Trained SAEs Barely Exceed Random Baseline

Y-axis: **PWMCC (Feature Overlap)**

Trained SAEs: **0.309±0.002**

Random SAEs: **0.300±0.001**

p = 9.13e-

- - - Random Baseline (0.30)