

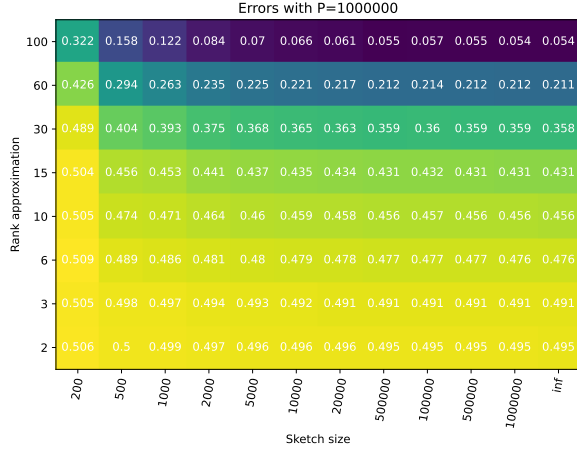


Figure 1: Number of parameter $p = 10^6$, true rank $R = 100$. Error (lower is better) of sketched projection norm (Equation 10) from ground truth projection norm (Equation 9) averaged over 100 random test points. Different values of s (x-axis) and r (y-axis) shown in log-log scale. For a fixed memory $s \cdot r$ sketching noise is outweighed by a higher-rank.

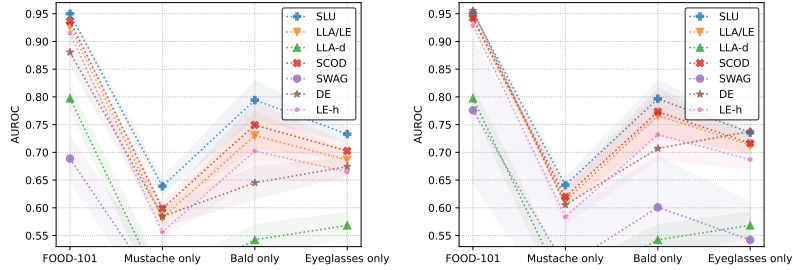


Figure 2: Vision Transformer with $4M$ parameter trained on CelebA. SLU and baselines AUROC performance with fixed memory budget of $3p$ (left) and $10p$ (right) against 4 OoD datasets.

	BUDGET $3p$		BUDGET $10p$	
	FOOD-101	Hold-out (avg)	FOOD-101	Hold-out (avg)
SLU (us)	0.95 ± 0.003	0.72 ± 0.02	0.95 ± 0.003	0.72 ± 0.02
LLA	0.93 ± 0.001	0.67 ± 0.02	0.94 ± 0.001	0.70 ± 0.02
LLA-D	0.80 ± 0.02	0.52 ± 0.03	0.80 ± 0.02	0.52 ± 0.03
LE	0.93 ± 0.001	0.67 ± 0.02	0.94 ± 0.001	0.70 ± 0.02
LE-H	0.91 ± 0.002	0.64 ± 0.03	0.93 ± 0.002	0.67 ± 0.02
SCOD	0.94 ± 0.00	0.68 ± 0.02	0.94 ± 0.002	0.70 ± 0.02
SWAG	0.69 ± 0.04	0.46 ± 0.04	0.78 ± 0.13	0.55 ± 0.09
DE	0.88 ± 0.02	0.63 ± 0.02	0.96	0.68

Table 1: AUROC scores of Sketched Lanczos Uncertainty vs baselines on CelebA. Error bars over 3 random seeds. DE in $10p$ -budget error bars requires 30 models.