

Errata: Piecewise Deterministic Markov Processes for Bayesian Neural Networks

Ethan Goan

Errata Corrected on arXiv version of paper [1].

There is an incorrect negative sign in the dynamics for the Boomerang Sampler, it should read,
 $\Psi(\omega, \mathbf{v}, t)_\omega = \omega_* + (\omega^i - \omega_*) \cos(t) + \mathbf{v}^i \sin(t)$

Eq. 10: $t_i = \Lambda^{-1}(\Lambda(t_{i-1}) - \ln U)$

Eq. 11: $t_i = -b_i/a_i + \sqrt{b_i^2 + a_i^2 t_{i-1}^2 + 2a_i b_i t_{i-1} - 2a_i \log(1 - U)}/a_i$

Bug: When preparing my thesis, I found a bug in the code where gradients were not being appropriately scaled by the mini-batch and dataset size. Experiments rerun with results on arXiv.

I also added some clarification to the event rate sampling algorithm which clarifies some of the results of how thinning work. This is captured by the R term in Algorithm 2, and we explore this more in Appendix A. This R term was initially there during debugging of the approximate thinning algorithm, which acted to say that if the acceptance ratio exceeded R , then we should reject the proposal and update our approximate upper bound to try and correct for it. This is why the acceptance rates for the boomerang sampler are so tightly clustered around 1; the proposal exceeded this threshold and we rejected this proposal to try and improve our upper bound. This will be due to the non-linear dynamics of the boomerang sampler, meaning it will induce greater bias to samples from the approximate posterior.

Code updated on Github as well.

References

- [1] E. Goan, D. Perrin, K. Mengersen, and C. Fookes. Piecewise deterministic markov processes for bayesian neural networks. *arXiv e-prints*, pages arXiv-2302, 2023.