Marginal log likelihood wrt iterations train_mll valid_mll test_mll -3.996-3.999-4.002-4.005-4.008100 200 300 iter