First of all, we would like to thank all three reviewers for their feedback. It is clear that all of them engaged with the paper and managed to give truly thoughtful feedback. In what follows we try to address some of the concerns described in the reviews:

(1) **Relationship between RLC and relative Lipschitz continuity:** All the reviewers mentioned we should have discussed the relation between RLC and relative Lipschitz continuity. We agree with that and the lack of this discussion made it hard to understand the relationship between the results of our paper and the one by Antonakopoulos et al. [2020]. Reviewer 2 mentioned a possible equivalence (or at least close relationship) between the Riemannian and relative Lipschitz continuity definitions. In one direction, if a convex (differentiable) function f is RLC w.r.t. a metric g and a function R is strongly convex w.r.t. the metric g, then f is relative Lipschitz w.r.t. R. However, in the case where f is *not* differentiable, it is not clear if this implication holds. Without differentiability, the definition of RLC becomes more intricate. Even in the differentiable case, relative Lipschitzn continuity is at least as general as RLC by the previous argument. For the reverse direction, it is not in general obvious how to find a proper Riemannian metric to make the function RLC given relative Lipschitz continuity. Moreover, the simplicity of our proofs allow us to easily extend the results to the case with composite functions. Thus, in the revised version of the paper we plan to better discuss the relationship between RLC and our results based on the following points:

1. The results that require only relative Lipschitz continuity are at least as general as the ones for RLC;
2. Given a function that is relative Lipschitz, it is not in general obvious how to find a metric to make it RLC;
3. This works focuses on non-differentiable cost functions, where the relationship is less obvious.

(2) **Lack of motivating examples:** The most interesting applications we currently know of relative Lipschitz continuous functions, such as SVM training and Ellipsoid intersection detection, are described in the work of Lu [2019]. The inverse Poisson problem in Antonakopoulos et al. [2020] is also an application, given the discussion in the previous point. We plan to add a more interesting discussion of applications in the revised version, and given the generality of the results we expect to see more applications in the future.

(3) **Significance of our logarithmic regret bounds:** Reviewer 2 pointed out that requiring a function to be both Lipschitz continuous and strongly convex relative to the same regularizer may be too restrictive since it implies that the Bregman divergence between an optimal solution and any point in the feasible set $\mathcal{X}$ should be bounded. As they mentioned, this does not hold when $\mathcal{X}$ is unbounded or when the Bregman divergence explodes on the boundary of $\mathcal{X}$. However, we note that the same argument holds in the Euclidean case. This implies that even classical logarithmic regret results can only be applied in the case where $\mathcal{X}$ is bounded. Hence, it is not surprising that we cannot have an unbounded feasible set in the relative case.

It is worth mentioning that even the sublinear regret bound from Antonakopoulos et al. [2020] for the inverse Poisson problem requires a bounded feasible region, since the strong convexity w.r.t. the metric used only holds in a bounded set (they show that it holds on $[0, 1]^n$ and one can get strong convexity with worse constants by scaling). That is, boundedness of the feasible region is needed even for $O(\sqrt{T})$ regret guarantee in this important example. So the necessity of a bounded feasible region for logarithmic regret does not seem significantly more restrictive.

Reviewer 2 is correct that our logarithmic regret bounds cannot hold where the Bregman divergence explodes at the boundary, but the relative setting can be useful beyond such case. For example, one path of future research is to devise new algorithms for optimization problems that already have logarithmic regret bounds although with a bad dependency on the dimension. Our bounds for the relative setting are similar to the ones for the classical settings, but the Lipschitz and strong convexity constants are deeply related to the regularizer used. Thus, different regularizers might yield widely different dependency on the dimension of the problem. A classical example is the experts problem: we have $O(\sqrt{Tn})$ regret guarantee with n experts when using gradient descent (squared ℓ_2-norm as a regularizer) but a $O(\sqrt{T \log n})$ regret guarantee when using the multiplicative weights update method (negative entropy as the regularizer). It is our hope that the results for the relative setting can yield new application with better theoretical dependence on the instance's parameters, such as the dimension.

References

K. Antonakopoulos, E. V. Belmega, and P. Mertikopoulos. Online and stochastic optimization beyond lipschitz continuity: A riemannian approach. In *8th International Conference on Learning Representations, ICLR*, 2020.

H. Lu. “Relative continuity” for non-lipschitz nonsmooth convex optimization using stochastic (or deterministic) mirror descent. *Informs Journal on Optimization*, pages 265–352, 2019.