We thank the reviewers for taking the time to read this work and for their thoughtful reviews, Below we address some of the specific comments raised by the reviewers..

## Reviewer 1

- **The ideas presented in the paper rely heavily on the choice of hypothesis class** Note that the result immediately applies to any class that contains 1d thresholds as a subclass. In particular it extends to linear classifiers in any dimension (which is often used in practice).
- **Also, in practice, people sometimes (often?) discritize the parameter space and then use the PAC bound on that discritized hypothesis class.** Here, our result implies that the PAC Bayes guarantee will deteriorate with the precision of the discretization (as opposed to VC bounds).

## Reviewer 2

- **Should explain earlier and more intuitively why having the adversarial distribution choice not depend on P is challenging and more interesting** We did try to address this in the paper, but evidently it is not emphasized enough. We will gladly highlight it earlier and in more detail.

  We address this fact by stating that the prior may depend on the distribution (which is the same as saying that the bounds hold for a distribution which is chosen independent of the prior). E.g. in the introduction starting line 34 "We stress that…", line 40: "We emphasize… " Line 84: "Remarkably . . . .". Then, in terms of the technical challenge in lines 121 to lines 127 we explain why choosing the prior as a function of the distribution leads to the main technical difficulty.

  In any case we will definitely highlight this in the full version and rearrange things.
- **Is it possible to show non-uniform learning bound?** This is an excellent question. We will give this some thought and see if our techniques imply such bounds. In any case we will discuss this question in the final version. Thanks!
- **The relationship between uniform learning, PAC-Bayes, non-uniform learning, and these new lower bounds should be further explored in the set up of the problem.** We will follow this advice and highlight the points raised.

## Reviewer 3

- **Notably, Seeger (2002)'s ... Is the given impossibility theorem and its proof remain the same for these settings?** We note that the adversarial distribution we choose is realizable (as stated, see corollary 1 and thm2). So the result holds for this setting as well. We will gladly add further discussion on the implication of our result to these more modern bounds. Thanks for suggesting this.
- **Several published references are cited as arXiv** Thanks, we will correct this.

## Reviewer 4

- **The method proposed in this paper is limited in the simple linear classification and can be seen as a specific counter-example of PAC-Bayes analysis.** The contribution of this paper is an impossibility (hardness) result for the PAC Bayes framework. By definition, hardness results follow by exhibiting a specific counter-example, and we don't understand in what sense it is a weakness.
- **It is suggested to explore whether there exist other learnable tasks which cannot be proved with a PAC-Bayes analysis and specify the scope of effectiveness of PAC-Bayes analysis.**

  A natural future direction is indeed to characterize which classes exactly are amenable to PAC-Bayes analysis (see also our answer to reviewer 1 for a related question). We also discuss this future challenge in section 5, where we highlight recent results that might hint that understanding the role of Littlestone dimension can help to characterize the classes that are amenable to PAC-Bayes analysis. But again, we are not sure what is the weakness here.
- **so it is suggested that providing the complete proof in the main body rather than in the supplementary.**

  With the 8-page limitation of Neurips this is not reasonable.
- Overall we could not understand what is the justification for the low score provided by the reviewer. We hope the reviewer will reconsider her/his score.