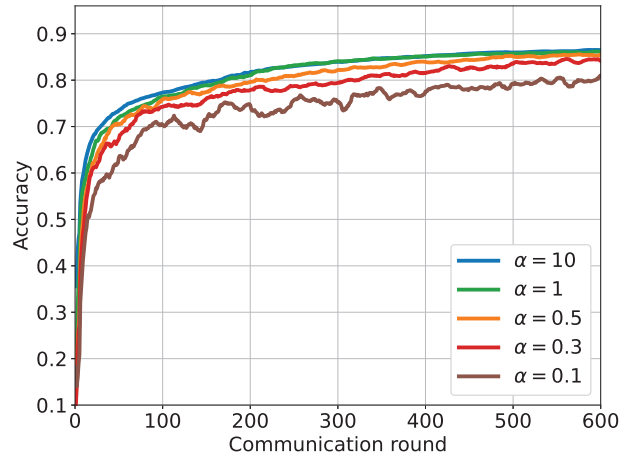


(a) CIFAR-10 dataset



(b) FMNIST dataset

Figure 7: Test accuracy versus communication round under different data heterogeneity levels on CIFAR-10/FMNIST dataset. Note that as data heterogeneity is simulated by a Dirichlet distribution $Dir(\alpha)$, a smaller α corresponds to a higher degree of non-i.i.d. distribution of the dataset.