

1 We thank the reviewers for their very constructive and detailed feedback on our manuscript. In this work, we propose
 2 a novel and scalable method for inferring a continuous target as well as representations for epistemic and aleatoric
 3 uncertainty, without sampling during inference. Our method does not require any out-of-distribution (OOD) data during
 4 training (unlike Dirichlet Prior Networks [24]) and performs on-par with or better than state-of-the-art (SoA) approaches.
 5 We demonstrate learning well-calibrated measures of uncertainty on various benchmarks, scaling to high-dimensional
 6 vision tasks, as well as robustness to new OOD and adversarial test samples. As a sampling-free and performant method,
 7 our work will enable key advances in resource constrained areas, such as robotics, where sampling is infeasible.

8 **R1: 3.1. Pseudo-counts:** The overall evidence is presented as the sum of pseudo-counts
 9 [32]. We could equivalently average, by applying a constant re-scaling directly captured
 10 by λ , without changing any of our results. **3.2a. Regularizer and results:** We agree
 11 that our method provides no guarantees that it will definitively yield high epistemic
 12 uncertainty far from in-domain regions; however, we believe that the extensive empirical
 13 results achieved with our method, and results from similar related approaches which
 14 also train on *only in-distribution data* (i.e., [32], [Joo, T. et al. '20]), support the claim
 15 that uncertainty increases on OOD data. Our approach will undoubtedly improve by
 16 leveraging the supervision of OOD data during training, closer to [24, 3]; however,
 17 as we (and R2) point out, the need for OOD training data is often a critically limiting
 18 assumption. **3.2b. "Confused evidence":** As R1 correctly states, the regularizer captures scenarios where the evidence
 19 is leading to the incorrect target (i.e., incorrect or "confused" evidence, not lack of evidence). We fully agree with this
 20 excellent point and have updated the manuscript to reflect this. However, we do not believe that the approach "conflates
 21 aleatoric and epistemic uncertainty" and provide results from the suggested experiment to support our claim (Fig. 1),
 22 using the *standard score* instead of L1 error. Further details and analysis are added to the manuscript. **3.3. Other
 23 metrics:** Leveraging evidential distributions to compute M.I. or even differential entropy is a great idea, as these are
 24 rich statistics that our method captures. We focus on first order moments for more direct comparability to existing SoA
 25 baselines and leave further analysis of richer statistics to future work. **4.1. Performance:** RMSE for our method (and
 26 baselines) is in fact provided in Tab. S1, Figs. 4B, 6A. We observed little to no performance loss based on RMSE and
 27 will certainly include the other metrics as suggested. **4.2. AUC:** The histograms (and CDFs) provided in Figs. 5, 6, and
 28 S5 (as in [21], [Nalisnick, E. et al. '18], and others) are richer performance statistics and directly reduce to the requested
 29 AUC-ROC scores. To address these concerns, we have added all AUC-ROC values to our performance charts. **4.3
 30 Adversarial:** We updated the implementation details of the attack method (FGSM). While we can evaluate additional
 31 attacks, our paper is *not* proposing a new defense (neither are any of our baselines), and thus this would be out of scope.
 32 The goal of Sec. 4.3.1 is solely to evaluate on additional OOD samples based on a basic adversarial method.

33 **R2: 1. Fig. 3 aleatoric:** Within the training region there are very few differences, which can be attributed to intrinsic
 34 randomness and initialization during training. OOD there is much more variability, aligning with MVE [18, 28]. Since
 35 there is no training data in this region, we do not expect consistent results for aleatoric uncertainty, unlike epistemic
 36 uncertainty as is pointed out. **2. Relation to Kendall. et al:** This is an excellent and very important point; we apologize
 37 for the confusion. To clarify, estimating aleatoric uncertainty using NNs without sampling has a well-accepted solution
 38 dating back to 1994 with MVE (see [28]). This is the same approach used in Kendall et al. [18] and is what we
 39 compare against in our work (Fig. 3, and elsewhere when evaluating aleatoric uncertainty). Further, [18] proposes
 40 jointly learning MVE with a sampling-based epistemic uncertainty estimator (in their paper, dropout [5]). Thus, in [18]
 41 aleatoric can be estimated sampling-free, but epistemic requires sampling. We believe all our provided benchmarks do
 42 indeed accurately compare against [18]. The majority of our results focus on epistemic comparisons since our method
 43 uses a Gaussian lower-order distribution which achieves aleatoric estimation similar to [18] using MLE. In contrast,
 44 we jointly learn a *sampling-free* epistemic estimate which is not the case in [18], representing a key contribution of
 45 our work. Sampling approaches, including [18], are the current SoA and we agree with R2 that the benchmarking we
 46 provide on these methods is absolutely critical. **3. Intuition of parameters:** Thank you, additional exposition has been
 47 added. **4. Baselines:** Please refer to #2 above, which clarifies the incorrect point about missing baselines.

48 **R3: 1. Gaussian assumption:** Thank you, explanations will be added. **2. Additional comparisons and prior work:**
 49 Discussion on these works will be added. Note, [20] proposes a way to calibrate a given uncertainty method as opposed
 50 to a new uncertainty estimator, and can be used to strengthen any uncertainty estimator - it is not a competing method.

51 **R4: 1. Other distributions:** Excellent point to be included in the manuscript. **2. Performance on OOD:** Results for a
 52 variant of the proposed experiment can be found in Fig. 5. Further metrics have also been added via R1 #4.2.

53 **Summary:** Thank you for running our software. R3: "This is so far the only code I was able to run among the ones I
 54 have to review. Authors really went to a great length to provide runnable code, and this is commendable." We believe
 55 this work supports new research through its broad applicability and accessible, easy to use code. We hope the rebuttal
 56 and new experiments address all concerns (esp. R1), and appreciate all comments which have improved the manuscript.

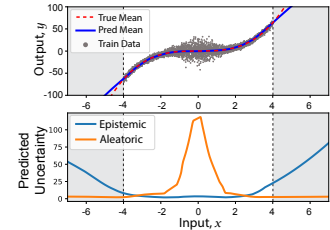


Fig. 1. Disentangled uncertainty