We thank the reviewers for their careful consideration of our paper and their positive feedback. Below we address individual comments/questions by the reviewers.

Reviewer 1: We thank the reviewer for their positive feedback.

Technical novelty: We would like to point out that "multi-filtering" is an algorithmic framework and not a specific algorithm. As the reviewer notes, this framework was first introduced in a STOC'18 paper by Diakonikolas, Kane, and Stewart. The [DKS18] paper obtained list-decodable mean estimation algorithms for *identity covariance Gaussian distributions*. We emphasize that enhancing this framework to obtain our efficient algorithm (with near-optimal error guarantees) for the broad family of *bounded (and unknown) covariance distributions* requires overcoming a number of obstacles. To do so, we develop new technical tools that were not present in [DKS18] or in the heavy-tailed robust mean estimation algorithms for $\alpha > 1/2$. Please see lines 105-160 of our submission for an overview.

We termed our algorithm "potentially practical", given that it is iterative (with each iteration being fast, running in near-linear time) and based on prior experience regarding the experimental performance of filtering algorithms. We agree with the reviewer that an experimental evaluation of our algorithm would be interesting.

Related work: We will make sure to cite and compare with the contemporaneous work by Cherapanamjeri, Mohanty, and Yau in the final version of our paper. We would like to point out that this work first appeared on the arXiv around a week before the NeurIPS deadline. Briefly, we note that this concurrent work uses very different techniques and achieves runtime $\tilde{O}(nd)/\operatorname{poly}(\alpha)$, for some unspecified degree polynomial in $1/\alpha$ (from our reading, the degree of the $\operatorname{poly}(\alpha)$ dependence seems to be at least six). This runtime is better than that of our algorithm as a function of n, but worse as a function of $1/\alpha$. We note that the runtime dependence on $1/\alpha$ is equally significant in some key applications of list-decodable learning (e.g., in learning mixture models with many components).

Reviewer 2: We thank the reviewer for their positive feedback.

Regarding comparison to contemporaneous related work: Please see the *Related Work* paragraph in the response to Reviewer 1.

Regarding implementability of our algorithm: Our algorithm is simple to implement and we believe it can scale in high dimensions, even for small values of the parameter $\alpha < 1/2$. That said, an experimental evaluation is left for future work. The contribution of our paper is theoretical and boils-down to developing the first algorithm for this fundamental learning problem that avoids the ellipsoid method and has a low-degree polynomial runtime.

Reviewer 3: We thank the reviewer for their positive feedback.

Due to space limitations, in Section 1.1 we provide the necessary background. In particular, we summarize the prior work in algorithmic high-dimensional robust statistics that is most closely related to the results of our paper. We will be happy to add more context in the final version.

The median-of-means based techniques are typically used in a related, but different, context. The prototypical example is when we have i.i.d. samples from a heavy-tailed (bounded covariance) distribution and the goal is to obtain a *high-confidence* estimator of its mean. This notion of robustness is different than the one considered in our paper where the majority of the input dataset consists of adversarial outliers. We will clarify this connection in our final version.