Train and test negative log likelihood Train NLL Test NLL 2.2 2.0 1.8 1.6 1.4 1.2 250 500 750 1000 1250 1500