

7 Examples of Assumption A.4

Here, we show 3 examples that approximately satisfy Assumption A.4. Figure 1-3 show the distribution of $\Delta\delta$ when AE is trained on these datasets with 100 training epochs, $\text{batch_size} = |D|$ and 0.001 learning rate.

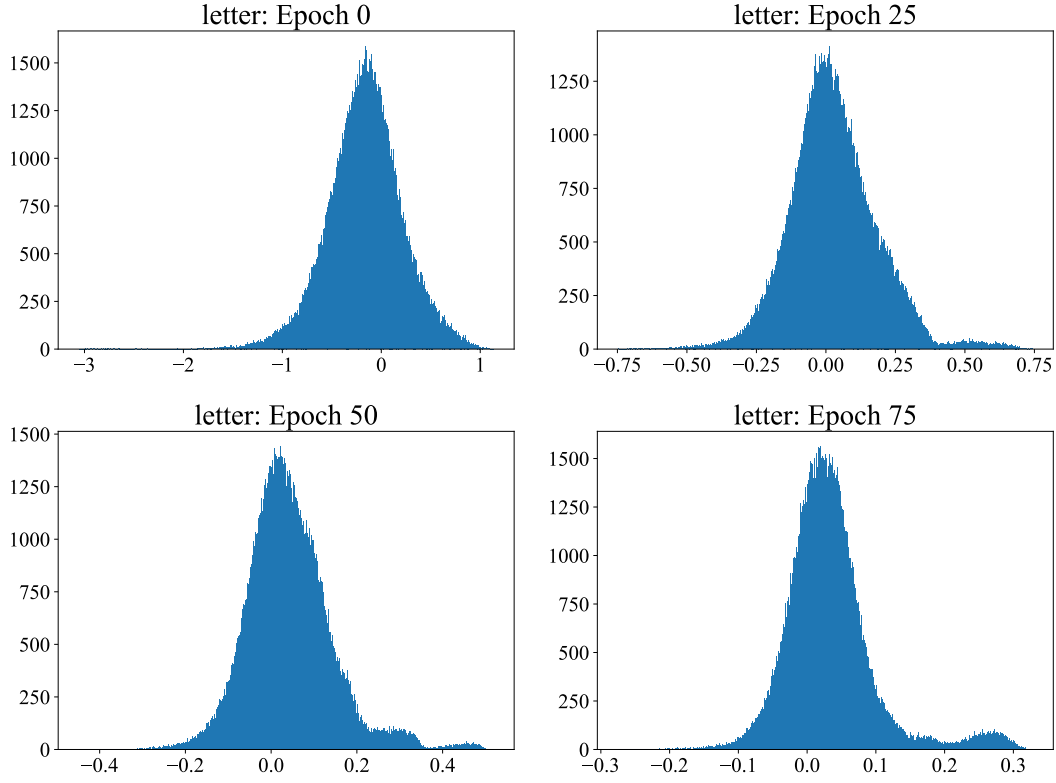


Figure 1: The distribution of $\Delta\delta$ for training AE on *letter*.

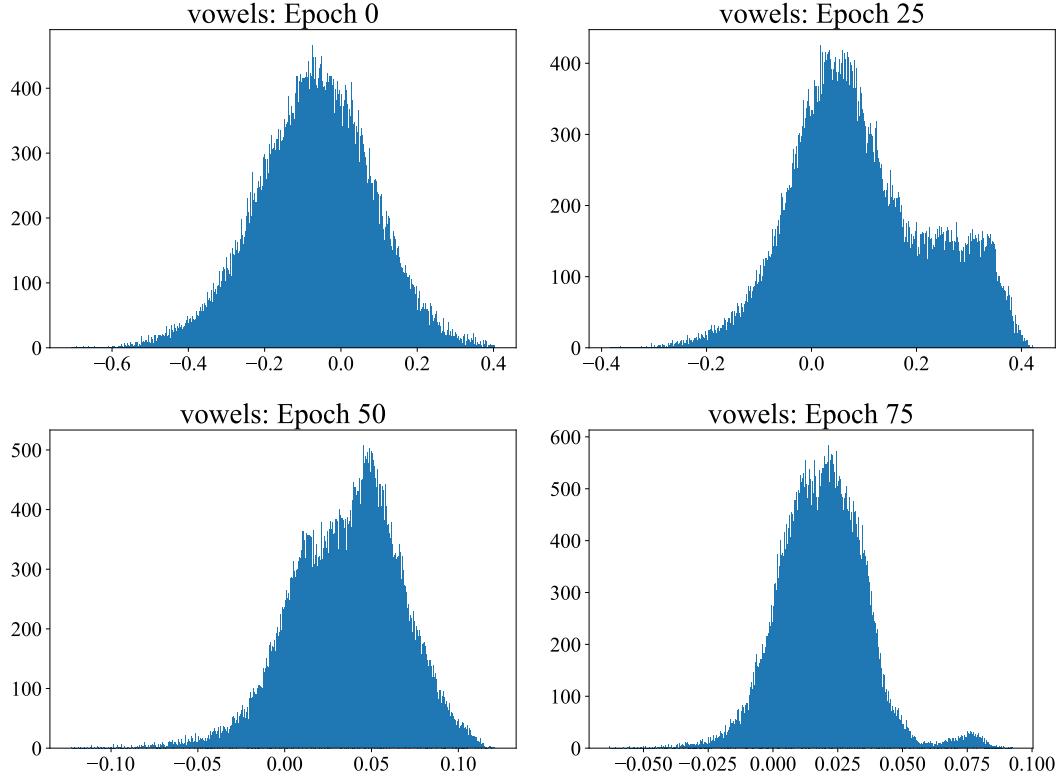


Figure 2: The distribution of $\Delta\delta$ for training AE on *vowels*.

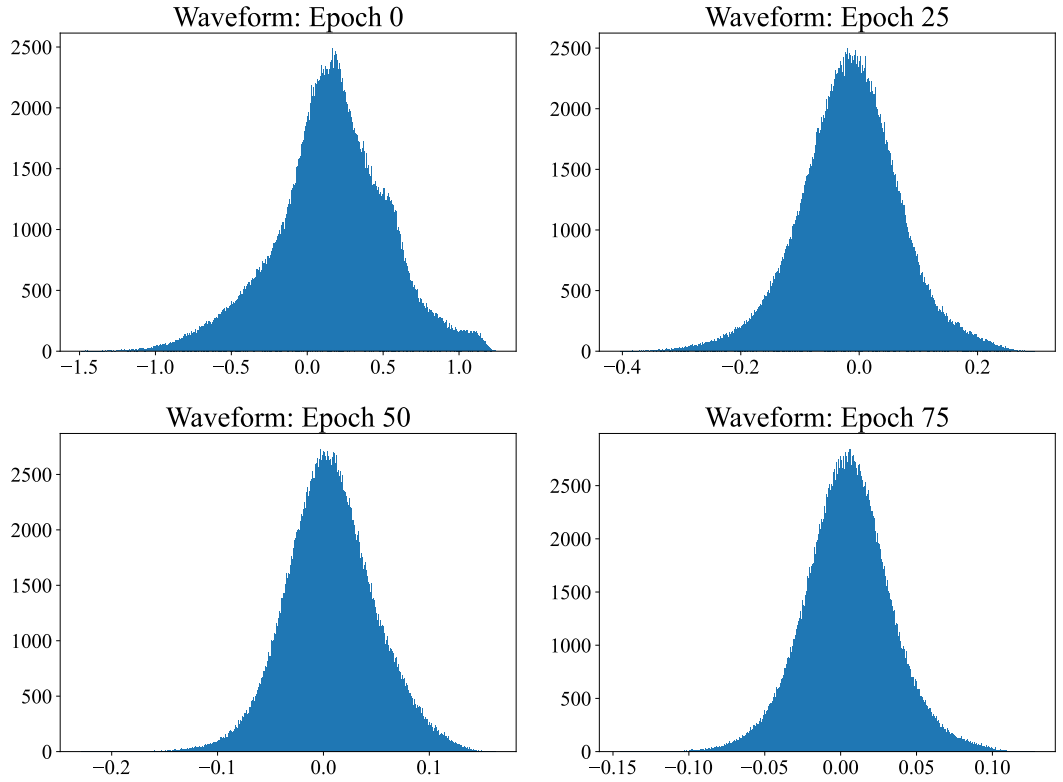


Figure 3: The distribution of $\Delta\delta$ for training AE on *Waveform*.