We thank all referees for their careful reading, constructive remarks and globally positive appreciations.

$\sharp$ 1 and $\sharp$ 4. We will include a short description of different privacy mechanisms. They can be non-interactive (also known as private-coin) when the users randomize independently the sample they receive, or interactive in the sense that some information is shared. We consider a large class of sequentially interactive mechanisms where some information (e.g. the privatized sample) can be transmitted from one user to all the next. From this point of view, the public-coin is a particular case of sequentially interactive mechanisms where the shared information is the original seed the first user employed. Our results imply that sequentially interactive mechanisms sharing more information than the seed (e.g. sharing the previously privatized samples) cannot improve on public-coin mechanisms, as the reviewer $\sharp$ 4 pointed out.

$\sharp$ 1 Line 49: The difference between upper and lower bounds e.g. in the non-interactive setup is that we get $\sum_{j>j_*} p_0(j)$ in the upper bounds and $\ell_* p_0(\ell_*)$ in the lower bounds. We cannot currently exclude the possibility of pathological cases where these terms strongly differ. There is a logarithmic difference if $p_0(j) \propto j^{-1}$ for $j \leq d$.

Line 74-76: the papers mentioned up to that point obtained slower rates than ours. Broadly, Gaboardi and Rogers (2017) uses a standard chi-squared statistic calculated on noisy data, while Sheffet uses a standard randomized response mechanism which performs poorly in high dimensions, even when paired with the test of Valiant and Valiant that is optimal in the non-private case.

Line 106: in case $x = x'$ the ratio is 1 and the constraint is still checked.

Lines 120-121 and Table 1: references will be included.

$\sharp$ 2 Testing composite hypotheses is certainly most challenging but beyond the scope of this paper.

We will include a reference for the concentration inequality we use.

Indeed, our results can be stated for discrete distributions with infinite support. We wanted to state the rates in terms of d in order to compare with existing literature. However, our proofs hold for j in $\mathbb{N}$ instead of j from 1 to d.

$\sharp$ 3 The reduction due to Goldreich (further developed by Acharya et al., AISTATS 2019) gives a way of transferring upper bounds from uniformity testing to general identity testing when measuring separation using the $\mathbb{L}_1$ norm. The results in Goldreich establish upper bounds for the general problem that are within a constant factor of the upper bounds for uniformity testing, though it is known that such upper bounds are generally suboptimal. The extension of this reduction by Acharya et al. [arxiv:1905.08302, Appendix D] can provide better upper bounds for non-uniform p_0, though the optimality of this approach was not proved and lower bounds do not follow. We directly provide upper and lower bounds for both $\mathbb{L}_1$ and $\mathbb{L}_2$ norms that are explicit in their dependence on p_0. The lower bounds due to Acharya et al. (AISTATS 2019) in the uniform case apply to public-coin mechanisms, a very specific type of sequentially-interactive mechanism, while ours hold more generally. We can expand the corresponding discussion in the paper (second paragraph of Section 1.2) to more clearly discuss our novelty in a revision.

We agree with the reviewer that our test statistic is an ℓ_2 statistic and to call it chi-square is an abuse of notation that will be fixed. Indeed, weighted ℓ_2 statistics are called chi-square and in the non-private setup particular weights depending on the distribution under the null, p_0, have to be employed. We tried to explain in the reduced space available, lines 160-167, that the variance of the statistic corresponding to an outcome j depends on $p_0(j)$ in the non-private setup - hence the weights, but it is free of $p_0(j)$ in the private setup (homoscedasticity) and therefore, no weights are required here.

We will replace some formulas by text in order to explain the algorithms.

Lines 142 and 186: by 'and/or' we mean that it is not an exclusive or (xor), so that both conditions may hold simultaneously. It is probably sufficient to keep 'or' instead of 'and/or'.

The envelope classes considered in arxiv:0801.2456 can be used in our upper bound results in order to state uniform results with respect to p_0 belonging to such an envelope class. However, the lower bounds cannot hold uniformly for such classes, which means that the upper bounds will be suboptimal for many distributions in the envelope class. For example, the exponentially decreasing distribution belongs to the envelope class with polynomially decreasing envelope but the optimal rate for testing it is much faster. This question is related to the point made by reviewer $\sharp$ 2 about composite null hypotheses.

$\sharp$ 4 Thank you for the details on private vs. public coin privacy mechanisms. We will include a discussion of this in a revision.

We will include the additional references, correct the typos and update the full references in the final manuscript.