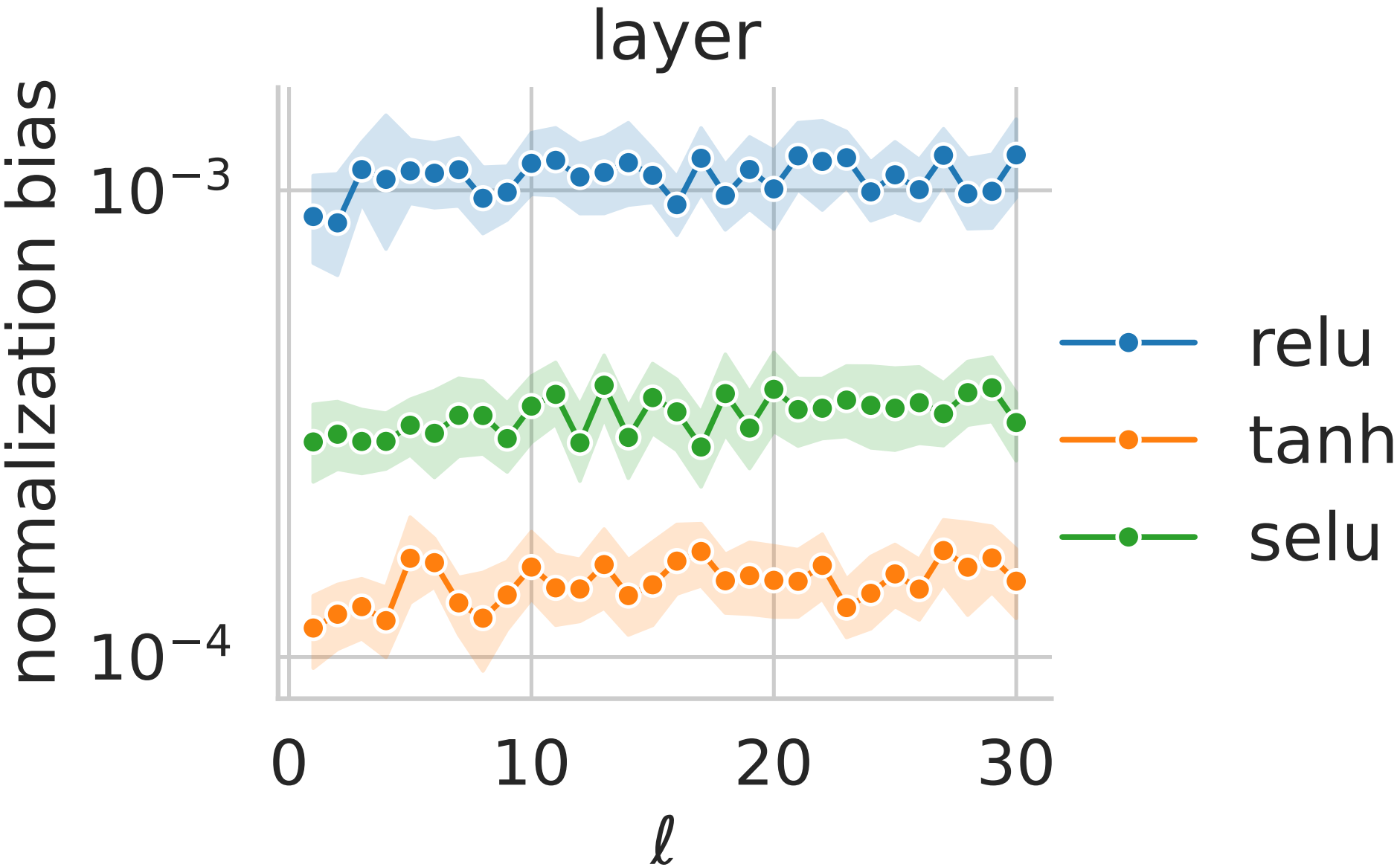
layer
normalization bias
10^{-3}
10^{-4}
0
10
20
30
ℓ
relu
tanh
selu