Experiments: Backdoor

Setting 1: Reproduce the setting considered by Sun et al.

Setting 2: Consider an increasing fraction of malicious participants, aiming to show how effective defenses are against varying numbers of attackers.

Compute both the main task and backdoor accuracy

Apply: Norm Bounding, Weak DP, CDP, LDP
(on all participants, on non-attackers)

55

Experiments: Backdoor

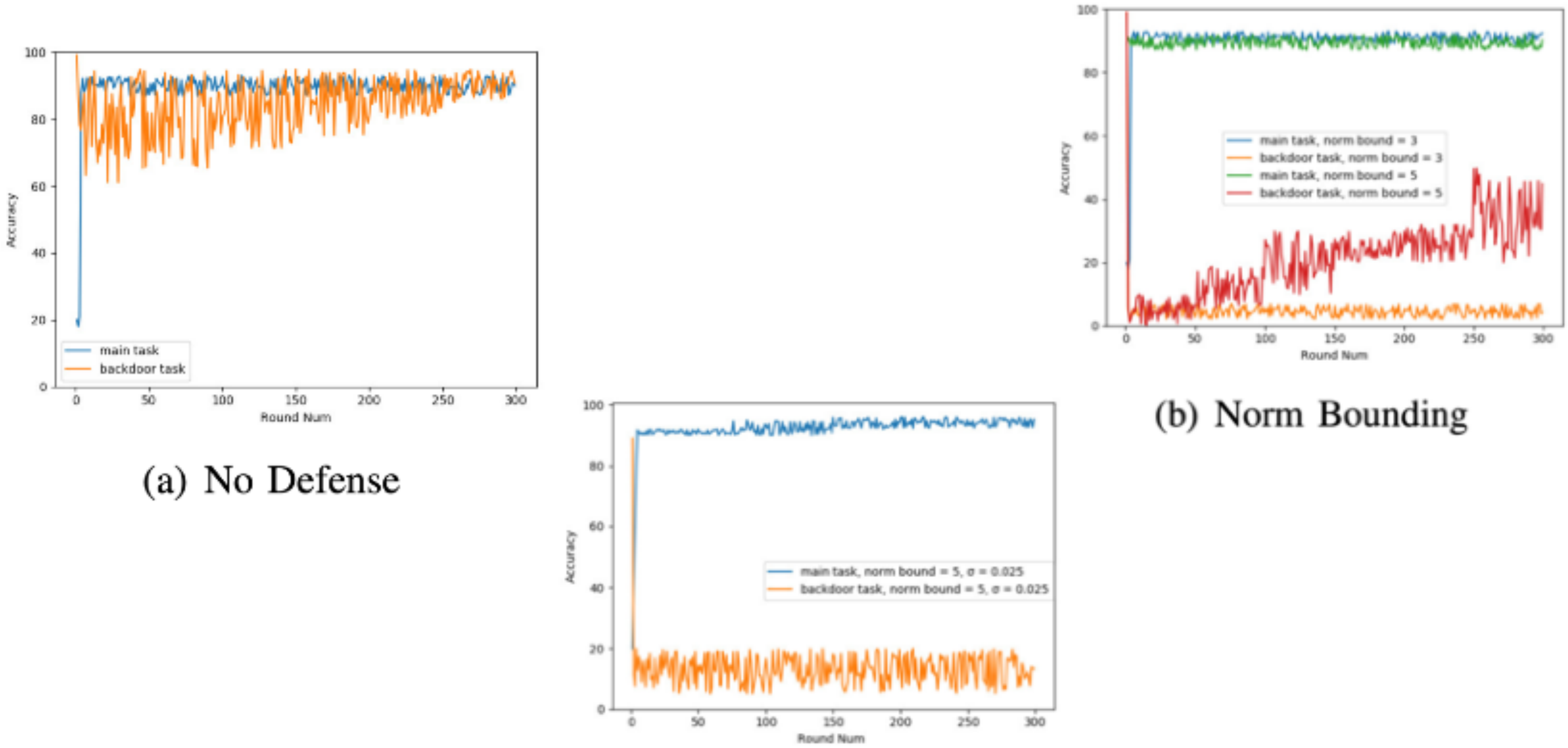
Setting 1: Reproduce the setting considered by Sun et al.

Setting 2: Consider an increasing fraction of malicious participants, aiming to show how effective defenses are against varying numbers of attackers.

Compute both the main task and backdoor accuracy

Apply: Norm Bounding, Weak DP, CDP, LDP
(on all participants, on non-attackers)

Setting 1 (on EMNIST)



(a) No Defense

(b) Norm Bounding

(c) Weak DP