We thank all the reviewers for their feedback!

Our paper formalizes a data acquisition problem when one cannot verify the true labels of the collected data. Understanding this problem is important as high-quality data are crucial for high-quality output of machine learning systems. The writing of our submission focused more on the basics to ensure clarity for the general, diverse Neurips readers. Most of the technical results were either deferred to the appendix or compressed to fit in the page limit. We thus want to elaborate a bit more on some of our technical contributions.

One of our major technical contributions is the explicit sensitivity guarantee of the peer-prediction style mechanisms. Sensitivity is the key property we want for a data acquisition mechanism, which prevents data providers from harmful misreporting. In this work, we provide explicit and verifiable conditions for sensitivity, which is absent in the previous work. In [1], checking whether the mechanism will be sensitive requires knowing whether a system of linear equations (which has exponential size in our problem) has a unique solution. So it is not clear how likely the mechanisms will be sensitive. In our data acquisition setting, we are able to give much stronger guarantees:

1. When θ has a finite support, we give a sufficient condition for sensitivity. The condition only uses the distribution of θ and a single data point (Corollary 5.1). The basic message is that when there is enough "correlation" between other people's data and θ, the mechanism will be sensitive. Corollary 5.1 quantifies the "correlation" using the k-rank of the distribution matrix. It is arguably not difficult to have enough "correlation": a naive relaxation of Corollary 5.1 says that assuming different θ lead to different data distributions, the mechanism will be sensitive if the total number of other people's data points $\geq |\Theta| - 1$.[1]
2. For an exponential family distribution, we give the explicit necessary and sufficient condition for the mechanism to be sensitive (Theorem 5.3), which is based on a function of the normalization term of the exponential family probability.

This kind of sensitivity guarantee is possible because of the special structure of the reports (or the signals): each dataset consists of i.i.d. samples (conditioning on theta), despite the fact that the signal space for our problem is much larger (the number of possible realizations of a dataset is exponentially large).

Besides sensitivity, another important property of our mechanisms is budget feasibility.

1. For the one-time acquisition, the log PMI payment rule requires PMI to be bounded. We give a polynomial-time method (a trivial method would take exponential time) to find the bounds for finite-size Θ (Appendix C.2).
2. For the multiple-time mechanism, budget feasibility can be guaranteed for any underlying distribution. This is achieved by carefully choosing the convex function f in the f-mutual information gain to have a bounded derivative.

If the paper is accepted, we will add more explanation about our technical results and factor out some omitted key steps in the proofs. We thank all the comments on the writing/related work/typos etc. We will revise the paper accordingly.

Other minor comments:

About the proof of Theorem 5.1. Yes, showing budget-fixed from budget bounded is straightforward based on our normalization. We will make it clear in the final version. Thanks for pointing this out.

About Example 3.2. It is a claim about knowledge: we may not need to know the entire prior to compute $p(\theta|\mathbf{x}_i, y_i)$. In this example, it is right we need to know $p(\theta)$, but we do not necessarily need $p(\mathbf{x}_i, y_i|\theta)$ because $p(y_i|\mathbf{x}_i, \theta)$ will be sufficient.

References

[1] Yuqing Kong and Grant Schoenebeck. Water from two rocks: Maximizing the mutual information. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 177–194, 2018.

[1] Here we have found a typo in our statement of Corollary 5.1, the ineuqality should be $\sum_{j\neq i}(rank_k(G_j) - 1)\ N_j + 1 \geq |\Theta|$ rather than $\sum_{j\neq i} rank_k(G_j)\ (N_j - 1) + 1 \geq |\Theta|$. We will correct this in our final version.