	Method	Dataset					
		ASCADv1 (fixed)	ASCADv1 (random)	DPAv4 (Zaid version)	AES-HD	OTiAiT	OTP
	Random	111.6 ± 0.3	108 ± 5	14 ± 2	$\boxed{127 \pm 1}$	1.21 ± 0.04	1.05 ± 0.02
First-order parametric methods	SNR	117.2 ± 0.6	116.7 ± 0.7	11.4 ± 0.2	$\boxed{126 \pm 2}$	$\boxed{1.10 \pm 0.02}$	1.0125 ± 0.0007
	SOSD	114.9 ± 0.5	105 ± 2	$\boxed{8.04 \pm 0.08}$	$\boxed{126 \pm 2}$	1.14 ± 0.03	1.027 ± 0.002
	CPA	111.5 ± 0.4	114 ± 1	11.5 ± 0.3	$\boxed{126 \pm 2}$	1.49 ± 0.04	1.0125 ± 0.0007
Neural net attribution	GradVis	$\boxed{\underline{107.0 \pm 0.5}}$	$\boxed{\underline{95 \pm 2}}$	12.1 ± 0.3	$\boxed{\underline{127 \pm 1}}$	1.4 ± 0.2	1.0142 ± 0.0008
	Saliency	$\boxed{\underline{107.1 \pm 0.5}}$	$\boxed{\underline{95 \pm 2}}$	11.8 ± 0.3	$\boxed{\underline{127 \pm 1}}$	1.39 ± 0.04	1.014 ± 0.001
	Input $*$ Grad	$\boxed{\underline{107.2 \pm 0.5}}$	$\boxed{\underline{95 \pm 2}}$	11.8 ± 0.4	$\boxed{\underline{127 \pm 1}}$	1.36 ± 0.04	1.0141 ± 0.0009
	LRP	$\boxed{\underline{107.2 \pm 0.5}}$	$\boxed{\underline{95 \pm 2}}$	11.8 ± 0.4	$\boxed{\underline{127 \pm 1}}$	1.36 ± 0.04	1.0141 ± 0.0009
	1-Occlusion	$\boxed{\underline{107.1 \pm 0.5}}$	$\boxed{\underline{95 \pm 2}}$	10.1 ± 0.2	$\boxed{\underline{127 \pm 1}}$	1.36 ± 0.04	1.0141 ± 0.0009
	m-Occlusion†	$\boxed{\underline{107.4 \pm 0.4}}$	$\boxed{\underline{94 \pm 2}}$	18.0 ± 0.8	$\boxed{\underline{127 \pm 1}}$	1.39 ± 0.03	1.014 ± 0.002
	2^{nd}-order 1-Occlusion	$\boxed{\underline{107.0 \pm 0.4}}$	$\boxed{\underline{95 \pm 2}}$	10.0 ± 0.2	$\boxed{\underline{127 \pm 1}}$	1.34 ± 0.04	1.0138 ± 0.0008
	OccPOI	111 ± 1	105 ± 4	11 ± 2	$\boxed{\underline{127 \pm 1}}$	$\underline{1.14 \pm 0.03}$	1.043 ± 0.008
	OccPOI (Released result)	110.0 ± 0.7	104 ± 3	n/a	$\boxed{\underline{127 \pm 1}}$	n/a	n/a
	OccPOI-Extended*	109.7 ± 0.4	104 ± 1	$\boxed{\underline{9 \pm 2}}$	$\boxed{\underline{127 \pm 1}}$	$\underline{1.13 \pm 0.04}$	$\boxed{\underline{1.011 \pm 0.001}}$
	GradVis (ZaidNet)	108.8 ± 0.8	n/a	9.3 ± 0.2	$\boxed{\underline{126 \pm 2}}$	n/a	n/a
	Saliency (ZaidNet)	108.8 ± 0.8	n/a	9.3 ± 0.2	$\boxed{\underline{126 \pm 2}}$	n/a	n/a
	Input $*$ Grad (ZaidNet)	109.0 ± 0.6	n/a	9.2 ± 0.2	$\boxed{\underline{126 \pm 2}}$	n/a	n/a
	1-Occlusion (ZaidNet)	109.3 ± 0.6	n/a	9.2 ± 0.2	$\boxed{\underline{126 \pm 2}}$	n/a	n/a
	2^{nd}-order 1-Occlusion (ZaidNet)	108.8 ± 0.4	n/a	9.1 ± 0.2	$\boxed{\underline{126 \pm 2}}$	n/a	n/a
	OccPOI (ZaidNet)	110 ± 1	n/a	10 ± 4	$\boxed{\underline{127 \pm 2}}$	n/a	n/a
	OccPOI-Extended* (ZaidNet)	108.9 ± 0.6	n/a	$\boxed{\underline{8.5 \pm 0.6}}$	$\boxed{\underline{127 \pm 1}}$	n/a	n/a
	GradVis (WoutersNet)	109.9 ± 0.5	n/a	9.6 ± 0.3	$\boxed{\underline{126 \pm 2}}$	n/a	n/a
	Saliency (WoutersNet)	109.8 ± 0.5	n/a	9.6 ± 0.3	$\boxed{\underline{126 \pm 2}}$	n/a	n/a
	Input $*$ Grad (WoutersNet)	109.7 ± 0.5	n/a	9.4 ± 0.3	$\boxed{\underline{126 \pm 2}}$	n/a	n/a
	1-Occlusion (WoutersNet)	109.7 ± 0.4	n/a	9.4 ± 0.3	$\boxed{\underline{126 \pm 2}}$	n/a	n/a
	2^{nd}-order 1-Occlusion (WoutersNet)	109.2 ± 0.2	n/a	9.2 ± 0.2	$\boxed{\underline{126 \pm 2}}$	n/a	n/a
	OccPOI (WoutersNet)	109 ± 1	n/a	13 ± 9	$\boxed{\underline{127 \pm 1}}$	n/a	n/a
	OccPOI-Extended* (WoutersNet)	109 ± 1	n/a	$\boxed{\underline{8 \pm 1}}$	$\boxed{\underline{127 \pm 2}}$	n/a	n/a
	ALL (ours)	$\boxed{\underline{107.5 \pm 0.3}}$	101 ± 2	12.2 ± 0.4	$\boxed{\underline{126 \pm 2}}$	1.23 ± 0.03	1.0161 ± 0.0009

Table 1: Performance of leakage localization algorithms according to the Fwd-DNNO (forward DNN occlusion) test (**smaller is better**). To compute this metric, we first train a supervised DNN classifier to map emission traces to the sensitive variable. We then occlude all its inputs and incrementally un-occlude them *from most- to least-leaky* (the opposite order as for the Rev-DNNO test) as estimated by the method under test, and at each step compute its performance (quantified by rank, lower is better) on the test dataset. The Fwd-DNNO metric is given by the average value of these performance assessments (lower is better, because it indicates that claimed leaky features indeed had utility to the classifier). Of the two DNN occlusion metrics, this is more sensitive to *true/false positive* leakiness measurements because the performance of the classifier tends to jump and stay up as soon as it sees leaky measurements. Best result is boxed and best deep learning result is underlined. Results are reported as mean $\pm$ std. dev. over 5 random seeds. This metric appears to have high variance and little discriminative power compared to the oSNR and Rev-DNNO metrics (as indicated by the large number of tied 'best' methods), and there is no clear best method according to Fwd-DNNO. *The algorithm as proposed in (Yap, 2025) is very time-consuming because it requires $O(T^2)$ non-parallelizable passes through the dataset. Due to time constraints, for ASCADv1 (random) and AES-HD we stop it early after $10\times$ the runtime of ALL. †For each dataset, we use m in `np.arange(1, 51, 2)` which maximizes the oSNR metric for m-occlusion. Specifically, we use $m = 3$ for ASCADv1 (fixed), $m = 7$ for ASCADv1 (random), $m = 43$ for DPAv4, $m = 33$ for AES-HD, $m = 3$ for OTiAiT, and $m = 7$ for OTP.