High-level comments.

We thank the reviewers for their thoughtful feedback. The main focus of our paper has been an extensive theoretical analysis of this rich research area. This required tackling a variety of different questions and deriving a number of new results in order to handle sample-dependent priors with reasonable generality. We agree that our theory can and should be developed into applications, and this is the focus of ongoing work.

Reviewer 4.

- **Main limitation:** Please see high-level comments above. We will also work on improving the readability of the final version of the paper.
- **Typos:** We thank the reviewer for catching them (lines 118, 121, 125, 148, 154, 158) and will fix them all.
- **Line 161:** We agree that "the paper could give more detail about how to go from the Rademacher complexity defined for sets of hypothesis to the Rademacher complexity for the randomized classifier" and will make this more explicit.

Reviewer 6.

- **Weakness:** Please see high-level comments above.
- **Line 162:** We agree that line 162 is missing a subscript (S, as the reviewer pointed out, plus μ).
- **Broader Impact:** We will also expand on the Broader Impact section.

Reviewer 7.

- **Weakness I:** Please see high-level comments above. Additionally, we thank the reviewer for the suggestion about adding some theoretical applications as in Foster et al. (2019). In our final version, we will initiate a discussion of the choice of the parameters (and their implications) and seek to outline a theoretical application along these lines.
- **Weakness II (Theorem 4):** We agree with the reviewer. Indeed, this implies a finite class, although the bounds would be non-trivial even for a fairly large class, since the dependence on $1/\eta$ is only logarithmic. Motivated precisely by this concern, in Appendix C, we present an alternative bound without any assumption on the minimum probability, at the price of a slightly worse dependence on ϵ. On the other hand, note that Theorem 4 already yields a non-trivial generalization bound based on the observation that $|\mathcal{H}| \leq 1/\eta$. This means that $\mathfrak{R}_m(\mathcal{H}) \leq \sqrt{\frac{2\log(1/\eta)}{m}}$, and hence Theorem 4 yields a generalization bound scaling as $O\left(\frac{1}{\sqrt{m}}\right)$ even for a somewhat modest $\epsilon = 1/\sqrt{m}$, comparable to Dziugaite and Roy (2018a).
- **Empirical transductive Rademacher complexity in Theorem 2:** That is a good suggestion. A version of the theorem with an empirical transductive Rademacher complexity would be more useful in practice and we will include that in the final version. That is straightforward to derive using the fact that the empirical transductive Rademacher sharply concentrates around its expectation.
- **Clarifications for Theorem 3:** We will move these clarifications from the appendix (in the proof) to the main text and further explain the special structure, which is that the loss function in our setting, $Q \mapsto \langle Q, \ell \rangle$, is *linear*. This linearity yields the bounds in the displayed equations after line 497.
- **Discussion of Theorem 4:** As described in the Related Work section, the results of Dziugaite and Roy (2018a) are for a completely different setting than ours. In the final version, we will give a more explicit discussion of Theorem 4 and its implications, specifically contrasting Theorem 4 with the results of Dziugaite and Roy (2018a).
- **Clarification for $\mathcal{Q}_m$ and $\mathfrak{R}_m^\diamond(\mathcal{Q}_m)$:** $\mathcal{Q}_m$ is a family of sets of distributions (lines 202-203). This is in contrast with $\overline{\mathcal{Q}}_{U,m,\mu}$, which is a union (line 398, Appendix). Equation (6) is actually the definition of $\mathfrak{R}_m^\diamond(\mathcal{Q}_m)$ (lines 210-211).
- **Proof of Theorem 2:** We will revise the proof of Theorem 2 in the appendix so that all steps are more clear.