

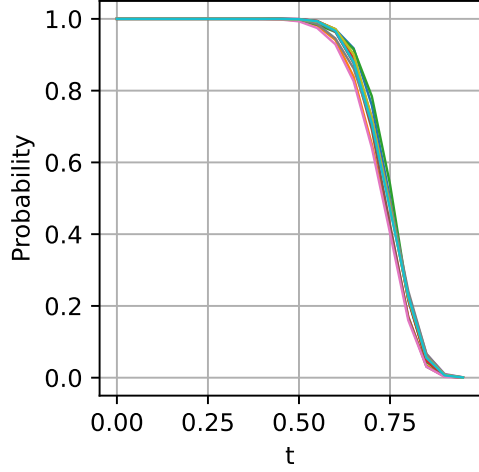


Figure 1. For reviewer **YNcT**. Decay rate for $\langle \mathbf{X}_i, \mathbf{X}_j \rangle$.

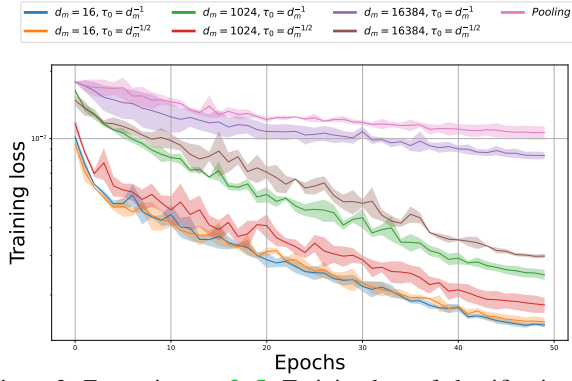


Figure 2. For reviewer **p3n5**. Training loss of classification task (with cross-entropy loss) on MNIST dataset.

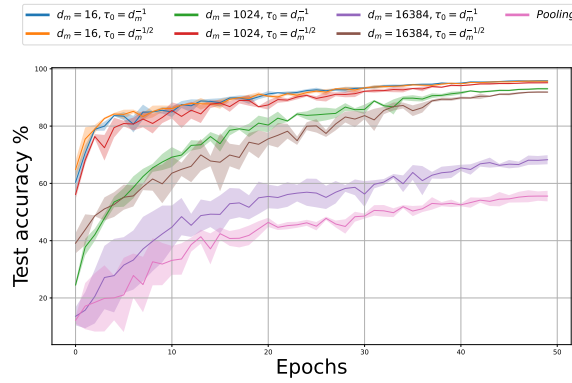


Figure 3. For reviewer **p3n5**. Test accuracy of classification task on MNIST dataset.

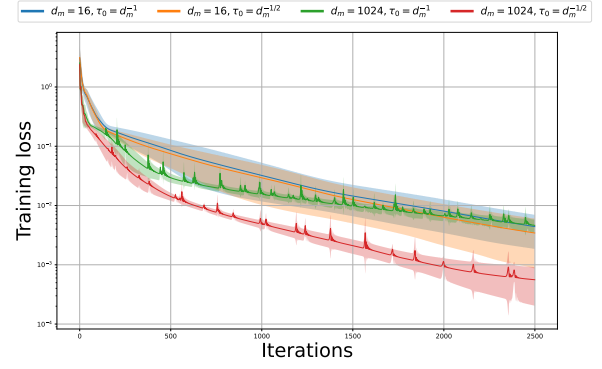


Figure 4. For reviewer **p3n5**. Regression task (with MSE loss) on MNIST dataset.

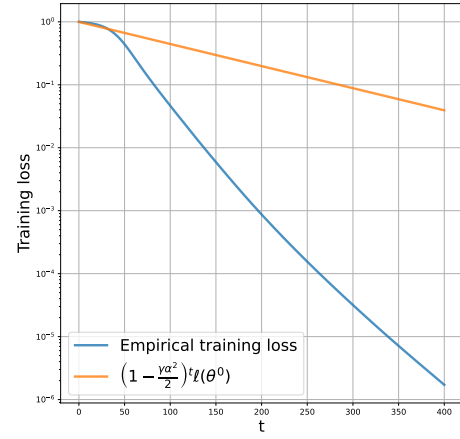


Figure 5. For reviewer **9ERM**. The comparison between the empirical convergence and the quality $(1 - \frac{\gamma\alpha^2}{2})^t \ell(\theta^0)$ in the experiment ($d_m = 100, \tau_0 = d_m^{-1/2}$) at sec.5.1.