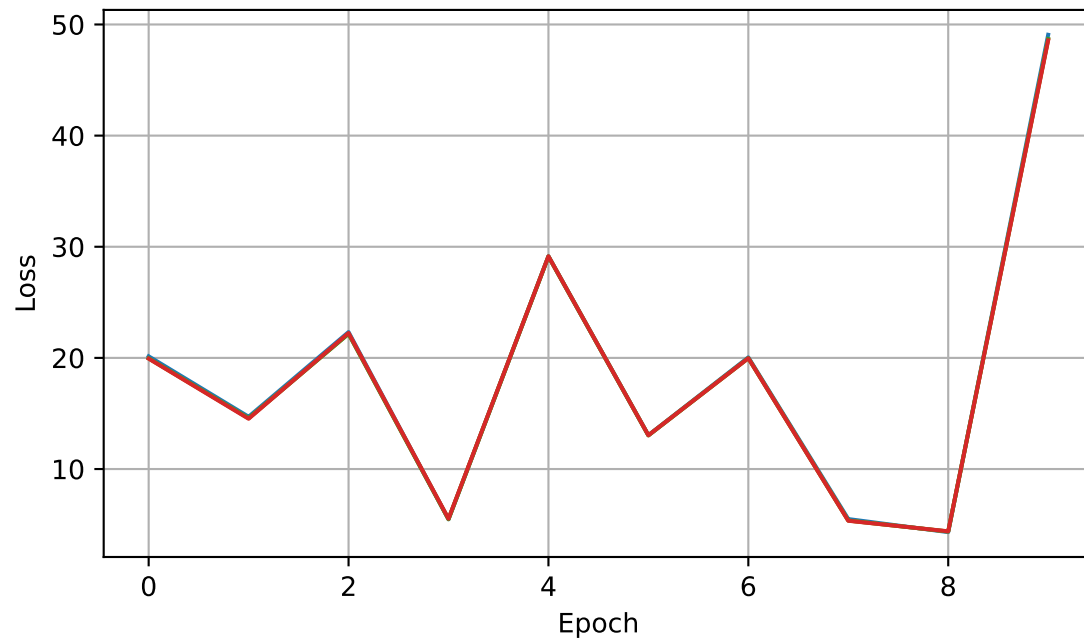
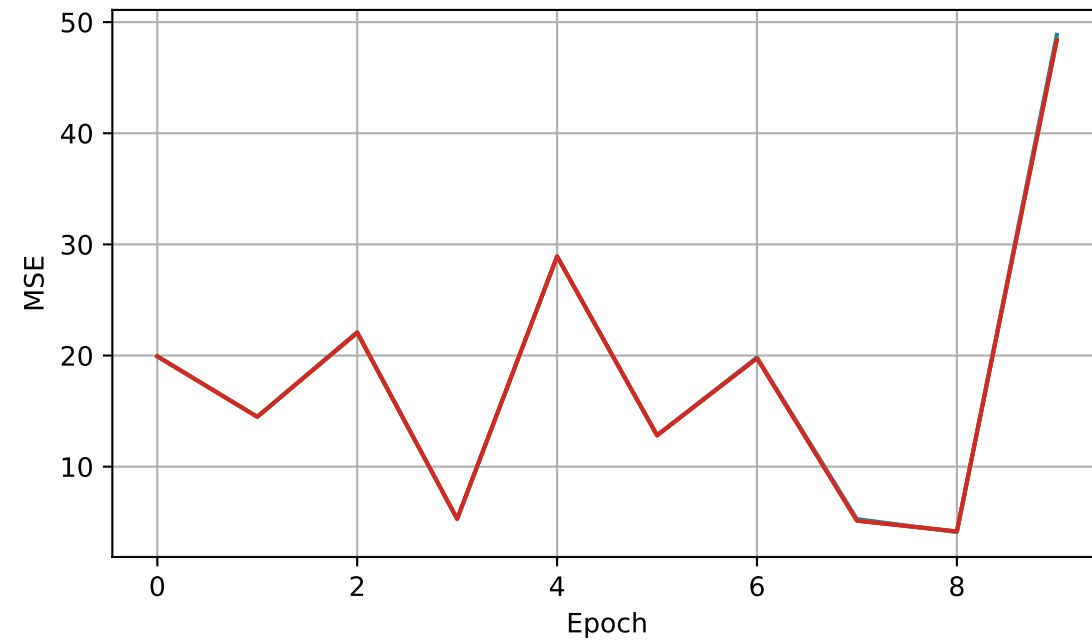


constant, $\beta=1.00$, warmup=5.0 linear, $\beta=1.00$, warmup=5.0 sigmoid, $\beta=1.00$, warmup=5.0
cosine, $\beta=1.00$, warmup=5.0

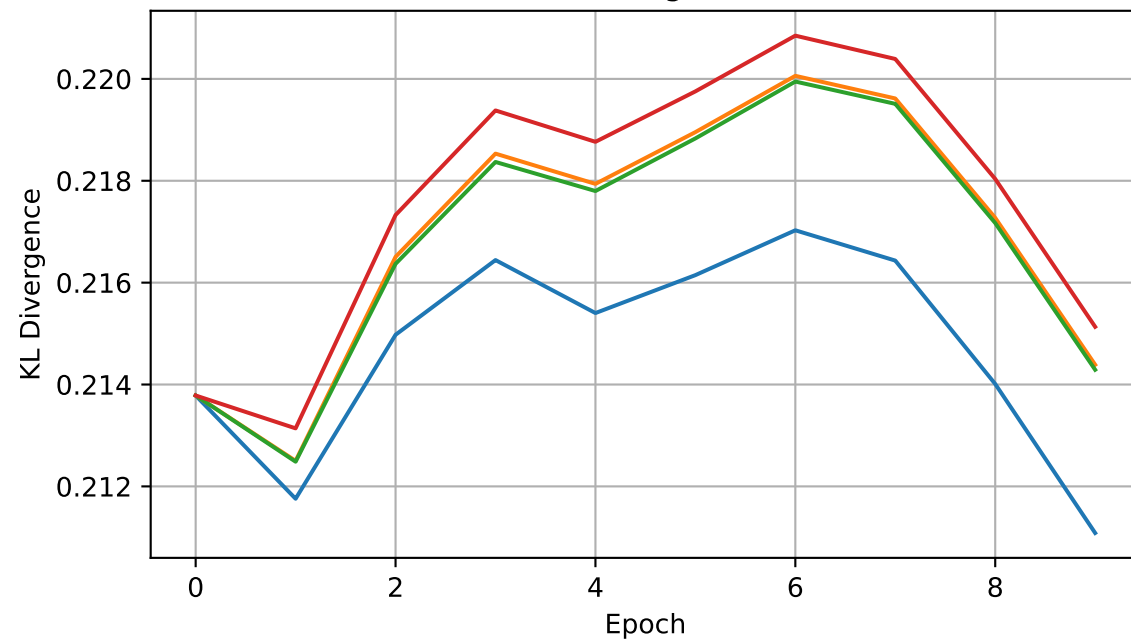
Loss



MSE



KL Divergence



β

