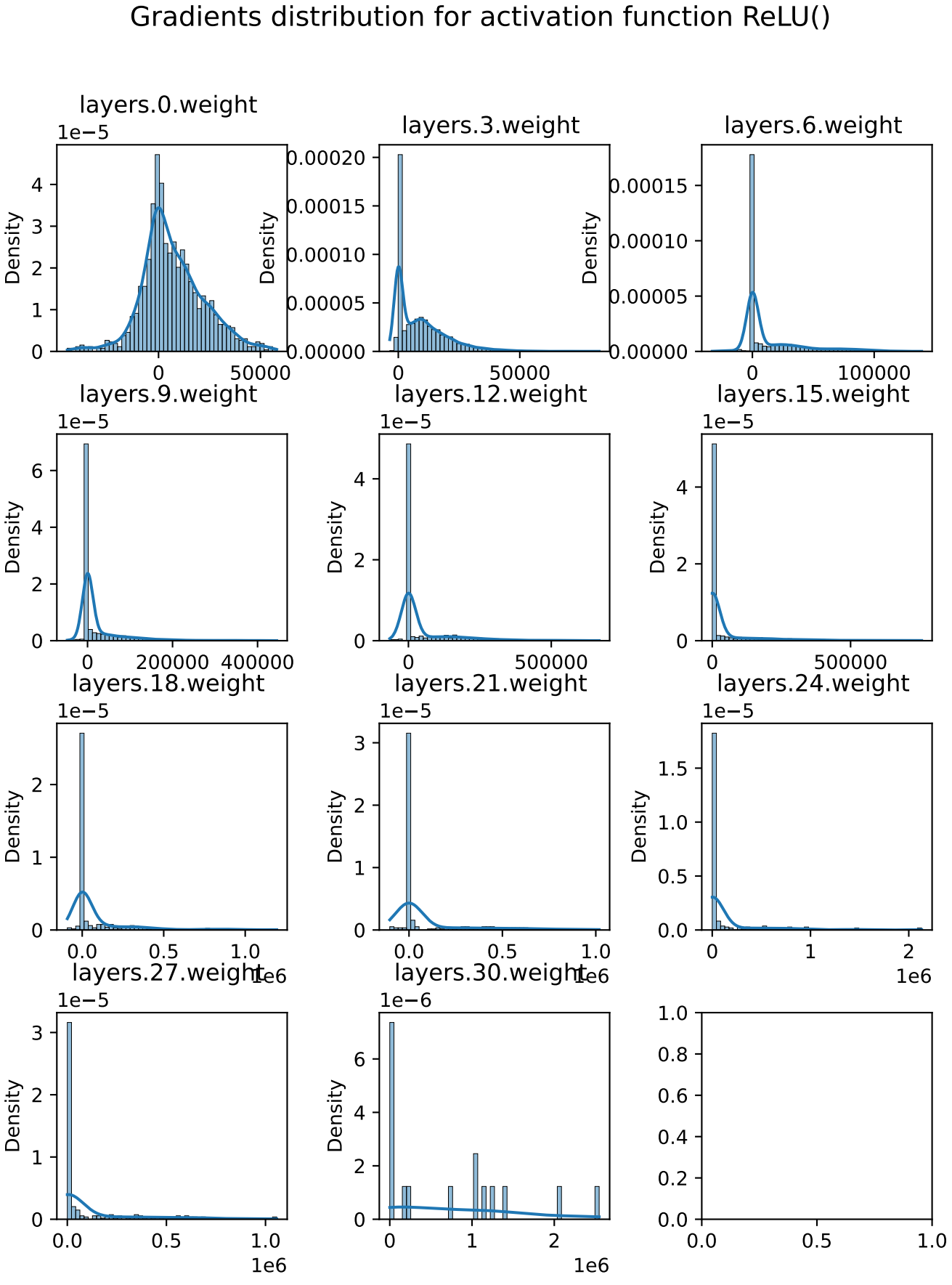
Gradients distribution for activation function ReLU()
layers.0.weight
layers.3.weight
layers.6.weight
layers.9.weight
layers.12.weight
layers.15.weight
layers.18.weight
layers.21.weight
layers.24.weight
layers.27.weight
layers.30.weight
Density
1e−5
1e−6
1e6