Double Descent LLA KFAC p = 0.1Classical Regime Modern Regime Variance-Bias Tradeoff Larger is Better 3 Train/Test loss Critical Region Interpolation Threshhold 0 0.2 0.0 0.4 0.6 0.8 1.0 1.2 1.4 $\times 10^{6}$ Parameters Train loss Gibbs loss Log Marginal Laplace Bayes loss Bound Log Marginal Upper Bound