We would like to thank the reviewers for their constructive feedbacks and we will correct the typos raised and include the suggestions for improving the paper readability accordingly. We would appreciate if the reviewers positively updated their ratings if our responses were satisfactory.

1 Reviewer #1

Full (exact) conformal set vs. split or cross-validated conformal set *Full (exact)* conformal prediction set is important and worth studying since statistical efficiencies are lost both in the model fitting stage and conformity score rank computation stage in split or cross-validated approach. This is visible in many experiments conducted in previous papers [14, 15] and confirmed in ours.

Non-connectedness of the conformal prediction set. In practice, we consider the convex hull of the conformal set, which is always an interval. This was initially suggested in [18, Remark 1]. The lack of interpretability is a real issue, but it is an intrinsic default of the conformal set which remains in our proposed computation.

Choice of $[y_{\min}, y_{max}]$. We follow the actual practice in the literature [14, Remark 5]. We choose $y_{\min} = y_{(1)}$ and $y_{\max} = y_{(n)}$. In that case, we have $\mathbb{P}(y_{n+1} \in [y_{\min}, y_{\max}]) \geq 1 - 2/(n+1)$. This implies a loss in the coverage guarantee of $2/(n+1)$, which is negligible when n is sufficiently large. We did not observe violations.

Direct fitting of C_L and C_U. We do not agree that computing the whole path, with homotopy, is not needed for computing the lower and upper bound (even for Lasso and Ridge). Defining $C_L = \inf\{y > y_{\min} : \pi(y) > \alpha\}$ and $C_U = \sup\{y < y_{max} : \pi(y) > \alpha\}$, computing C_L and C_U is as hard as computing the exact conformal set. Indeed, for simplicity assume that the full exact conformal set is an interval. Then, we have

$$[C_L, C_U] = \hat{\Gamma}^{(\alpha)}(x_{n+1}) \cap [y_{\min}, y_{max}] \ . \tag{1}$$

Unless an explicit and simple enough formula for $\pi(y)$ (i.e. $\hat{\beta}(y)$) is available, such computation are intractable and was, so far, limited to class of regression problem where the entire solution path for y can be computed exactly. Our contribution, based on approximated solution and convex hull, can be interpreted as a direct estimation of C_L and C_U.

2 Reviewer #2

Paper readability. Thanks, we will fix the notation issues and put some remarks and details in the appendix to ease the reading flow. We will also summarize the proposed algorithm in a direct pseudo-code. While important, we believe that the presentation issues pointed out here can be properly corrected in the camera-ready version without changing the main technical body of the work.

Contributions compared to prior work. As discussed in line 99, the papers [18] and [14] are restricted to Ridge and Lasso where *explicit* solution are available as a piece of linear function of y. In that case, approximation is no longer needed since the exact conformal set can be efficiently computed (with *only* a single model fitting). Our main contribution (which is not limited to linear model) is *not* to provide more computationally efficient method than existing ones but to provide an easily computable conformal set, based on approximated solution, when the exact set *cannot* be computed (we have presented the `logcosh` and `Linex` loss function as examples).

Figure 1 merely illustrates the trade-off between the statistical efficiency (length of the interval) and optimization error ϵ (in case of Ridge where the exact set can be computed). We will clarify this; including suggestions of Reviewer #4.

Contrarily to previous methods, our approach provide a simple, general and *unified* framework for computing full conformal set under mild assumptions on the loss function and *any* convex regularization Ω, along with a transparent complexity analysis (which is still unknown for the exact homotopy in Lasso; note in general that *worst case* complexity of *exact* homotopy can be exponential in the dimension of the underlying optimization problem (Gartner et al., 2012)). Thus, we generalize [18] and [14] to a much wider class of machine learning problems (See answer to Reviewer #1).

3 Reviewer #4

Improvement of the numerical experiments. Thanks, we highly appreciate your critiques. The suggested experiments will be investigated and added in our paper. By doing a quick check, in the Lasso case, we observe consistent results when the coverage level α varies. This is an important point and we will perform a more complete experiments and properly report the results and our understanding.