Importance nested sampling with `nessai` for gravitational-wave inference

Michael J. Williams, John Veitch, Chris Messenger
Institute for Gravitational Research, University of Glasgow
LISA Data Analysis Workshop 2022

Contact: m.williams.4@research.gla.ac.uk

IGR

University of Glasgow | School of Physics & Astronomy

Gravitational-wave inference

In gravitational-wave inference, we infer the parameters of a signal via Bayes' theorem

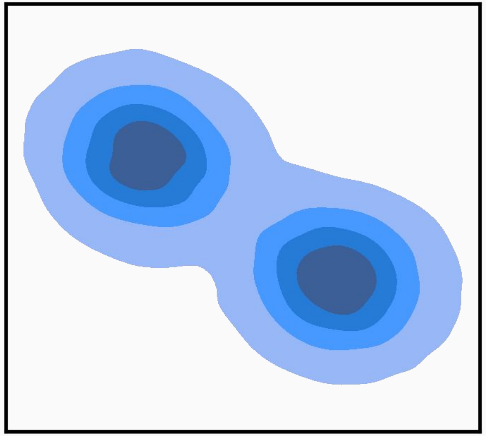
$$p(\vec{\theta}|d, H) = \frac{p(d|\vec{\theta},H)p(\vec{\theta}|H)}{\boxed{p(d|H)}}$$

Evidence $\mathbf{Z}$

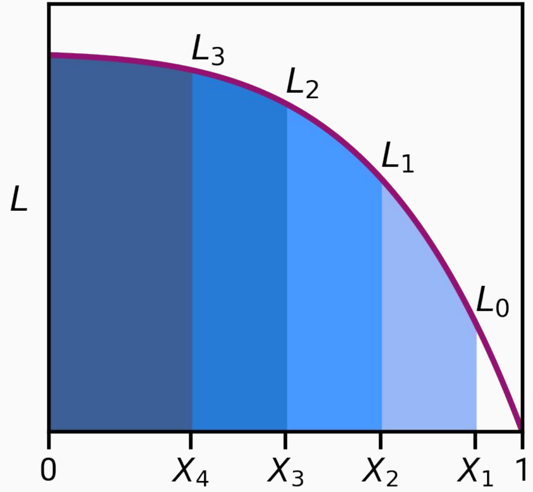
In nested sampling, the evidence is re-written as a 1-d integral in terms of prior volume X

$$Z = \int_0^1 L(X)dX$$

This allows the integral to be approximated



Parameter space



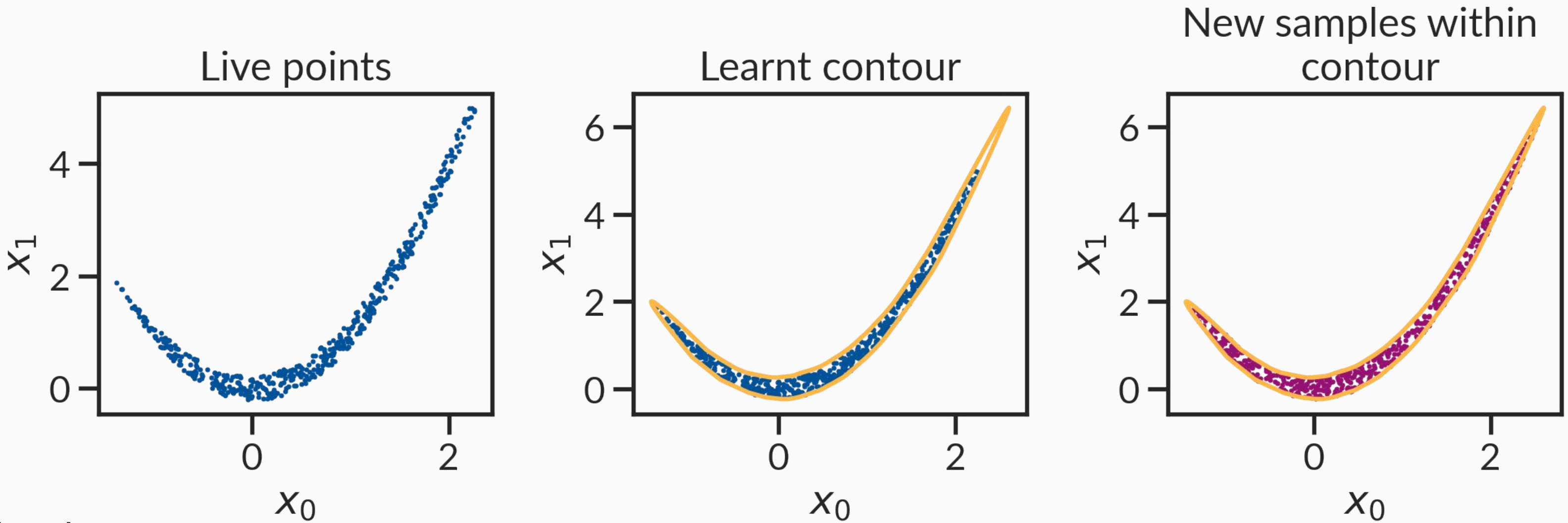
Adapted from Skilling 2006

Main challenge is drawing new samples

nessai

nessai: Nested Sampling with Artificial Intelligence

Core idea: train a normalising flow to learn likelihood contours during nested sampling, and then sample directly from those contours to produce new samples according to the prior.



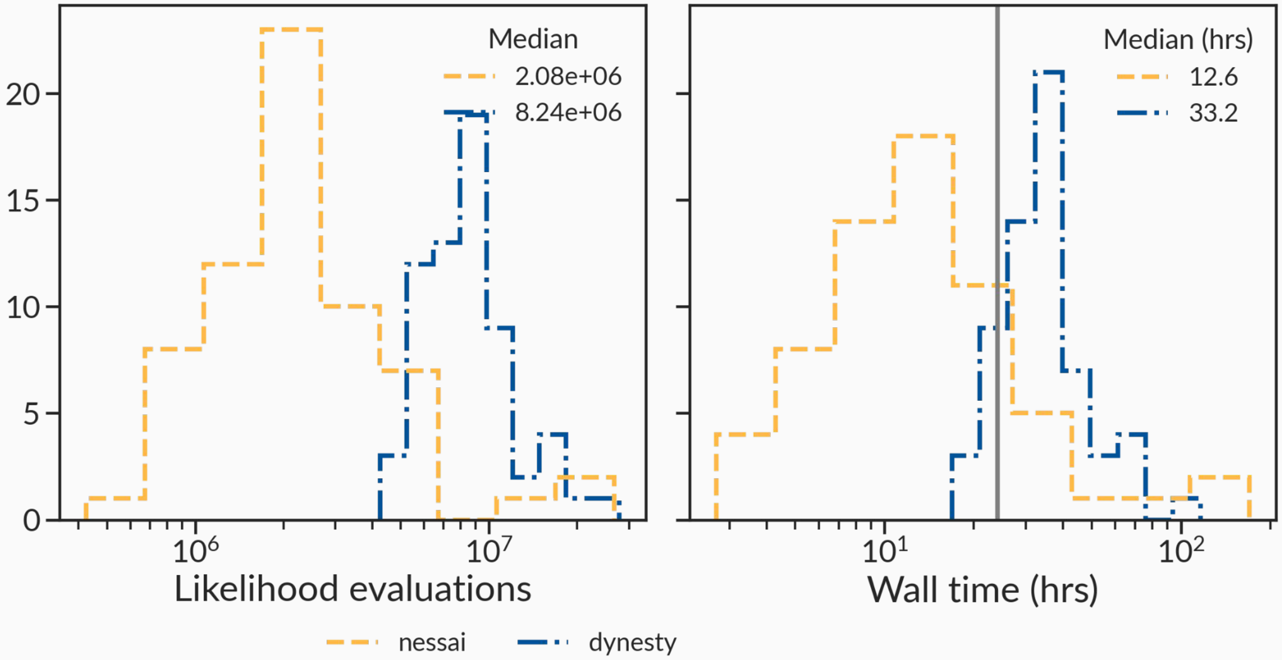
For details see: [arXiv:2102.11056](https://arxiv.org/abs/2102.11056)

Applying **nessai** to GW inference

We compared **nessai** to **dynesty** (arXiv:1904.02180) on a set of 64 BBH injections

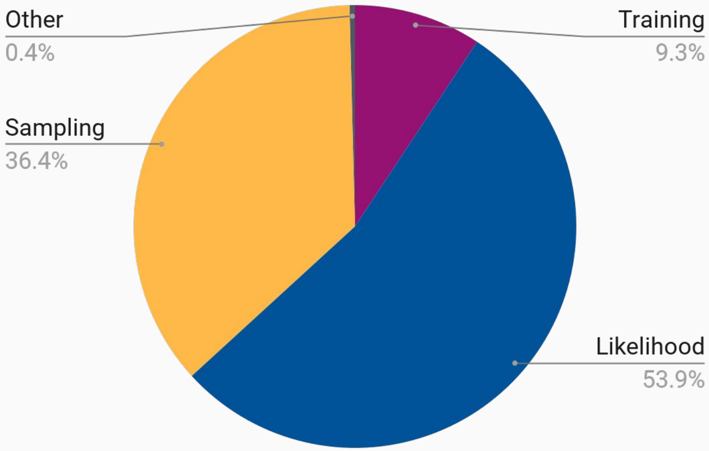
We observed a ~**4 times improvement** compared to **dynesty**

Results produced using **bilby** and **bilby_pipe**



Limitations in `nessai`

- Standard nested sampling requires samples that are distributed according to the prior
  - → Requires rejection sampling
  - → Can be a bottleneck
- Normalising flows can process batches of samples, but standard nested sampling is serial
  - → Limits parallelisation scaling



Wall time breakdown for CBC inference

Can we address these limitations?

Importance nested sampling

Aims:

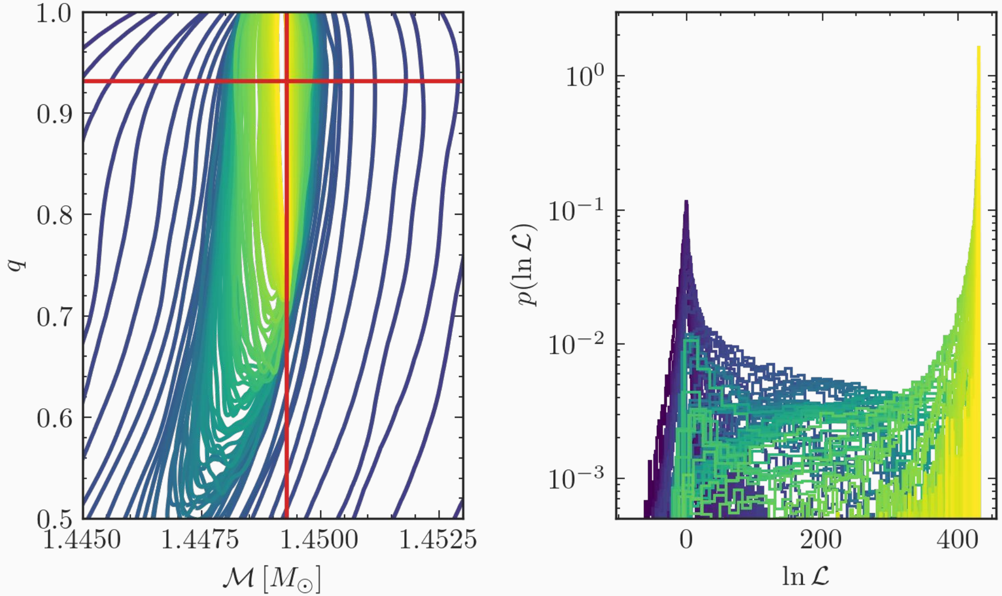
- Remove requirement for samples to drawn from the prior
- Design an algorithm that uses batches of samples
- Allow samples to be added in any order
  - → Not strictly increasing in likelihood

Method:

- Combine elements from
  - → MultiNest (arXiv:1306.2144)
  - → Diffusive nested sampling (arXiv:0912.2380)
  - → Sequential Monte Carlo (arXiv:1805.03924)
- Think of nested sampling in terms of a set of discrete levels (or distributions)

i-nessai

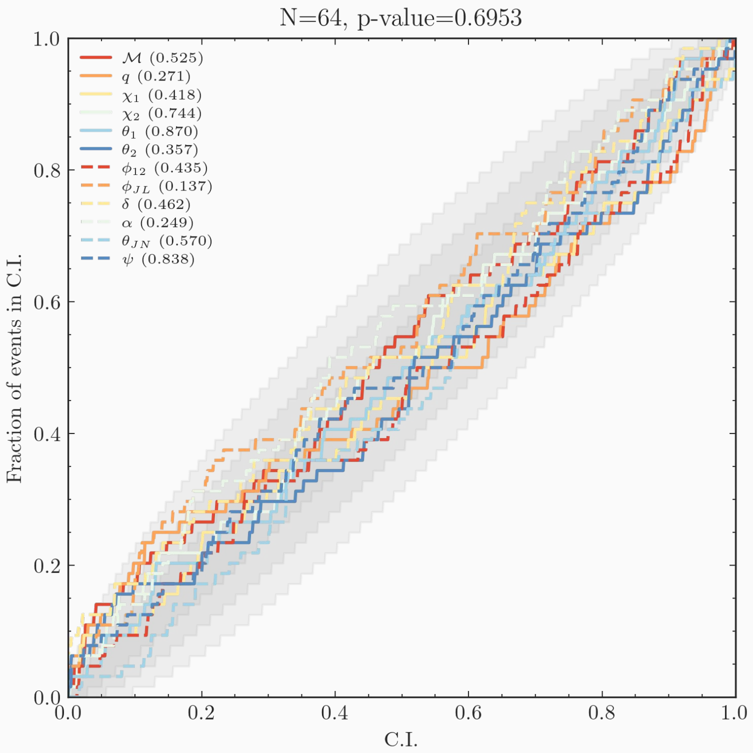
- At each iteration, batches of samples are added
  - → Updates the overall distribution
- Samples are drawn from a normalising flow
- The final result is combination of *N* distribution (normalising flows)
  - → Akin to *nested samples*



Williams et al. (In prep.)

Validating `i-nessai`

- Analyse the same BBH events using **`i-nessai`**
- Validate the results using probability-probability plots
- If the posterior distributions are unbiased, then the lines should be approximately diagonal



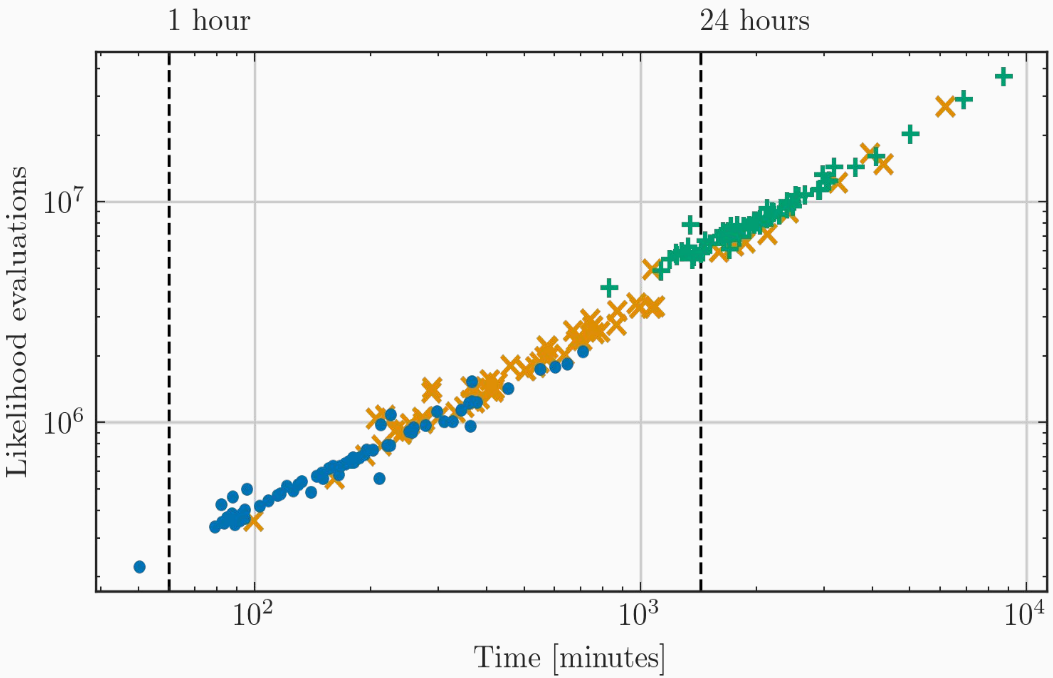
Comparing to other samplers

We compare results for BBH inference

Median time of ~2 hours with ~ 600,000 likelihood evaluations

Improvement:

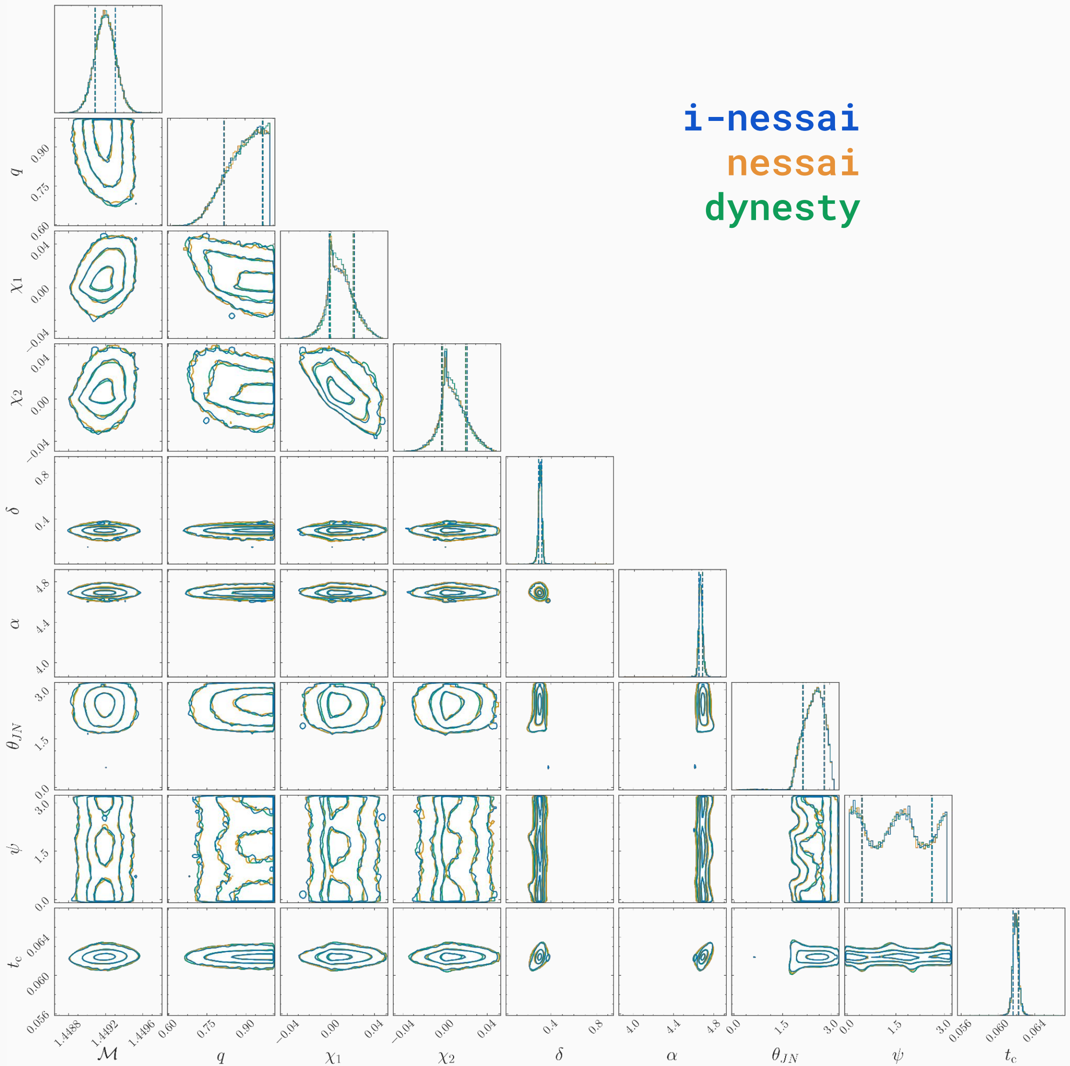
- **~2.5x** compared to **`nessai`**
- **~10x** compared to **`dynesty`**



Williams et al. (in prep.)

Application to BNS Inference

- GW190425-like BNS
- 80-second duration sampled at 8096 Hz
- Assume aligned spins and low-spin priors (<0.05)
- Use an ROQ basis and run on 16 cores
- Inference takes ~30 minutes
- Requires ~1 million likelihood evaluations compared to, ~1.8 million for **nessai** and >20 million for **dynesty**



Summary

- **`nessai`** can speed up existing analyses without any pre-training.
  - It is general purpose
  - But there are bottlenecks
- Importance nested sampling can help address these
- **`i-nessai`** can reduce the number of likelihood evaluations by an order of magnitude compared to **`dynesty`**
- Paper on **`i-nessai`** coming (very) soon

```
$ pip install nessai
```

https://github.com/mj-will/nessai

https://nessai.readthedocs.io/

arXiv:2102.11056

Try **`nessai`** out

launch binder

Extra Slides

Importance nested sampling equations

In importance nested sampling, the evidence is given by:

$$\hat{Z} = \frac{1}{N_{\mathrm{tot}}} \sum_{i=1}^{N_{\mathrm{tot}}} \frac{L(\theta_i)\pi(\theta_i)}{Q(\theta_i)}$$

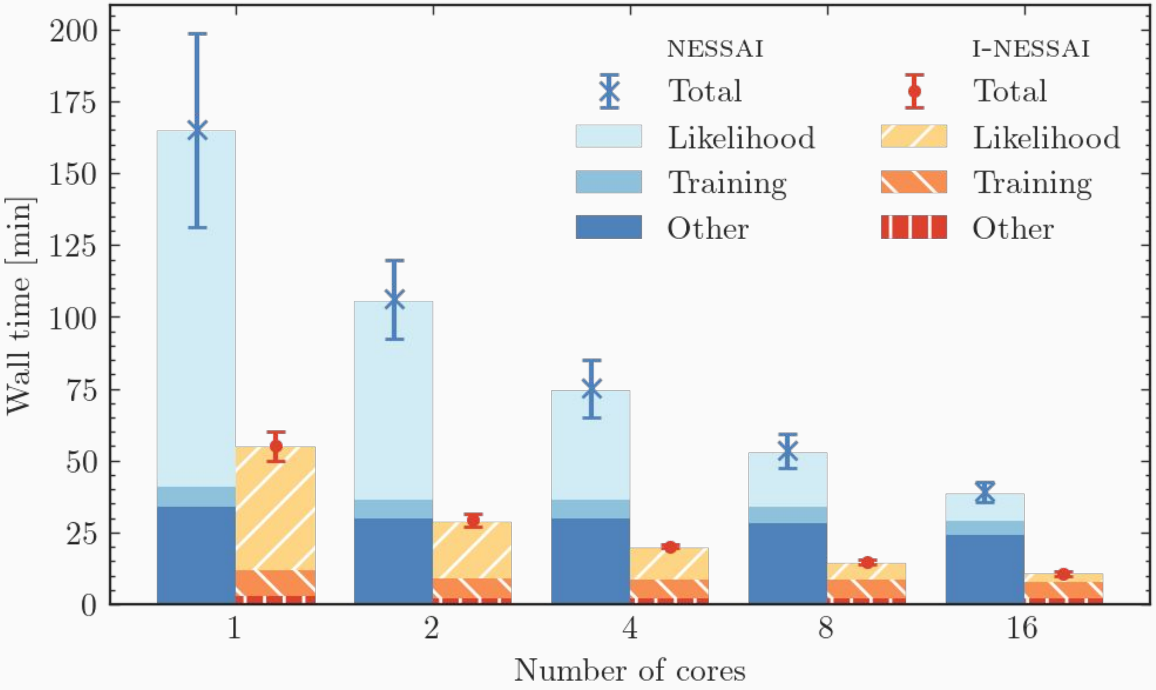
In **`i-nessai`**, *Q* is defined as

$$Q(\theta_i) = \sum_{j=1}^{N} \alpha_j q_j(\theta_i)$$

where *q* are the normalising flows.

Parallelisation

- Standard nessai is limited by rejection sampling
- How does **`i-nessai`** compare?
  - Better scaling than standard **`nessai`**



Scaling when applied to BBH analyses - Williams et al. in prep.