We thank the reviewers for their thoughtful feedback and for their appreciation of the novelty of considering query-efficiency in finding homology of decision boundaries using active learning[1].

R1: Realism of assumptions and usefulness of upper-bounds. The assumptions made are known to give strong indications of the practical applicability of the algorithms and are standard in literature (see, e.g., the seminal paper [11]). In section 6 of the supplement, we provide numerical complexity comparison and show that the proposed framework uses ***ten times fewer labels*** than passive learning in the example considered. This is also anticipated in the real dataset provided that the conditional number $1/\tau$ (intrinsic complexity) of the manifold is high (lines 63-68). **R1: Extending Theorem 1 to persistence diagrams.** This is an excellent point. Persistence diagrams encode the birth and death times of topological features as a function of ϵ of the LC-complex. Our experimental results in Figure 4 and Figure 7(b) of the main paper show that samples found by active learning generate persistence diagrams closer to ground-truth ones than passive learning. Directly relating our theoretical results to persistence diagrams is a more fundamental question in manifold learning, and is out of the scope of the theory of the current manuscript; it is certainly an interesting avenue for future work – we will remark on this in the final version.

R2: The containment of ∂C and LČ complex in the tubular neighborhood of $\mathcal{M}$. We very much appreciate the R2's detailed review and technical comments. First, we want to clarify a typo in line 122 where $\mathcal{D}^0$ there actually refers to points of class 0 in the LČ complex but not the entire dataset. Nevertheless, as R2 points out, assumption 1(a) (line 118) does not guarantee that all samples of ∂C or the LČ complex falls within the tubular neighborhood of $\mathcal{M}$ (the same issue occurs in [3]). However, this can be mitigated by requiring a minor extra constraint on the tubular neighborhood of $\mathfrak{D}$ – under assumption 1(a), we only require that $3r$ is bounded from above by $(\sqrt{9} - \sqrt{8})\tau$. To see this, let $Tub_{r'}(\mathfrak{D})$ to denote a r'-radius tubular neighborhood of $\mathfrak{D}$. For samples in a covering ball $B_{r/2}(\mathbf{x})$ on manifold $\mathcal{M}$, a k-radius nearest neighbor graph requires $k >= 2r$ to have two furthest samples of opposite labels connected. That said, after constructing a k-radius nearest neighbor graph G to satisfy lemma 1 (line 172), the smallest region covered by ∂C of G (line 165) is $\mathfrak{D} + Tub_{2r}(\mathfrak{D})$. This should guarantee all samples of ∂C come from maximum allowed tubular neighbourhood of $\mathcal{M}$. We remark here that the same fix applied to the LČ complex will help correct [3]. We will make this change in final version.

R3: Importance of density near the decision boundary. Assumption 1(a)(line 118) indicates density near decision boundaries is nonzero. As a result, provided sufficient samples from the density, the proposed framework will succeed; notice that the focus here is on labeling efficiency and not sample complexity. **R3: Using topology to guide active sample acquisition.** This paper presents the first analysis of active learning for homology recovery with efficient labeling, and we adopted a simple but effective two-stage framework. Using the topology statistics to guide active learning is a very compelling avenue for future work. **R3: On realism of the model marketplace, comparison to other statistics, and other applications.** Our "model marketplace" application is different from [3]. As R3 suggests, we trained a bank of classifiers with the ***same set of training data*** (line 290-294) and verified model selection with the validation data from the same distribution; this is compelling evidence for the proposed framework, and it could indeed be made stronger by comparison with other statistics. Currently, our experiments are used to demonstrate that we find the homology of decision boundaries with fewer labels. Other applications where our work applies can be label efficient implementation of a topological regularizer [2], complexity measure [6], and finding coresets that preserve the homology of the decision boundary. In figure 1 below, we show results on the coreset application in binary classification in MNIST. As observed, predominantly active learning outperforms passive learning from coresets.

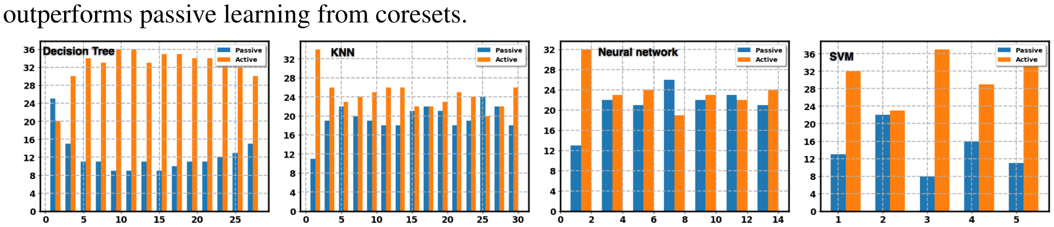

Figure 1: Training classifiers with a coreset of 300 data points sampled by active learning/passive learning. The orange (resp. blue) bars represent the number of MNIST classification cases (out of 45) where active (resp. passive) learning outperforms the other in a certain range of parameters (line 271-276) of the classifier.

[1] all references refer to the main manuscript