

Extra Experiments

In this experiments, we compare our new algorithm **DASHA-PP** with previous baselines **MARINA** and **FRECON** in the partial participation setting. We consider **MARINA** and **FRECON** because they are the previous SOTA methods in the *partial participation setting with compression*. We consider the same optimization problem and setup as in Section A of the paper.

Stochastic Setting

In Figure 1, we compare all three methods in the stochastic setting on two different datasets: *real-sim* and *MNIST*. The parameter s means the number of clients participating in each round. We can see that **DASHA-PP** has better convergence rates than **FRECON**. **DASHA-PP** with $s \in \{5, 10\}$ even improves **MARINA** that works in the full participation regime (one should expect that convergence rate of **MARINA** with partial participation will be even worse).

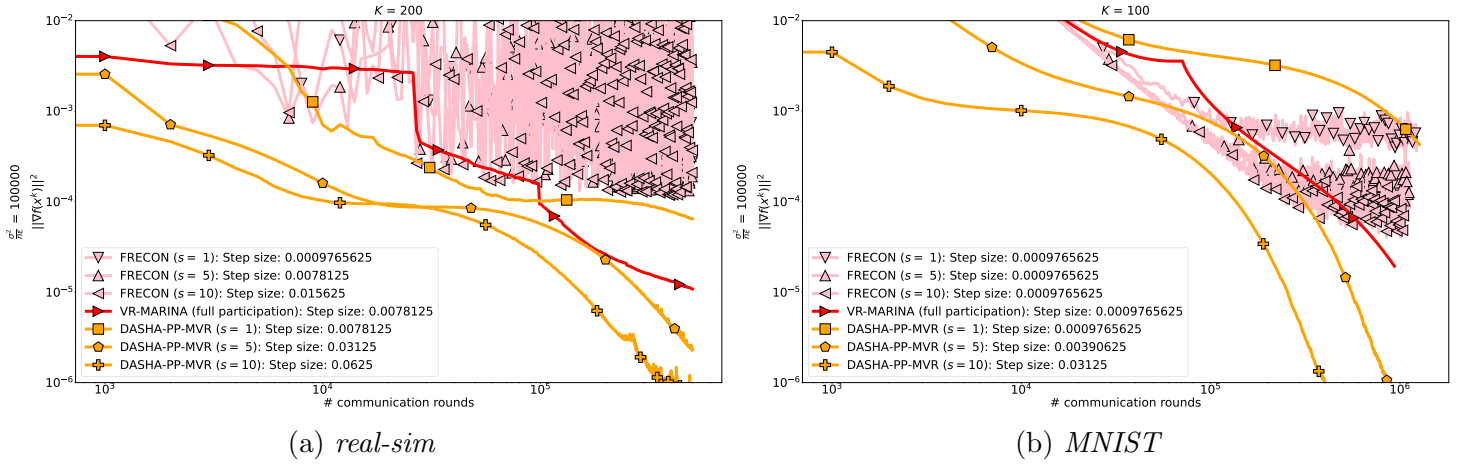


Figure 1: Classification task with the *mushrooms* dataset.

Finite-Sum Setting

In Figure 2, we consider the finite-sum setting. We do not have experiments with **FRECON** because the authors of **FRECON** do not consider it in the finite-sum setting. We can see that **DASHA-PP** with $s \in \{50, 90, 100\}$ converges faster than **MARINA** (that works in the full participation regime).

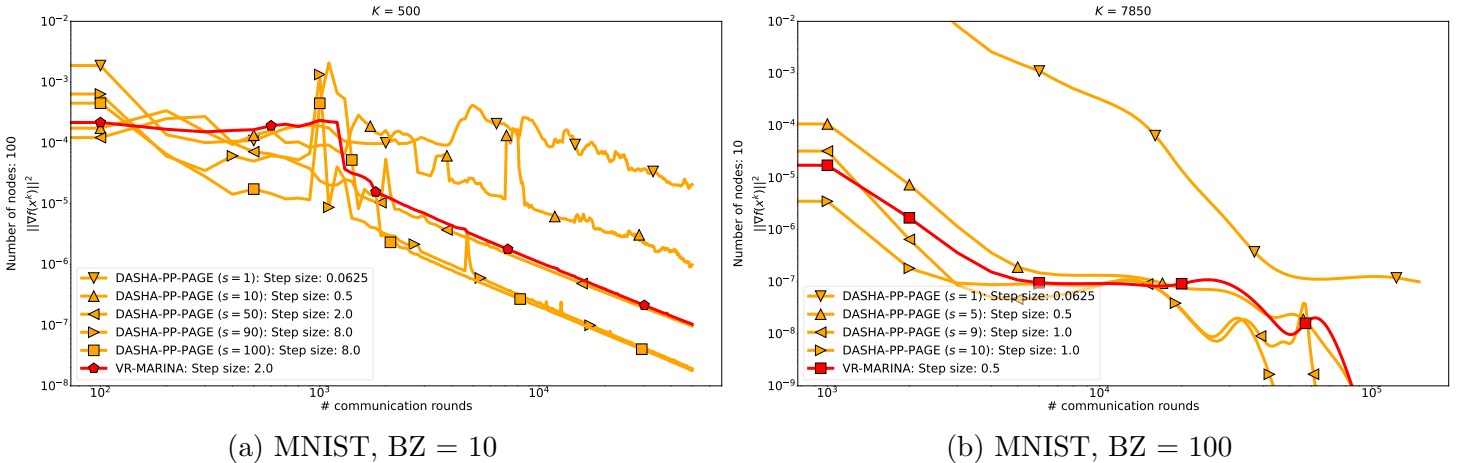


Figure 2: Classification task with the *mushrooms* dataset.