We thank all reviewers for their constructive feedback.

Techniques (Reviewer 2) and comparison to the concurrent work of [DKTZ20] (Reviewer 1):

1. We propose to sequentially minimize a series of "potential functions" $f_{u,b}(w)$ (see Appendix F for definition) for progressively smaller values of b. We believe that our time-varying setting of parameter b in combination with this new potential function is the key to achieve the state-of-the-art $O\left(\mathrm{poly}(\frac{1}{1-2\eta}, \log\frac{1}{\epsilon})\right)$ label complexity for active learning under Massart noise, which has not appeared before.
2. We also believe that our application of online linear optimization in (active) learning halfspaces is quite novel: in order to optimize $f_{u,b}(w)$, we construct an online linear optimization problem whose *negative benchmark performance* is an upper bound of $f_{u,b}(w)$ (Lemma 6)! To the best of our understanding, the work of [DKTZ20] considers optimizing the global expectation of a sigmoid-like loss function (denoted as $L(w)$), and its key observation is that the gradient of the expected loss $\nabla L(w)$ is large whenever w and u have a large angle. We are not sure if there is a connection there (although any connection would be very interesting). To efficiently solve the online linear optimization problem in a sparsity-adaptive manner, we use online mirror descent with a squared ℓ_p-norm regularizer (with p close to 1), which is well-known to promote attribute-efficiency in prior works on online learning.

We will add more discussions and explanations on the above points in the final version.

As Reviewer 3 suggests, we will add a brief introduction to online mirror descent in the final version. The extra hard thresholding step suggested by the reviewer is not necessary (although it would not hurt either - the final error guarantee would only change by a constant factor); we refer the reviewer to the proof of Theorem 4 for more details.

References

[DKTZ20] Learning Halfspaces with Massart Noise Under Structured Distributions, Diakonikolas, Kontonis, Tzamos, Zarifis, arXiv:2002.05632