Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Speeding up Bayesian inference for the nHz SGWB: Importance nested sampling with Normalizing Flows

Eleonora Villa, Golam Shaifullah, Andrea Possenti, Carmelita Carbone

Spoke 3 II Technical Workshop, Bologna Dec 17 -19, 2024

Scientific Rationale

- Pulsar Timing Arrays collaborations reported evidence for the presence of a stochastic Gravitational Wave Background (SGWB) at nHz frequencies
  - constrain properties of the astrophysical sources
  - probe potential cosmological sources

- Current EPTA dataset: 25 pulsars, 61k ToA measurements, ~28-50 parameters per pulsar

≥ 2 WEEKS FOR A FULL BAYESIAN INFERENCE!

- Future IPTA dataset: 100+ pulsars, 1M+ ToA measurements, 50+ parameters per pulsar

SPEEDING UP BAYESIAN INFERENCE!

FOCUS ON NESTED SAMPLER WITH ML: NESSAI

NESSAI in a nutshell

NESTED SAMPLING IDEA

$$p(\vartheta, d) = \frac{\mathcal{L}(d,\vartheta)\pi(\vartheta)}{\int \mathcal{L}(d,\vartheta)\,\pi(\vartheta)d\vartheta} \quad \rightarrow Z = \int \mathcal{L}(d,\vartheta)\,\pi(\vartheta)d\vartheta = \int_0^1 \mathcal{L}(X)dX \qquad X = \int_{\mathcal{L}>\mathcal{L}^*} \pi(\vartheta)\,d\vartheta$$

1. Standard nested sampling: sequential sampling from the likelihood-constrained prior $\pi(\vartheta)$
https://arxiv.org/pdf/2102.11056

2. Importance nested sampling: $Z = \int \frac{\mathcal{L}(d,\vartheta)\pi(\vartheta)}{Q(\vartheta)} Q(\vartheta)d\vartheta$ independent of the $\mathcal{L} > \mathcal{L}^*$ constraint
https://arxiv.org/pdf/2302.08526

WHERE (GENERATIVE) ML COMES IN: NORMALIZING FLOWS

$$p(\vartheta, d) = \frac{\mathcal{L}(d,\vartheta)\pi(\vartheta)}{\int \mathcal{L}(d,\vartheta)\,\pi(\vartheta)d\vartheta} \quad \rightarrow Z = \int \mathcal{L}(d,\vartheta)\,\pi(\vartheta)d\vartheta = \int_0^1 \mathcal{L}(X)dX \qquad X = \int_{\mathcal{L}>\mathcal{L}^*} \pi(\vartheta)\,d\vartheta$$

$$Z = \int \frac{\mathcal{L}(d,\vartheta)\pi(\vartheta)}{Q(\vartheta)} Q(\vartheta)d\vartheta$$

The normalizing flow f learns the distribution of a set of live points via mapping from an auxiliary known simple distribution q

Standard ns: $\pi(\vartheta) = q(f(\vartheta))|det\,J|$

Importance ns: $Q(\vartheta) = q\big(f(\vartheta)\big)|det\,J|$

- **Explicit expression of the learnt distribution → normalized**
- **Convenient when training cost << sampling computational cost**

Targets and KPI in M9: 100% in M9, 25% of the project

❖ April 24 – June 24: definition of the science case

- Incorporation of NESSAI in ENTERPRISE

✓ ACHIEVED - KPI 3.2 (i)

- Hands-on training on Neural Networks with Pytorch and nflows

✓ ACHIEVED (and in progress): 10+ hands-on courses completed - KPI 3.2 (ii)

- Test the configuration of the Normalizing Flows part of the algorithm

✓ ACHIEVED (and in progress) – KPI 3.2 (iii)

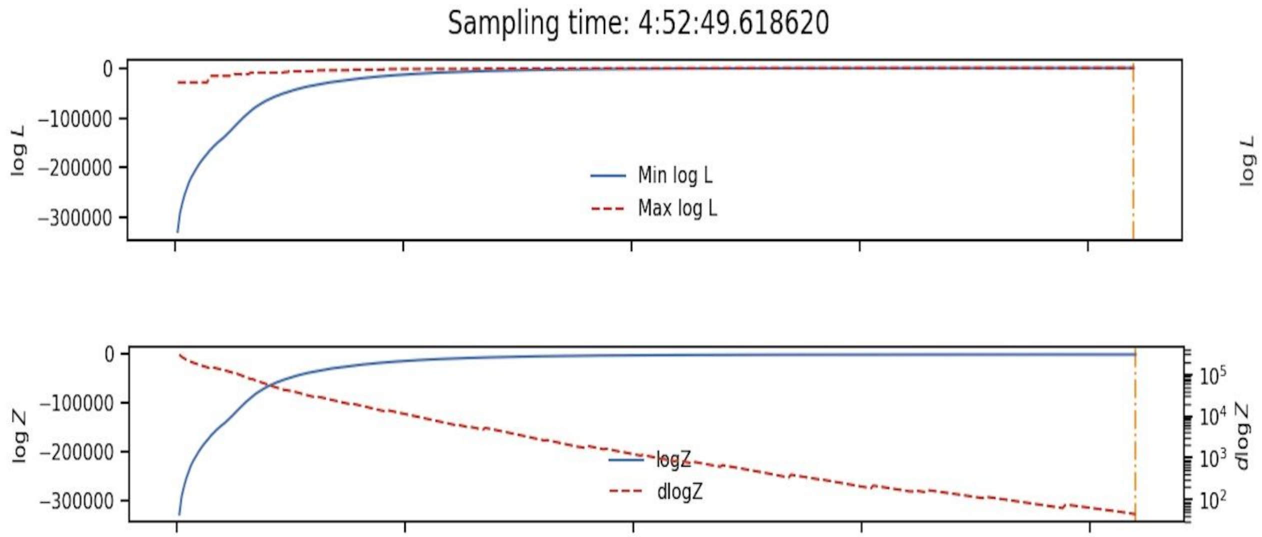
Testing NESSAI

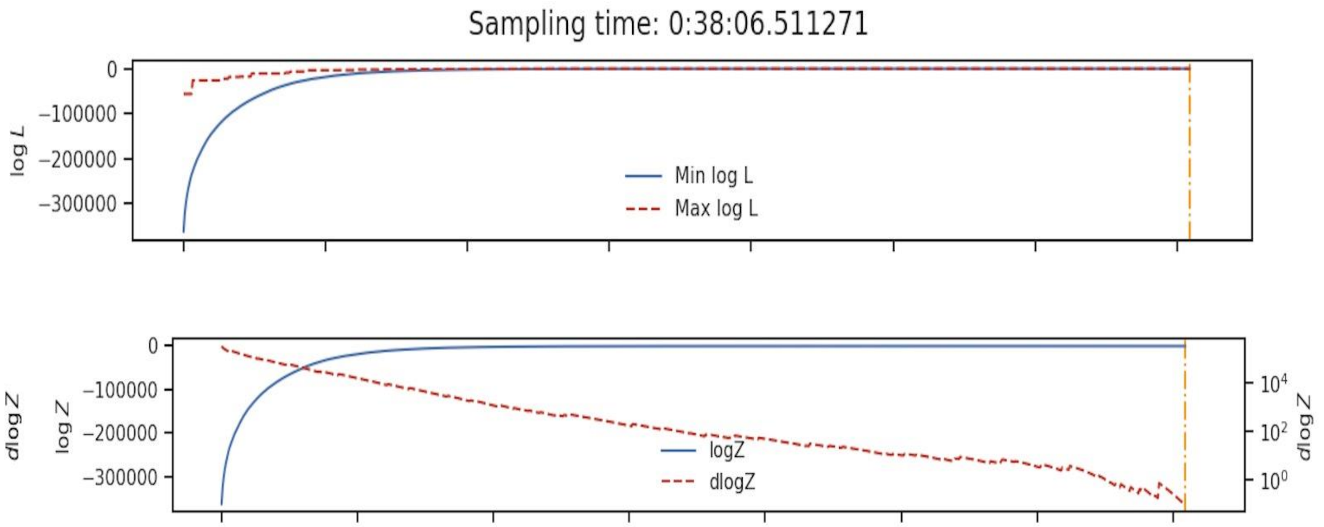
- **Rosenbrock likelihood**
  - standard benchmark, easy to compare different methods
  - challenging function to sample
  - highly correlated parameters (as we expect for PTA likelihood)

- **Fixed coupling-based NF algorithm (RNVP)**
  - computationally cheap in view of high dimensions
  - triangular Jacobian matrix
  - trivial to invert

Standard nested sampling vs importance nested sampling

Rosenbrock likelihood, n=18,~ 10^6 evaluations, 6000 live points



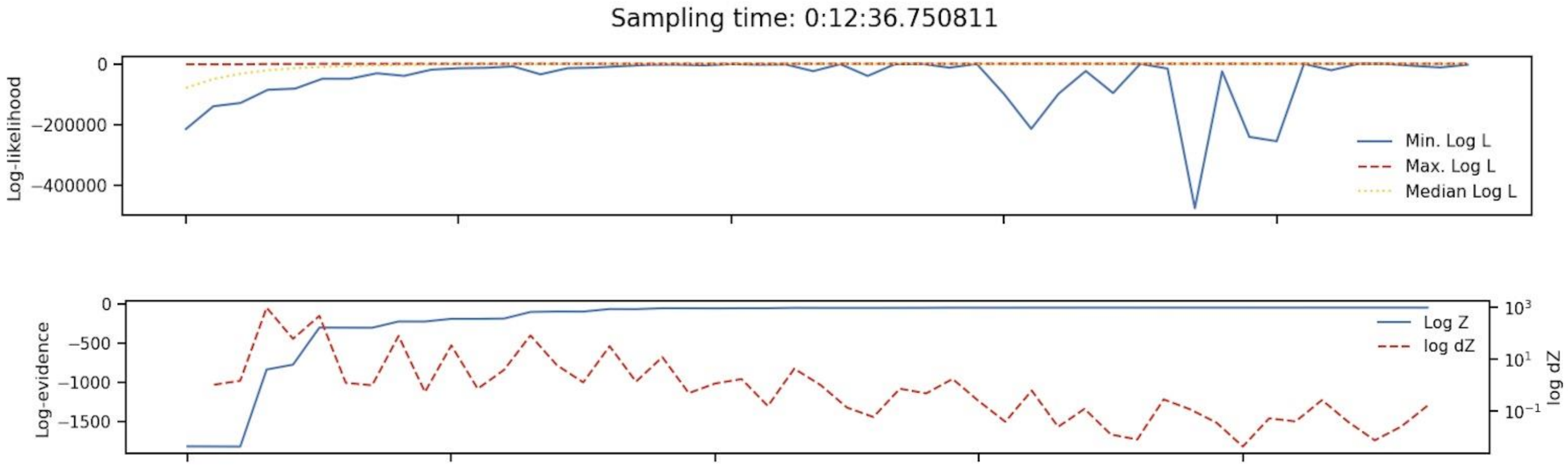


😊 Importance nested sampling much faster!
😕 Too much (relative) time spent on the plateau?
- ➢ Tune stopping criterion
- ➢ Play with NF configuration

HARDWARE:
ASUS ZenBook
32 GB RAM
CPU Intel Core i9 (20-core) 8 GB RAM

Nested importance sampling: parallelization of likelihood evaluations

Rosenbrock likelihood, n=12, 780000 evaluations, NF configuration fixed, Pytorch parallelization 8 threads



😊 It works!

😕 Fluctuations in the likelihood

➢ Configure user-definend parallelization

➢ Play with NF configuration

Road map to M10 intermediate and final targets

- **INTERMEDIATE TARGET**
  Customize the Normalizing Flows in NessAI for the PTA likelihood for the CURN model: white noise + intrinsic pulsar red noise + SGWB from autocorrelation between pulsars

- **PRESENT SITUATION: 2 weeks for a full Bayesian analysis without Normalizing Flows**

➢ **KPI: code on GitHub, paper**

- **FINAL TARGET**
  Customize the Normalizing Flows in NessAI for the full PTA likelihood: white noise + intrinsic pulsar red noise + the full SGWB signal
- **Challenge: theoretical scaling ~** $N_{pulsar} N^3_{para\ x\ pulsar} \rightarrow N^3_{pulsar} N^3_{para\ x\ pulsar}$

THANK YOU!

Back-up slides

NESTED SAMPLING IDEA

$$p(\vartheta, d) = \frac{\mathcal{L}(d,\vartheta)\pi(\vartheta)}{\int \mathcal{L}(d,\vartheta)\,\pi(\vartheta)d\vartheta} \quad \rightarrow Z = \int \mathcal{L}(d,\vartheta)\,\pi(\vartheta)d\vartheta = \int_0^1 \mathcal{L}(X)dX \qquad X = \int_{\mathcal{L}>\mathcal{L}^*} \pi(\vartheta)\,d\vartheta$$

Standard nested sampling: sequential sampling from the prior $\pi(\vartheta)$

1. Draw N live points ~ $\pi(\vartheta)$, calculate the $\mathcal{L}$ and choose the lowest value $\mathcal{L}^* = \mathcal{L}(\vartheta^*)$
2. Calculate the Z integral
3. Draw new points until $\mathcal{L} > \mathcal{L}^*$ and choose the lowest
4. Update the Z integral
5. Repeat until a stopping criterion is met, e.g. $\Delta\, ln\, Z < 0.1$

NESTED SAMPLING IDEA

Importance nested sampling: $Z = \int \frac{\mathcal{L}(d,\vartheta)\pi(\vartheta)}{Q(\vartheta)} Q(\vartheta) d\vartheta$

1. Draw N live points ~ $\pi(\vartheta)$ and calculate the $\mathcal{L}$
2. Use NF to get $Q(\vartheta)$
3. Calculate the Z integral
4. Repeat until a stopping criterion is met

Scaling of the Rosenbrock likelihood

parameters	standard ns	importance ns
5	3m04s	1m10s
10	9m57s	3m01s
12	40m09s	12m36s
16	1h48m	19m08s
18	4h52m	38m06
20	killed	40m21s

HARDWARE:
ASUS ZenBook
32 GB RAM
CPU Intel Core i9 (20-core) 8 GB RAM