Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy

Hunter Gabbard1, Chris Messenger1, Ik Siong Heng1, Francesco Tonolini2, Roderick Murray-Smith2

SUPA School of Physics and Astronomy, University of Glasgow, United Kingdom 1
School of Computing Science, University of Glasgow, United Kingdom 2

Introduction

Gravitational wave (GW) detection is now commonplace [1] and as the sensitivity of GW detectors improves, we will observe O(100) of transient GW events per year [2]. The current methods used to estimate their source parameters employ optimally sensitive [3] but computationally costly Bayesian inference approaches [4] where typical analyses have taken between 6 hours and 5 days [5]. We show for the first time that a conditional variational autoencoder (CVAE) [6, 7] pre-trained on binary black hole (BBH) signals can return Bayesian posterior probability estimates ~ 6 orders of magnitude faster than existing techniques.

Our Model

Figure 1: The configuration of the CVAE neural network. During training (left-hand side), a training set of noisy GW signals (\(y\)) and their corresponding true parameters (\(x\)) are given as input to encoder network \(q_\phi\). While only \(y\) is given to encoder network \(r_\phi\). The KL-divergence is computed between the encoder output latent space representations \(\mu_q\) and \(\mu_r\) forming one component of the total cost function. Samples \(q_\phi\) from the \(q_\phi\) latent space representation are generated and passed to the decoder network \(r_\phi\) together with the original input data \(y\). The output of the decoder \(\mu_r\) describes a distribution in the physical parameter space and the cost component \(L\) is computed by evaluating that distribution at the location of the original input \(x\).

Methods Overview

- Generate 10^5 training BBH waveforms in Gaussian noise.
- Generate 256 test samples with postersiors produced by 4 benchmark Bayesian sampler approaches (Dynesty, CPNest, Ptemcee and Emcee).
- Train model (Vitamin) to infer posteriors for 7 parameters describing a BBH merger.
- Hyperparameters chosen through Bayesian optimisation using Gaussian Processes.
- Produce 3 figures of merit:
  - Corner plots for each test sample superimposing results from Vitamin, Dynesty and Ptemcee.
  - P-P plot showing how consistent each approach is with the truth.
  - KL-divergence plot showing how similar posteriors are between different approaches (values closer to zero indicate high degree of overlap).

Results

Figure 2: Corner plot showing one and two-dimensional marginalised posterior distributions on the GW parameters for one example test dataset using Vitamin (red), Dynesty (blue) and Ptemcee (green).

Figure 3: Distributions of KL-divergence values between posteriors produced by different samplers. We show the distribution of KL-divergences computed between a single benchmark sampler and every other benchmark sampler over all 256 GW test cases (grey) and KL-divergence distributions between the single benchmark sampler and the Vitamin outputs (blue, green, purple, yellow).

Discussion

We have demonstrated that we are able to reproduce, to a high degree of accuracy, Bayesian posterior probability distributions generated through machine learning with the same quality of results as trusted benchmark analyses used within the LIGO-Virgo Collaboration. The significance of our results is most evident in the orders of magnitude increase in speed over existing algorithms. Given that the predicted number of future detections of binary neutron star (BNS) mergers (~ 180) [2] will severely strain the GW community’s current computational resources using existing Bayesian methods, it is imperative that some form of fast posterior generation be employed. Given the abundant benefits of this method, we hope that a variant of this of approach will form the basis for future GW parameter estimation.

Contact Information

- Email: h.gabbard.10@research.gla.ac.uk

References