We thank all reviewers for their work. We ask Reviewer #5 to reconsider his/her decision in the light of our feedback.

About our experiments. We agree that experiments on real-world datasets would add value. However, we remark that the main objective of our work is a theoretical investigation of the query complexity of cluster recovery with same-cluster queries. Performing a thorough evaluation on real-world data is a separate task, beyond the scope of this paper. The goal of our experiments is to certify that our algorithms are easy to implement, behave as predicted by the theory, and have no large hidden costs. We believe our empirical setting fulfills these objectives.

Reviewer #2.

Observation about low rank: The statement about linear combinations is correct and is equivalent to a low-rank assumption. A justification is given precisely by matrix factorization, which assumes the set of users (the rows of the user matrix) spans a low-dimensional subspace. Another justification is that often some features (dimensions) are linear combinations of others, and the actual clusters lie in a low dimensional subspace (namely, are contained in degenerate ellipsoids). In any case, we note that the exponential dependence on r is unavoidable, as our lower bounds show.

Other niceness notions: we agree on these future directions, and thank the reviewer for the suggestion.

Reviewer #3.

Low-stretch separators: no, we do not use them — they are a "bonus" theoretical contribution.

The heuristic: we agree and will clarify this.

Reviewer #4.

MVEE and Eq. 1: we agree and will clarify this.

References for k-center etc.: we agree and will add them.

Datasets: we thank the reviewer for the pointer.

Reviewer #5.

"Theorem 1 is difficult to follow" and "What happens when $W = I$?": we believe these comments do not identify "typical weaknesses", since they just concern the way a theorem is introduced and a special case of our results.

Theorem 1 is difficult to follow: we will add more intuition and simplify the statement of the theorem.

What happens when $W = I$?: our algorithms are not adaptive in this sense, so when $W = I$ they do not obey the bounds of [4]. We agree this is an interesting direction. Note however that, in the supplementary material, we show that the bounds of [4] hold when W has low condition number and thus is "close" to I.

We should compare to [3] instead of [4]: we disagree. The goal of [3] is to build a PTAS for k-means. Our goal is to recover a latent clustering. The two problems are incomparable: a good k-means value can be achieved by a clustering very different from the optimal one, and vice versa. And indeed, [3] does not require any margin, whereas our lower bounds clearly say that without margin one needs n queries. This would have implied a contradiction had the two works solved the same problem. Thus, [4] provides a baseline, but [3] does not.

Problem formulation: thanks for the suggestion. We will add a formal definition of the problem addressed in the paper.

Balanced clusters: we do not need any assumption on the cluster sizes. Since there are at most k clusters, there is always a cluster with $\Omega(n/k)$ points, and this is sufficient for our algorithm.

Knowledge of γ: we disagree. Our algorithms require knowledge of γ exactly as the algorithms of [4] do: if the value of γ passed to the algorithm is a lower bound on the actual margin of the instance, then the correctness is guaranteed, otherwise the algorithm is allowed to return an arbitrary clustering. We are not aware of prior work that does not require knowing γ for exact recovery.

The analysis crucially relies on d being a constant: this is not true. Our analysis does not make any assumption on d, and indeed d appears in our bounds. If the reviewer refers to the bounds in the abstract, then, as the abstract says, those are simplified versions of the full bounds.

We should experimentally compare to [30]: we disagree, [30] requires in input a similarity matrix W correlated to the latent clustering, which is not available in our setting.

The paper is not written clearly: we will do our best to clarify all definitions, assumptions, and main results earlier on in the paper.