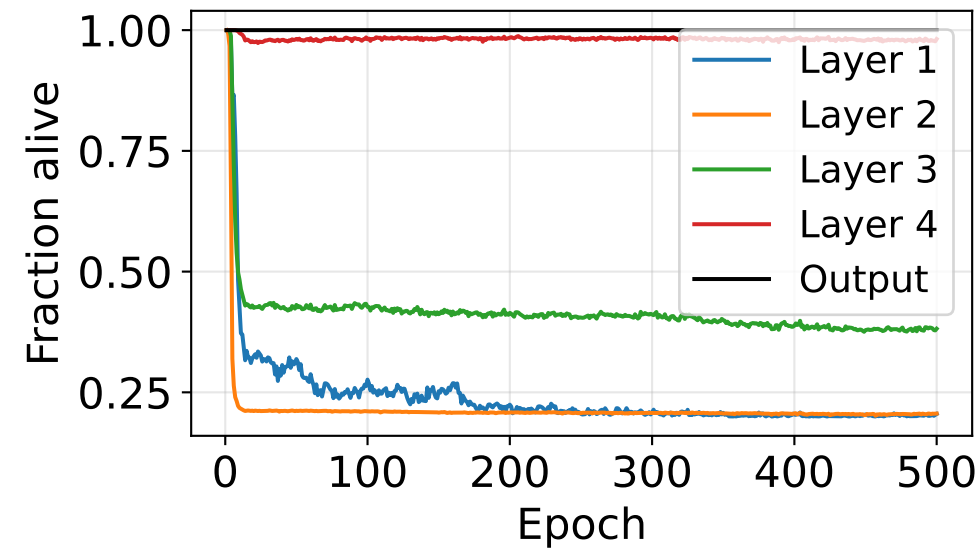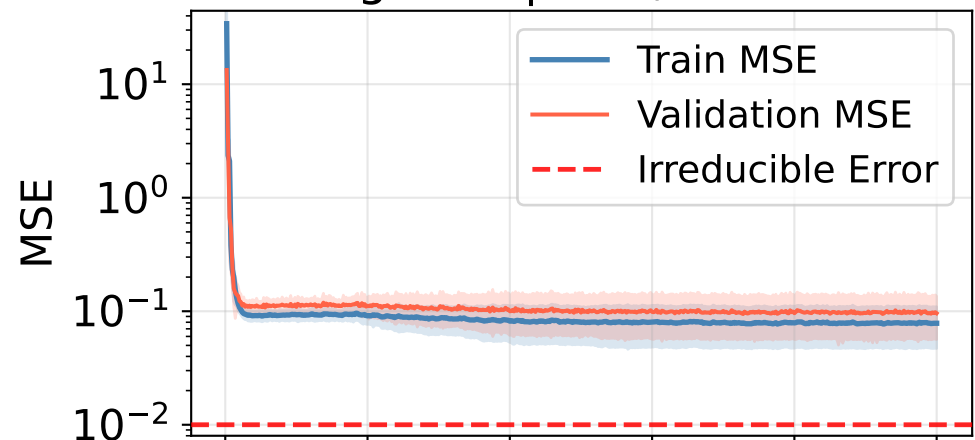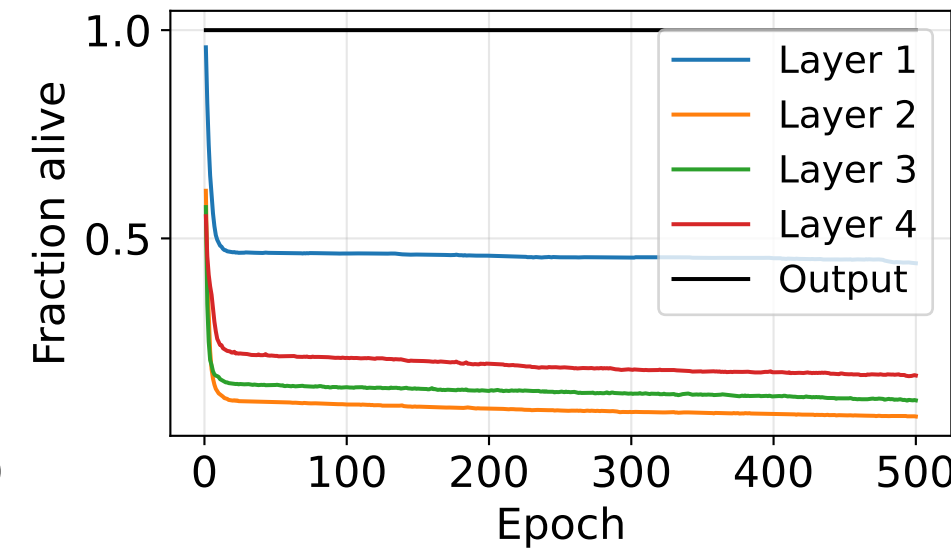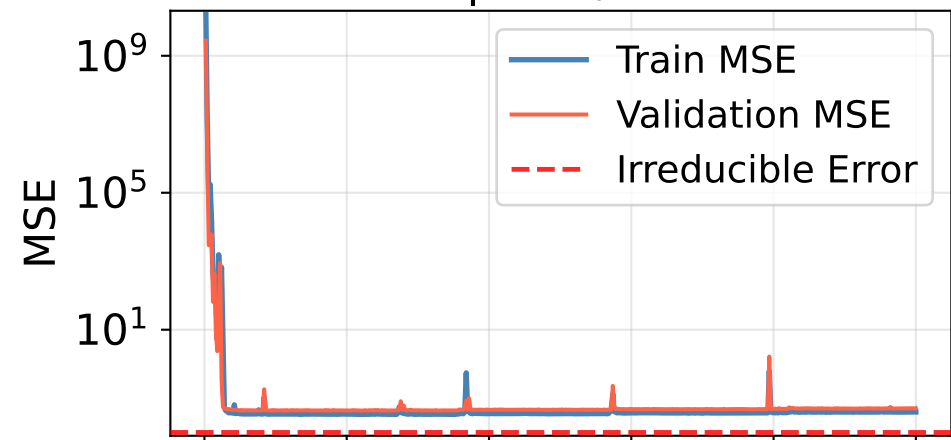**Learning vs Neuron Death (Without BatchNorm)**

Sigmoid | L=4, N=250 — ReLU | L=4, N=250 — LeakyReLU | L=4, N=250

Legends: Train MSE, Validation MSE, Irreducible Error; Layer 1, Layer 2, Layer 3, Layer 4, Output. Axes: MSE vs Epoch (top row), Fraction alive vs Epoch (bottom row).