



Figure 1: Out-of-distribution (OOD) performance of the ICL method in GLM Scenario 2. The x -axis shows the distribution shift between training and test distributions, quantified by C2ST. The y -axis displays the performance of the in-context learner, evaluated via C2ST, MMD, and \mathcal{W}_2 distances against HMC samples — where higher values indicate worse performance. OOD data is generated by gradually increasing the variance of the prior on the regression coefficients from an initial value of 1.0 to $\sqrt{2}$, 1.7, 2, 2.5, and 3.5.