Marginal log likelihood wrt iterations train_mll valid_mll test_mll -3.995-4.000-4.005-4.010-4.015100 200 300 400 500 iter