Dear referees and chairs,

We thank all referees for their close reading of our manuscript.

Reviewer #1:

The question on how OCO and Local Differential Privacy (LDP) are related is an important one. In training many machine learning models use an OCO algorithm. Without appropriate OCO algorithms these models can not be trained with LDP guarantees. Since noisy OCO perfectly captures the requirements to satisfy LDP guarantees noisy OCO and LDP seem to be a perfect fit. Because our bounds are adaptive to the unkown noise, data, and comparator our work is a step towards practically useful algorithms with LDP guarantees that have sound theoretical guarantees. We will make this connection more clear in the final version of the paper.

As mentioned by reviewer #3, designing data dependent bounds has been an important research topic in recent years. Results in the noiseless setting have been transitioning from the traditional worst case optimal $O(\|\boldsymbol{u}\|\sqrt{T})$ results to the more recent data dependent $O(\|\boldsymbol{u}\|\sqrt{\sum_{t=1}^T \|\boldsymbol{g}_t\|_\star^2})$. In the noisy setting we also have to adapt to the unknown parameters of the distribution of the noise for data dependent bounds. The adaptivity to the unknown parameters of the noise and data were open questions in the unconstrained setting before our paper.

Regarding novelty with respect to Jun and Orabona (2019): At a high level there are two similarities: 1) the use of the reward-regret duality (section 2.3) and 2) the use of the black-box reduction (section 3.1). Indeed, these techniques are cornerstones of much recent work in adaptive OCO. An early version of the reward-regret duality was introduced by (Mcmahan and Streeter, 2012) and has been used in for example McMahan and Orabona (2014); Orabona and Pál (2016); Orabona and Tommasi (2017); Cutkosky and Orabona (2018); and the black-box reduction comes from Cutkosky and Orabona (2018) and was also used by Jun and Orabona (2019).

Even with these techniques in hand, what was not known before our work, is how to obtain results that allow for different levels of differential privacy per user and per dimension and obtain data dependent bounds at the same time. As we mention in lines 107/108 of the paper, a partial result can be achieved by extending the techniques of Jun and Orabona (2019), but this result would be unsatisfactory, because their techniques crucially rely on knowing all the differential privacy parameters of the noise. Furthermore, this extension would still not allow for data-dependent bounds. As we argue in lines 24-26, these differential privacy parameters are themselves privacy sensitive (knowing how much someone cares about privacy may reveal that they are a celebrity for example), so we do not want to assume that they are known. We get around this issue by replacing the assumption of known privacy parameters by the alternative assumption that the noise has an arbitrary symmetric distribution for which we do not need to know the parameters or even the shape. With this new assumption we can handle all standard randomizers that are used for LDP, like for instance the Laplace randomizer.

Reviewer #2:

2. The "multiplying “$1 - \mathbb{E}[v\tilde{g}_t]$ for $t = 1, 2, \dots$” means that we multiply the bound in Lemma 4 to find the potential in equation (4).

3. The concrete form of the updates comes from working out the expectation in equation (5) for the conjugate and improper priors.

6. Here we allow the user to set $\tau_j = \infty$. While this does not give LDP guarantees for all attributes it does give LDP guarantees for attributes with $\tau_j < \infty$. One can imagine a situation in which part of the data is already public, but part of it is not. For example, a particular user might not care for privacy on social media posts but could be concerned about browsing history. Therefore, the user will set $\tau_j = \infty$ for j corresponding to social media posts, but set τ_j to be small for j corresponding to browsing history.

Thank you for pointing out the typos, we will fix them in the final version.

Reviewer #3:

We thank you for the positive review. We will try to address your comments in the final version of the paper.