Dataset	MNIST			FMNIST			CIFAR10		
ϵ	0.25	0.50	1.00	0.25	0.50	1.00	0.25	0.50	1.00
SH+SG (Us)	**9.6e-7**	**7.8e-6**	**6.7e-5**	**6.5e-7**	**4.0e-6**	**4.3e-5**	**9.8e-4**	**2.2e-2**	**1.2e-1**
SG	4.5e-6	4.1e-5	3.9e-4	3.0e-6	2.7e-5	2.6e-4	1.3e-2	8.6e-2	4.9e-1
SP (H + G)	1.2e-6	9.6e-6	8.1e-5	9.6e-7	7.5e-6	6.5e-5	2.1e-3	3.3e-2	2.5e-1
SP G	4.6e-6	4.1e-5	3.9e-4	3.2e-6	2.8e-5	2.6e-4	1.3e-2	9.0e-2	5.2e-1
Swish (H + G)	2.4e-6	2.0e-5	1.9e-4	2.1e-6	1.7e-5	1.8e-4	1.1e-2	3.3e-1	7.9e0
Swish G	5.6e-6	5.0e-5	4.9e-4	3.9e-6	3.5e-5	3.5e-4	4.9e-2	9.8e-2	6.0e-1
G	4.2e-3	1.7e-2	6.7e-2	3.8e-3	1.5e-2	6.0e-2	3.0e-1	1.2e-0	5.0e-0

Table 1: The average $\mathcal{P}_{MSE}$ achieved by first and second order methods are compared as a proxy for explainer quality. The predicted class logit is modeled for three datasets: MNIST, FMNIST and CIFAR10 at three radii each. A fully connected network is used for MNIST and FMNIST, wheras a ResNet18 was trained on CIFAR10. *We include new results for Swish smoothed ReLU networks, Swish Hessian + Gradient (Swish (H + G)) and Swish Gradient (Swish G).* Other methods include: SmoothHess + SmoothGrad (SH+SG,Us), SmoothGrad (SG), SoftPlus Gradient (SP G), SoftPlus Hessian + SoftPlus Gradient (SP (H + G)), and the Vanilla Gradient (G). *Our method SH + SG achieves the lowest $\mathcal{P}_{MSE}$ for each permutation of dataset and radii ε.*

ϵ	0.25	0.50	1.00
SH+SG (Us)	**1.3e-4**	**9.5e-4**	**6.9e-3**
SG	5.3e-4	4.3e-3	3.3e-2
G	3.8e0	1.5e+1	6.0e+1

Table 2: Using a ResNet101 trained on CIFAR10, the average $\mathcal{P}_{MSE}$ achieved by SmoothHess + SmoothGrad (SH + SG, Us), SmoothGrad (SG) and Vanilla Gradient (G), is evaluated as a proxy for explainer quality. Results are reported at three radii ε, for the predicted class logit. *Our method, SH + SG, achieves the lowest $\mathcal{P}_{MSE}$ for each of the three radii ε.* This indicates that the interactions SmoothHess captures improve the model of network behaviour over first-order approaches, *even for large networks such as ResNet101.* The camera ready version of Table 2 will contain additional results for SoftPlus and Swish, which we were not able evaluate for the rebuttal due to their expensive validation procedures described in the global response (G4).