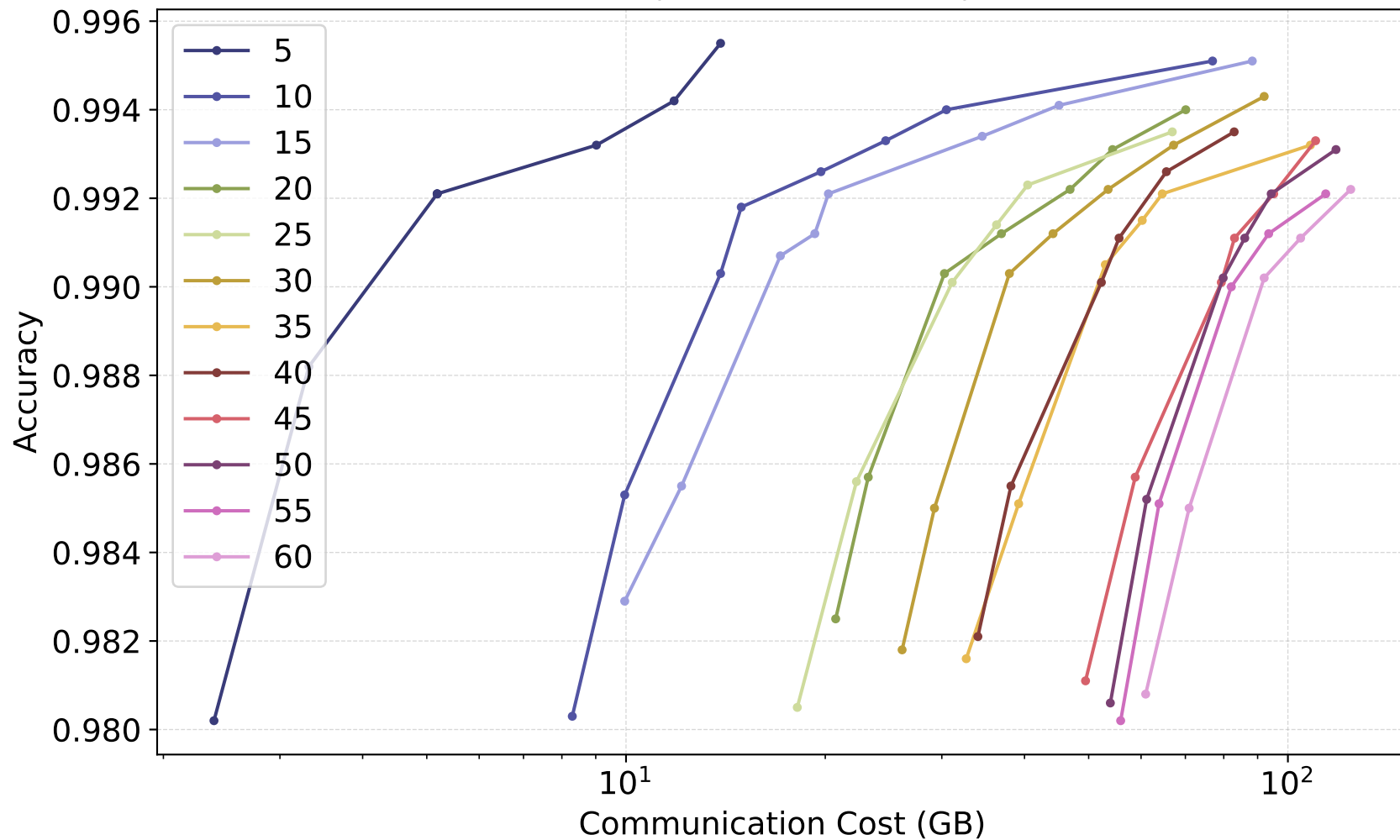
naive

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



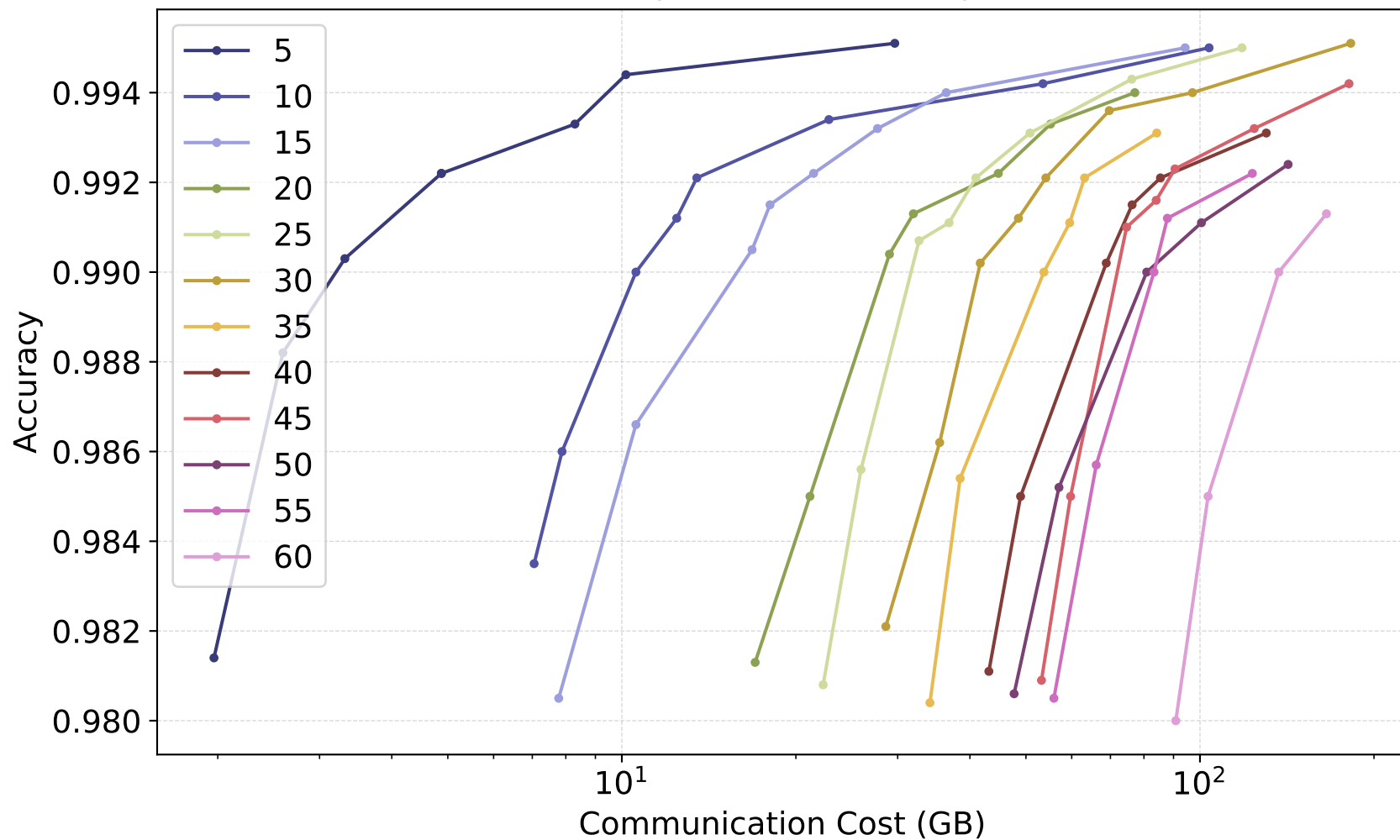
linear

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



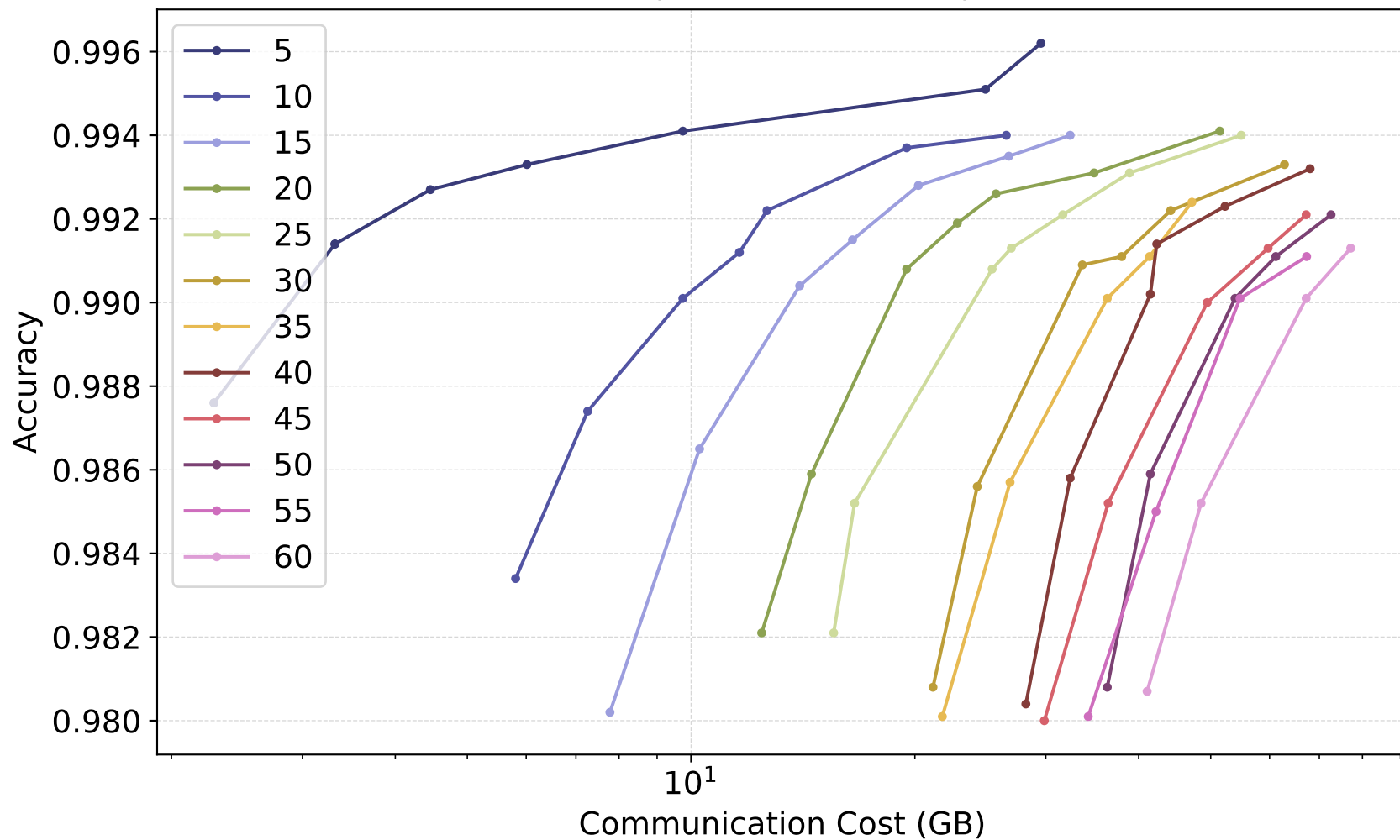
gm

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



naive

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

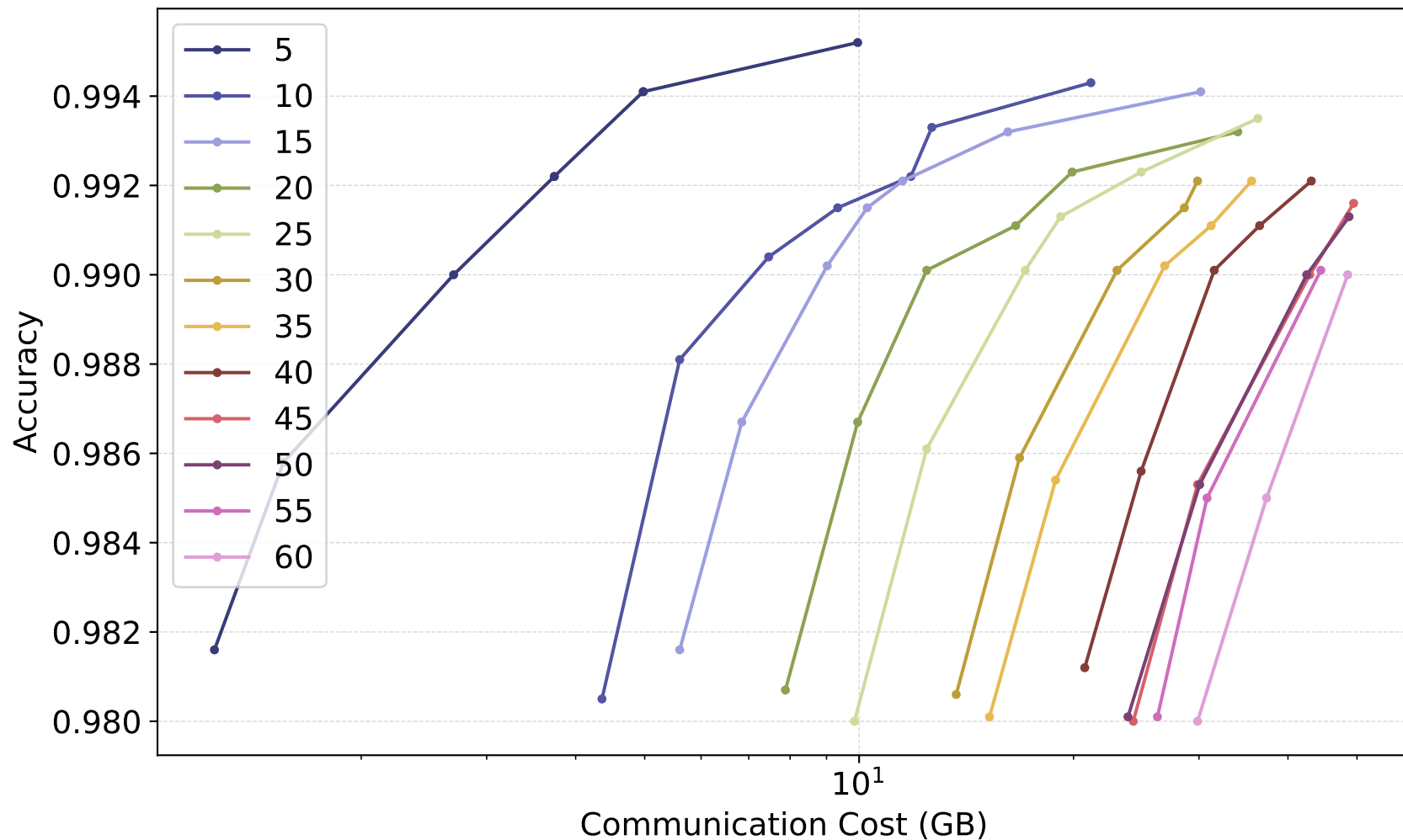






naive

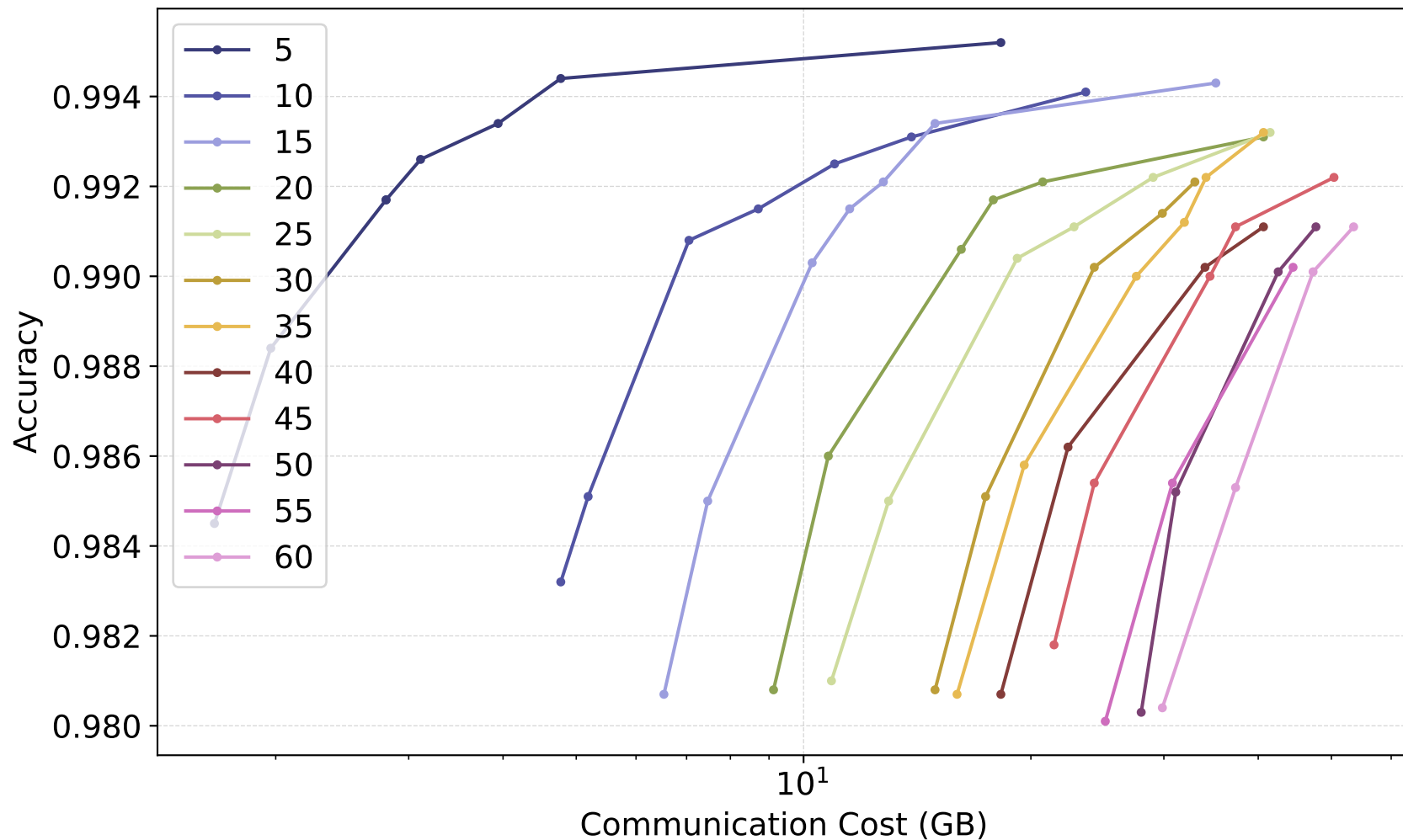
$\Theta$  : 75.0 , Batch Size: 32 , Bias: 0.9



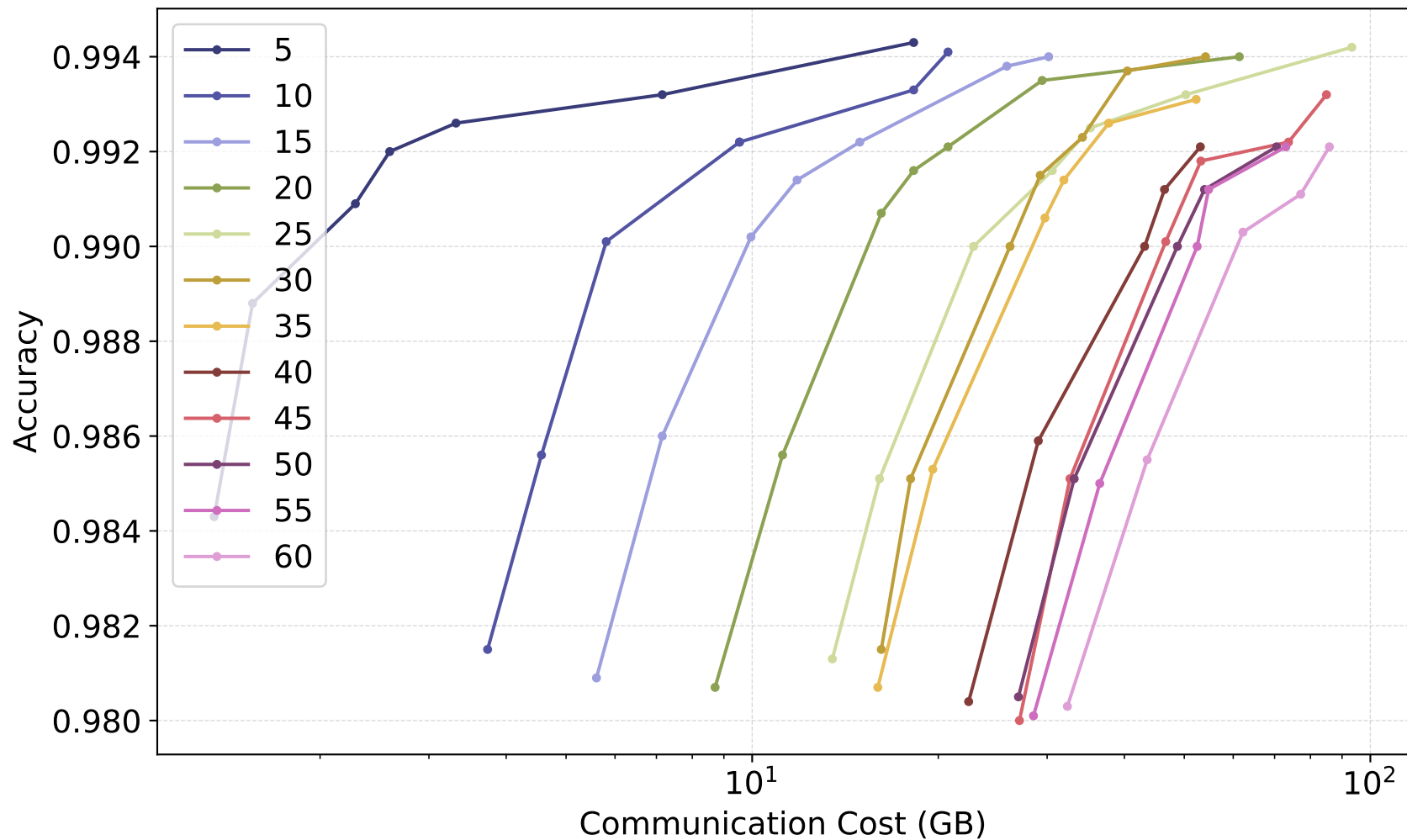


linear

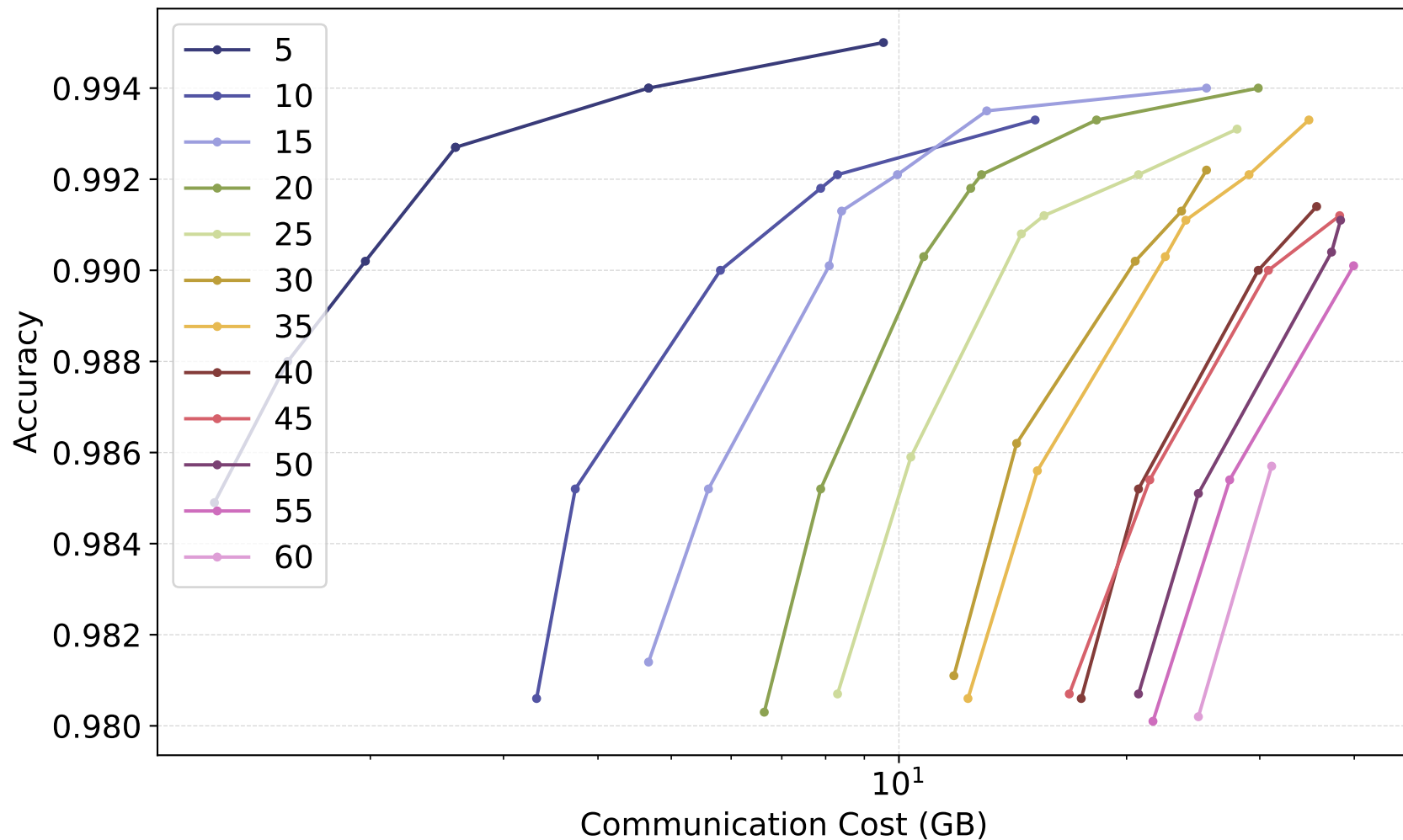
$\Theta$  : 75.0 , Batch Size: 32 , Bias: 0.9



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

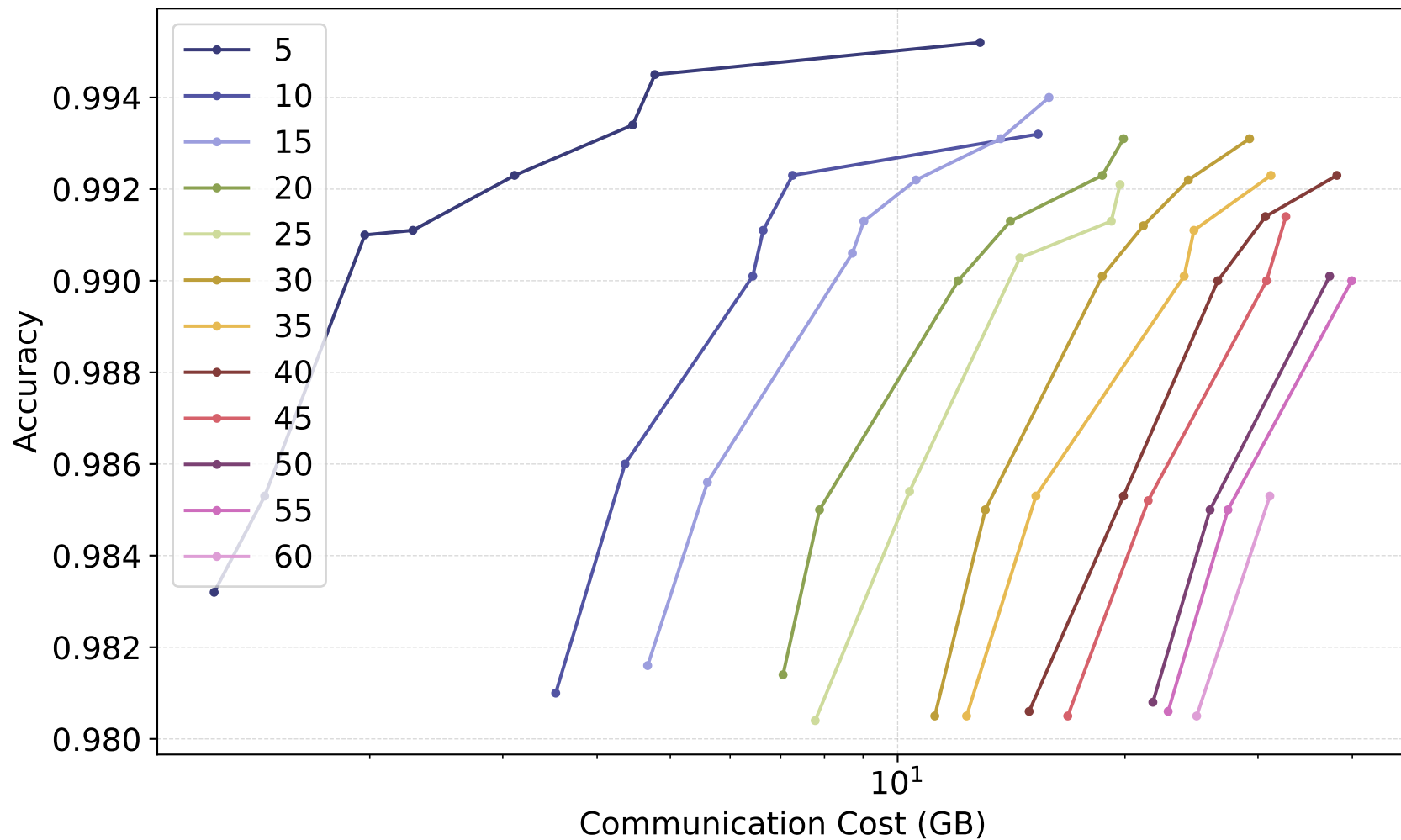


*Theta* : 100.0 , Batch Size: 32 , Bias: 0.9



linear

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



sketch

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

