CASPEN Exit Report: Accelerated Bayesian model selection for gravitational wave astrophysics and beyond*

Alicja Polanska

The goal for my visit to the Flatiron Institute Center for Computational Astrophysics was to carry out a research sprint, combining our learned harmonic estimator implemented in the `harmonic` Python package [3, 2] and Kaze Wong and collaborators' normalizing-flow enhanced sampling package for probabilistic inference FLOWMC [7, 6, 1]. The sampling method that is standard in the field currently is nested sampling [4], which allows to perform posterior inferences as well as estimate the marginal likelihood. Despite its many advantages, it does not scale well to higher dimensions and can be very computationally expensive. This is due to its serial nature, which is not easily parallelisable, so the algorithm cannot fully take advantage of accelerators. In [7], a fast way to perform parameter estimation is introduced, but it does not produce an estimate of the marginal likelihood. Our research group has recently proposed the highly robust learned harmonic mean estimator of the marginal likelihood with normalizing flows [3, 2]. The marginal likelihood, also called the Bayesian evidence, is a crucial quantity in Bayesian model comparison, and widely used within the field of astrophysics, including gravitational wave astrophysics.

As the first proof of concept during the visit, we produced a working toy example interfacing FLOWMC and `harmonic`. We considered a posterior that was a Gaussian mixture and obtained an estimate of the evidence for this problem, which was in agreement with the analytical value. Within the next few weeks, we plan to convert this Jupyter notebook to a tutorial available on the `harmonic` GitHub repository and documentation, so that users can easily interface our codes.

Following this simple example, we began working on using this pipeline on a real gravitational waves problem. We considered the examples from [8], who use the JIM framework for gravitational wave inference employing FLOWMC to perform parameter estimation for binary neutron star mergers, comparing to results obtained with the nested sampling based PARALLEL-BILBY [5], with significant speed-up. Gravitational wave events pose a unique challenge, as their posteriors are more complex than what was considered before with the learned harmonic mean. In particular, they tend to be multimodal and contain 'holes', or low density regions, surrounded by high density regions. This is a challenging case for the learned harmonic mean, as flows are topology preserving and can struggle with multimodal distributions. To remedy this, during the visit we experimented with introducing a Gaussian mixture as the base for the normalizing flow. The initial results, for this were promising. We obtained results self-consistent within `harmonic` but not with the standard method in some cases. Following from the visit, we are continuing this work planning to explore reparametrisation, flow architectures designed to handle multimodality and different training approaches to tackle this issue.

*This work was supported by collaborative visits funded by the Cosmology and Astroparticle Student and Postdoc Exchange Network (CASPEN).

Validating our method, also poses some challenges. In the analytical case, there is a ground truth available. However, for gravitational wave problems, the reference value is obtained using the standard method of nested sampling [4]. Nested sampling error on the evidence estimate can often be underestimated. We are exploring ways to compare our value to the nested sampling result, taking into account possible underestimation of the nested sampling uncertainty. One way we already explored to validate our results independent of the nested sampling value, is to ensure they are self consistent within `harmonic`, taking into account various diagnostics and plots of the trained flow. Another way is running the code with different seeds several times and taking the result to be the mean and the uncertainty the standard deviation across runs. Further, we can also consider a simulated example where we know what model should be preferred. We are planning to report this work in a paper that we will start writing once the results and validation are finalised. We expect this will happen within the next few months.

In addition to this research work, during the visit I had the opportunity to discuss and present my work on the learned harmonic mean to researchers at the Flatiron Institute. I presented our paper on the learned harmonic mean estimator at the Bayesian statistics reading group. I was able to have insightful discussions and receive feedback on the theoretical aspect of our work from researchers at the host institute, as well as promote my work and make connections that will hopefully lead to further collaboration between our institutions.

References

[1] Marylou Gabrié, Grant M. Rotskoff, and Eric Vanden-Eijnden. "Adaptive Monte Carlo augmented with normalizing flows". In: *Proc. Nat. Acad. Sci.* 119.10 (2022), e2109420119. DOI: 10.1073/pnas.2109420119. arXiv: 2105.12603 [physics.data-an].

[2] Jason D. McEwen et al. *Machine learning assisted Bayesian model comparison: learnt harmonic mean estimator*. 2021. arXiv: 2111.12720 [stat.ME].

[3] Alicja Polanska et al. *Learned harmonic mean estimation of the Bayesian evidence with normalizing flows*. 2024. arXiv: 2405.05969 [astro-ph.IM].

[4] John Skilling. "Nested sampling for general Bayesian computation". In: *Bayesian Analysis* 1 (2006), pp. 833–859.

[5] Rory J E Smith et al. "Massively parallel Bayesian inference for transient gravitational-wave astronomy". In: *Monthly Notices of the Royal Astronomical Society* 498.3 (Aug. 2020), 4492–4502. ISSN: 1365-2966. DOI: 10.1093/mnras/staa2483. URL: http://dx.doi.org/10.1093/mnras/staa2483.

[6] Kaze W. k. Wong, Marylou Gabrié, and Daniel Foreman-Mackey. "flowMC: Normalizing flow enhanced sampling package for probabilistic inference in JAX". In: *J. Open Source Softw.* 8.83 (2023), p. 5021. DOI: 10.21105/joss.05021. arXiv: 2211.06397 [astro-ph.IM].

[7] Kaze W. K. Wong, Maximiliano Isi, and Thomas D. P. Edwards. *Fast gravitational wave parameter estimation without compromises*. 2023. arXiv: 2302.05333 [astro-ph.IM].

[8] Thibeau Wouters et al. *Robust parameter estimation within minutes on gravitational wave signals from binary neutron star inspirals*. 2024. arXiv: 2404.11397 [astro-ph.IM].