We thank the reviewers for their insightful comments. We will address all these comments when updating the paper. Below we will address some of the specific concerns raised by the reviewers.

Reviewer 2 correctly pointed out that the results could be made more robust by considering the Wasserstein distance between distributions instead of statistical distance. Given that our hardness results are in fact for the discretized versions of the sampling problem, we can indeed extend all our results to the Wasserstein distance. We have verified the details and will update the paper to address this.

Reviewer 1 had concerns about the interesting-ness of the question. We will explain why we find the questions interesting. In the convex setting, the sampling and optimization problems have a very close connection. Many parts of modern ML deal with loss functions (or likelihood functions) that are not convex and in practice, we often want to solve sampling and optimization problems. For completely arbitrary f, these problems are likely unrelated. One might hope that for "natural" f, these two problems are not too different from a computational point of view, and one might hope to explain this by exploiting various properties that natural f satisfy. For example f's of interest may be efficiently computable, continuous, and smooth. Separations of the kind proved in our work show that just these conditions are not sufficient to establish any kind of computational equivalence. If for a specific problem, I want to argue equivalence, I would need to argue that the f's I care about have some additional properties that rule out the kind of separation results established in our work. Additionally, the framework may be useful to further understand what additional properties of natural functions are needed to avoid such separations.

Reviewer 1 remarked that *"the only proof which is more than a few lines long is for the result which shows that there exist some problem where sampling is hard and optimization is easy."*

We believe that the simplicity of the arguments is a feature. Indeed the previous work by Ma et al. had a long and complicated proof of a weaker version of the first result in our work.

Reviewer 1 also gave an alternate argument for one of the results in the work : *"Here is an alternative proof: take any problem where sampling is hard, change the value at* 0 *to be* $f(0) = -M$*. optimization became easy, but sampling is still equally hard."*

To our knowledge, we lacked the tools to prove that sampling is hard for some f's, so that it is not obvious how to start with a problem where sampling is hard. One of our contributions is to provide the tools to do that. Further, the function f defined by the reviewer is neither continuous nor Lipschitz and thus does not rule out equivalence of sampling and optimization for continuous and Lipschitz functions. More generally, our goal in this work was to establish a connection between the rich and well-developed area of complexity of discrete counting problems, and the question of sampling for continuous functions.

As suggested by Reviewer 3, we will add a section on future research directions.