We thank the reviewers for their comments.

Regarding the exponential dependence on $1/\alpha$, two of our colleagues have independently informed us that they can prove SQ lower bounds showing this dependence is necessary. We have encouraged them to write up their proofs as soon as possible.

Regarding the Hainline et al. paper, we thank the reviewer for bringing this paper to our attention and will cite it in the related work section. That said, we are a bit puzzled by the direct comparison to our work. In Hainline et al, they consider the problem of *conditional* linear regression, where the goal is to find a linear function with small square loss *conditioned* on a subset of training points whose 'indices' satisfy some constant-width k-DNF formula.

In our paper, there are simply inliers (points from the true distribution) and outliers, and we give an algorithm that outputs a list of linear functions, one of which has small error with respect to the inliers. We show that this algorithm succeeds if and only if the inliers are drawn from an anti-concentrated distribution. We do not place any 'computational constraints' on the training set (such as satisfiability by DNF formulas).

Finally, the Hainline et al. paper uses a brute force search subroutine over all subsets of the training set (of size 'r') and thus obtain consequences for regression where the coefficient vectors have a constant number of non-zero entries. The point of our work is to avoid brute force search by developing new techniques relying on the sum-of-squares method.