Normalized Distribution of FFN activations for grad\_norm 1e6 6 5 Frequency 2 1 0 0.2 0.6 8.0 1.0 0.0 0.4 Value