We thank all the reviewers for their comments and valuable feedback which will help improve the paper.

Reviewer 1

Explanation of the algorithm

We thank the reviewer for pointing out the typos. We will definitely improve the writing of the pseudo code and the prose in the final version. If the page limit becomes an issue, we will add a longer exposition in the appendix. We assure that we will address the reviewer's concerns in the final version and ensure that Section 4 and the pseudo code are reader friendly.

Concrete instances of Corollary 1, comparison to Acharya et al., and other applications

We will add concrete instantiations of Corollary 1 in the appendix for well studied symmetric functions and compare them to previous works. For a comparison between our work and Acharya et al., observe that since we release the entire histogram, our privacy mechanism can be used for many symmetric properties simultaneously, while Acharya et al.,'s work studies the problem for specific properties. Hence, our result for a specific symmetric property can be slightly worse. For example, consider entropy estimation. The main term in our privacy cost is $\tilde{\mathcal{O}}\left(\left(1/\alpha^2\epsilon\right)^{\frac{1}{1-2\beta}}\right)$ and Acharya et al's bound is $\mathcal{O}\left(1/(\alpha\epsilon)^{1+\beta}\right)$. Thus for $\beta = 0.1$, our dependence on ϵ and α is slightly worse. We agree with the reviewer that our work should also extend to other properties such distance to uniformity, which to the best of our knowledge has not been studied in the DP framework.

Doubly logarithmic dependence on k for entropy estimation

We thank the reviewer for catching this. We agree with the reviewer that dependence on $f_{\max}$ introduces an additive doubly logarithmic dependence on the domain size for entropy. We will modify line 173 to read "Furthermore, the increase is dependent on the maximum value of the function for distributions of interest and it does not explicitly depend on the support size".

Reviewer 2

Approximate vs pure DP

Since pure DP is strictly better than approximate DP, our algorithms also imply approximate DP guarantees. However, previous and our lower bounds do not hold in the approximate DP setting and we plan to pursue this in future. We thank the reviewer for raising this question.

Reviewer 3

We thank the reviewer for the stylistic comments and typos.

Comparison to Blocki et al.,'s (ϵ, δ)-DP result and other approaches

The algorithm we refer in "Pure vs approximate differential privacy" is due to Block et al., and as the reviewer stated it has an ℓ_1 error of $\mathcal{O}(\sqrt{n}/\epsilon + \log(1/\delta)/\epsilon)$. We improve on the dependence on ϵ compared to this work. Furthermore, our $(\epsilon, 0)$-DP guarantee is stronger than the (ϵ, δ)-DP of Blocki. et al.

We will also discuss previous algorithms and explicitly state which parts of our algorithm are new. To answer the reviewer's question: To the best of our knowledge both (i) splitting r around $\sqrt{n}$ and using prevalences in one regime and counts in another and (ii) the smoothing idea used to zero out the prevalence vector are new and have not been explored before. Of the two (i) is crucial for the computational complexity of the algorithm and (ii) is crucial in improving the ϵ-dependence from $1/\epsilon$ to $1/\sqrt{\epsilon}$ in the high privacy regime ($\epsilon \leq 1$). There are few subtle differences such as cumulative prevalences vs actual prevalences. We will explicitly highlight the above contributions in detail in the final version.

Finally, we note that Blocki et al., proposed an algorithm based on exponential and approximately exponential mechanisms on prevalences, whereas our algorithm is based on Laplace and Geometric mechanisms together with the splitting idea and smoothing methods described above. We will add the above discussion in detail in the paper. We hope that the above discussion clarifies the relation to the Blocki, Datta, and Bonneau paper to our work.

Even split of ϵ between ϵ_1, ϵ_2, and ϵ_3

There is no particular reason for ϵ_1, ϵ_2, and ϵ_3 to be equal and we chose those values for simplicity and easy readability. We will add a discussion in the appendix on better ways of splitting the privacy budget. For example, since ϵ_1 is just used to estimate n, analysis of the algorithm shows that ϵ_2, ϵ_3 affects utility more than ϵ_1. Hence, we can set $\epsilon_2 = \epsilon_3 = \epsilon(1 - o(1))/2$ and $\epsilon_1 = o(\epsilon)$ to get better practical results.