We sincerely thank the four reviewers for their valuable feedback.

Reviewer 1: We thank you for the useful comments:

Privacy:

Comparison with PRIVIX:

Experiments:

Reviewer 3: We thank the reviewer for the interest in our contribution:

Convergence Bounds:

Reviewer 4: We thank the reviewer for the thorough analysis. Our remarks are listed below:

Privacy: By privacy we mean that the adversary cannot get the exact data (as opposed to standard Federated Learning), since it is hidden in the random sketches. We will make it more clear in the revision.

Convergence Bounds:

Comparison with FedSGD: We stress on the observable gap between our method and FedSGD in the numerical runs. FedSGD is a method using the full gradient at each round of communication and thus displaying a higher computation cost than any other methods using sketches that we plot. While (50×100) may seem large, it still represents and $12\times$ compressing ratio, which is considerable. Under such communication reduction, for (50×100) sketch size, the test accuracy is very close (if not identical) to FedSGD in the bottom two figures in Figure 1 and 2. Thus, we believe our results validates the benefit of the proposed methods in practice.

Reviewer 8: We thank the reviewer for valuable comments. Our response is as follows:

Originality of our contribution: Our algorithmic contribution stands as a combination of two previous works. In [Ivkin et al. 2019], only the top-K coordinates (heavy hitters) are recovered, while in [Li et al. 2019], the whole model is compressed without specifically addressing the coordinates with largest magnitude. The HEAPRIX method combines the best of both worlds and we believe that it constitutes a novelty and a technical progress in federated learning design. A consequence of that novelty, a faster convergence is achieved with better empirical performance as displayed in our contribution. We stress on the fact that our methods can be used with any other sketching or compression techniques in-lieu of the HEAPRIX operation, yet no guarantees are provided that such resulting algorithm will have the same convergence behaviour. **The idea behind our contribution is to both leverage the sketching technique for privacy purpose and the unbiasedness of the operation for convergence purpose.**