

We thank the reviewers for their suggestions and positive feedback.

Reviewer #1, #3 - although both are notions of strong SQ, in the characterizations of [36,13] the distribution is fixed and the SQ-dim is computed over different classes. On the other hand, we fix the class and consider the SQ-dim over different distributions. This characterization was useful for us when relating it to bounded memory learning, which is the main focus of the paper.

Reviewer #4 -

- If the class is infinite, then most of the class hypotheses cannot be returned, as the memory is bounded. Therefore, the assumption that the class is finite is crucial. We know from [6] that distribution-dependent learning is equivalent to finite covering. So, as you wrote, one can focus on the cover of the class.
- By definition, small SQ-DIM implies weak bounded-memory learning (no need to use the general reduction from SQ to BM from [37]) since it suffices to find the correlation with a small number of hypotheses. The opposite direction is more challenging. Even though [17] is doing the heavy-lifting, there are still gaps even for weak learning, which we fill in this work.
- As Reviewer #2 pointed out, our contribution is conceptual; we introduce a new combinatorial definition and observe a tight characterization in a parameter regime. Obviously, it is desired to prove the characterization holds in the entire regime.