

---

We thank the five reviewers for their valuable feedback. We first discuss concerns shared by some reviewers:

**Reviewer 3, Reviewer 7 and Reviewer 8:**

– *On :*

**Reviewer 3:**

– *Various Notations and Typos:*

2) 4) We referred to Lemma 1 indeed. 5) 6) topk alone is biased which is what we want to avoid with HEAVYMIX. 7) We chose to encapsulate the update allowing to compute the sketch of the gradient vector. This allows to use any sketching method as a plug-and-play update. 8) HEAVYMIX does require to extract the exact  $Top_m$  values of the gradient. 9) See 7) 11) The function  $S_j(\cdot)$  is any sketching operator that depends on the random hash tables used. This will be clarified. 12) We will add a table to summarize the metrics from the plots.

**Reviewer 7:** We will proofread the paper and add the following remarks:

– *Privacy:* HEAPRIX sketches are based on random hashes that by construction ensures the attacker to access exact values of the gradients.

– *Compressing and Sketching:* Thank you for the references. Compression and quantization are two alternatives to the overall objective of making algorithms communication efficient. Though, our paper focuses on sketching techniques in the spirit of tackling the privacy issue of FL along its communication efficiency bottleneck.

– *Numerical Experiments:* 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  $(50 \times 100)$  may seem large, it still represents and  $12\times$  compressing ratio, which is high. Hence, not matching FedSGD performances while reducing by 12 the size of the transmitted vector is not an issue.

**Reviewer 8:**

– *Numerical Experiments:* We have developed numerical experiments on both MNIST and CIFAR-10 under iid and non-iid settings using shallow and deeper neural networks for completeness. Those runs are also reported in the supplementary material. More experiments are in progress for inclusion in the revised paper.

– *Comparison with other unbiased compressors:*

– *Heuristic behind FedSKETCH:*