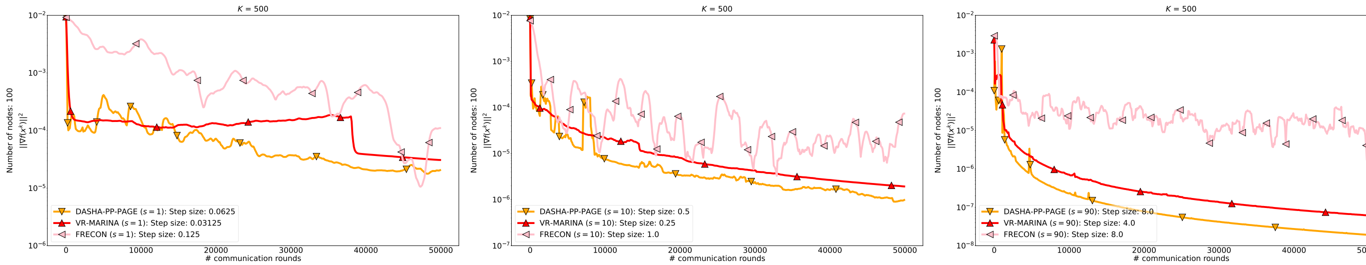
Extra Experiments

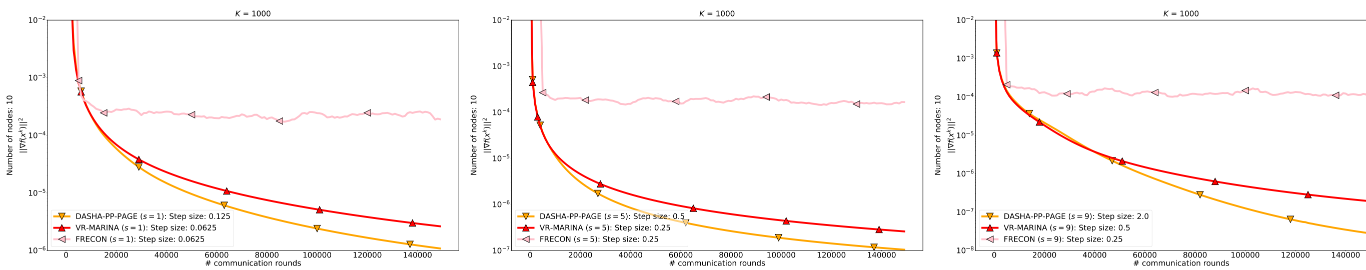
In this experiments, we compare our new algorithm DASHA-PP with previous baselines MARINA and FRECON in the partial participation setting (the comparisons of DASHA-PP with DASHA we present in Section A of the paper). We consider MARINA and FRECON because they are the previous SOTA methods in the *partial participation setting with* **compression**. We investigate the same optimization problem and setup as in Section A of the paper. One can find the details (loss functions, exact parameters) in Section A. All methods use the Rand$K$ compressor in these experiments.

1. **Finite-Sum Setting.** In Figures 1 and 2, we compare all three methods in the finite-sum setting (we use the function from Lines 408-409 of the paper) on two different datasets: *real-sim* and *MNIST*. The parameter s is the number of clients participating in each round that are selected randomly using the s-nice sampling (server chooses uniformly s nodes without replacement). We can see that DASHA-PP converges faster than MARINA. Since FRECON does not support variance reduction of stochastic gradients, it converges to less accurate solutions.



(a) 1 % of nodes participating (b) 10 % of nodes participating (c) 90 % of nodes participating

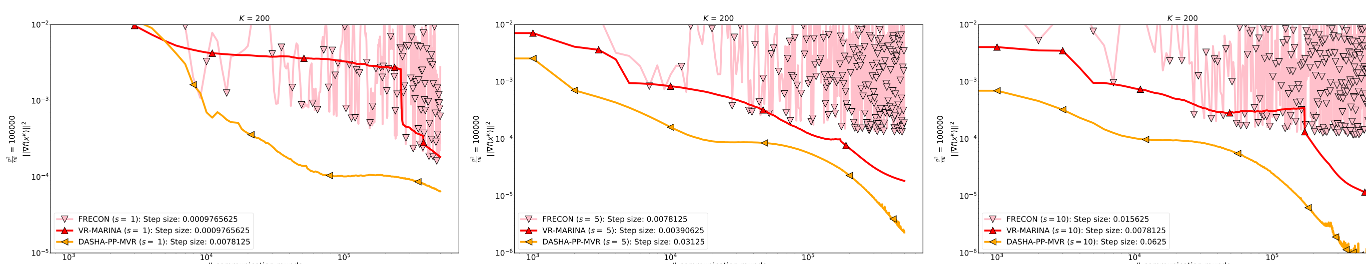
Figure 1: Classification task on *real-sim*



(a) 10 % of nodes participating (b) 50 % of nodes participating (c) 90 % of nodes participating

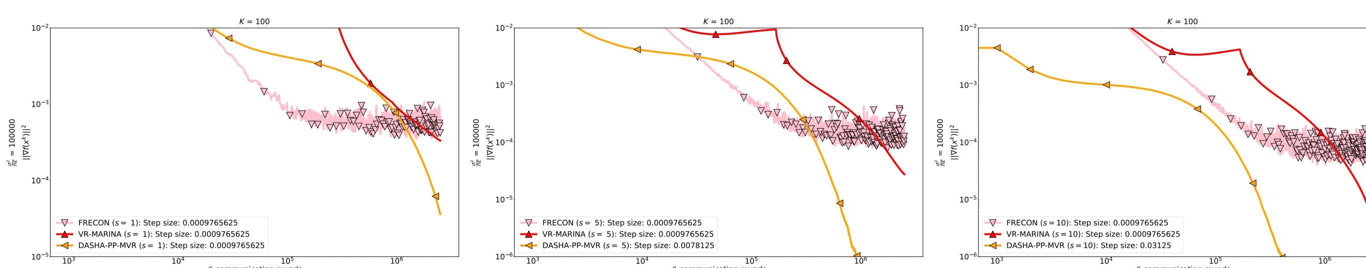
Figure 2: Classification task on *MNIST*

2. **Stochastic Setting.** In Figures 3 and 4, we consider the stochastic setting (we use the function from Lines 411-412 of the paper). We can see that DASHA-PP convergences to high accuracy solutions, unlike FRECON. Moreover, DASHA-PP improves the convergence rates of MARINA.



(a) 10 % of nodes participating (b) 50 % of nodes participating (c) 100 % of nodes participating

Figure 3: Classification task on *real-sim*



(a) 10 % of nodes participating (b) 50 % of nodes participating (c) 100 % of nodes participating

Figure 4: Classification task on *MNIST*