

A FedNano: Pseudocode Overview

Algorithm 1 *FedNano*. The K clients are indexed by k ; R is the number of communication rounds, and T is the number of local steps.

Server Update:

- 1: Randomly initialize \mathcal{A}_I^0 and \mathcal{A}_T^0 in *NanoAdapter*, and distribute to clients
- 2: **for** $r = 1$ to R **do**
- 3: **for** $k = 1$ to K **in parallel do**
- 4: $\theta_k^r \leftarrow \text{ClientUpdate}(\theta_{global}^{r-1}, D_k)$
- 5: Compute FIM F_k
- 6: **end for**
- 7: $\theta_{global}^r \leftarrow \text{ServerAgg}(\{\theta_k^r, F_k^r\})$ ▷ Eq. 1
- 8: **end for**

Client Update:

- 1: $\theta_k^{r-1} \leftarrow \theta_{global}^{r-1}$
 - 2: **for** local step $t = 1$ to T **do**
 - 3: Sample $\{(v_k, q_k, a_k)\}$ from D_k
 - 4: $\theta_k^{r(t)} \leftarrow \text{Optimization}(\theta_k^{r(t-1)}, v_k, q_k, a_k)$
 - 5: **end for**
 - 6: **return** θ_k^r
-