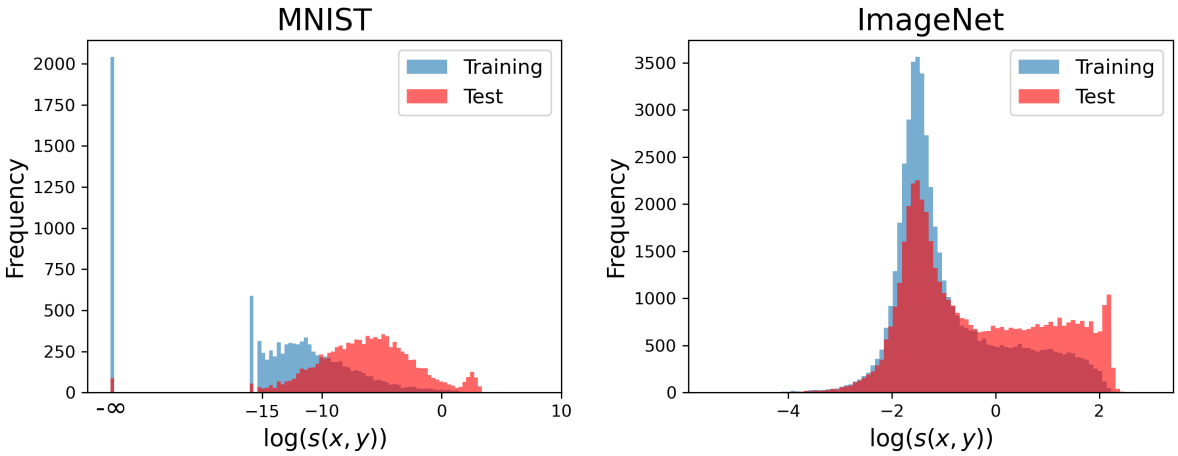

Figure 1: **Visualization of distribution shift in score space.** Score distribution shift plots on MNIST and ImageNet under (corruption ratio $p = 0.05$, noise level $u = 1.0$) perturbation in the data space. The score distribution obtained from the unperturbed data (red), and from the perturbed data (blue) are plotted in log scale. For ImageNet, we removed 18 negative-valued outliers ranging from -5.5 to -10 for visualization purposes.

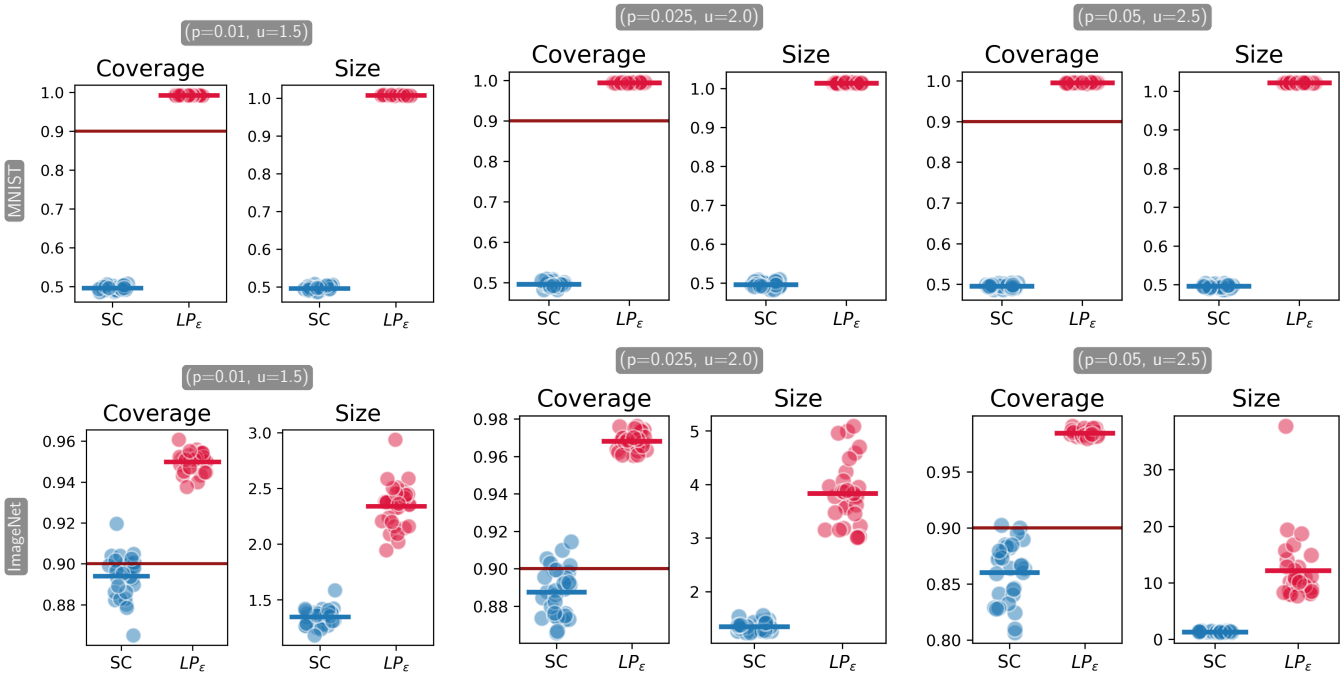

Figure 2: **Score-space distribution shift validity and efficiency.** We directly perturb the scores via noise of level u and corruption with ratio p as described in the paper. Desired $1 - \alpha$ coverage (long dark red line); empirical coverage and prediction set size for each split (scattered points); and mean coverage and prediction set size across 30 calibration-test splits (short colored horizontal lines) are plotted.