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