

Training loss for different choices of wd

