Training curves – ReLU activation (different depths)

Training MSE vs Epochs. Legend: 1 layer, 2 layers, 3 layers, 4 layers.