



*Figure 2. Visualization of decision boundaries of  $\ell_\infty$  adversarially trained models on 2D box example. The training datasets are randomly sampled from the same data distribution.*