Attack: None 0.50 0.25 0.00 Attack: FGSM targeting Ensemble Attack: PGD targeting Ensemble 1.00 1.00 0.50 0.50 0.25 0.25 0.00 0.00 30 35 40 20 35 45 50 20 30 50 52 Attack: FGSM targeting layers [20, 30, 35] Attack: PGD targeting layers [20, 30, 35] 1.00 1.00 0.50 0.25 0.25 0.00 30 Attack: FGSM targeting layer 20 Attack: PGD targeting layer 20 1.00 1.00 0.25 0.25 0.00 0.00 30 50 30 Attack: FGSM targeting layer 30 Attack: PGD targeting layer 30 1.00 1.00 0.50 0.50 0.25 0.00 20 30 35 20 35 50 52 Attack: PGD targeting layer 35 Attack: FGSM targeting layer 35 0.50 0.50 0.25 0.00 Attack: PGD targeting layer 40 Attack: FGSM targeting layer 40 1.00 0.50 0.25 0.25 0.00 0.00 40 30 35 45 50 20 30 35 50 52 Attack: FGSM targeting layer 45 Attack: PGD targeting layer 45 0.50 0.50 0.25 0.25 0.00 0.00 35 50 35 Attack: FGSM targeting layer 50 Attack: PGD targeting layer 50 1.00 0.50 0.50 0.25 0.00 0.00 35 40 45 50 52 52 Attack: FGSM targeting layer 52 Attack: PGD targeting layer 52 0.50 0.50 0.25 0.00 20 30 35 45 50 52 30 35 40 52

Dataset: cifar10

Accuracy