AE report on AOS2309-067 “Ensemble linear interpolators: The role of ensembling”

The paper considers ensemble linear interpolators with a specific focus on understanding the role of ensembling under the setting of the linear regression model. The full sample min-norm least square estimator is used as the benchmark. This study is closely related to the earlier work on ”Sketched ridgeless linear regression: The role of downsampling” by Chen, Zeng, Yang, and Sun (2023), which shares one of the authors with the current paper. Both papers concentrate on assessing the out-of-sample prediction risks of the sketched ridgeless least square estimator. Two differences are 1) the current paper also adds the additional component of multiplier-bootstrap-based bagging, and 2) Chen et al. (2023) primarily focuses on dense sketching matrices, while this paper considers diagonal and mostly singular sketching matrices.

This paper bears a strong resemblance to Chen et al.’s (2023) work. The model settings and the interpolators are the same. Both extensively establish the limiting out-of-sample prediction risk for sketching interpolators. Both papers initially explore the isotropic case before extending the analyses to the correlated case. The current paper introduces additional subsections to extend the results to the bagging estimator. Given the similarity, it is unclear to me whether the results and proofs in this paper are straightforward extensions of those presented in Chen et al. (2023).

A collection of results were presented, but most of them are not individually significant. Since there are so many such results, the paper seems to be lack of focus. The paper’s title and writing suggest a focus on the understanding of the role of bagging. However, the revealed message fails to offer much novelty. It is well understood that bagging can reduce variance. The bias reduction discovery seems to be new. I acknowledge that some intricate analyses may be involved in precisely determining variance reduction specific to this individual problem, but the extent to which these results provide new insights remains questionable.

I think an original contribution of this paper is that Bernoulli sketching optimizes the limiting out-of-sample prediction risk among the various sketching matrices explored here (diagonal and mostly singular matrices). Since the limiting out-of-sample prediction risk of Bernoulli sketching is the same as that of orthogonal sketching (whose optimality was previously proved by Chen et al. (2023) among the dense sketching matrices), and the computational cost of Bernoulli sketching is much lower than the orthogonal sketching, it is advantageous to use Bernoulli sketching. However, this result alone does not justify the significance of an AOS publication. Another concern of mine is this paper considers very simple model setting and estimator. It is unclear how likely the discoveries here can be generalized to a more realistic setting with more complicated models and estimators. For these reasons, I cannot recommend publication of the paper.