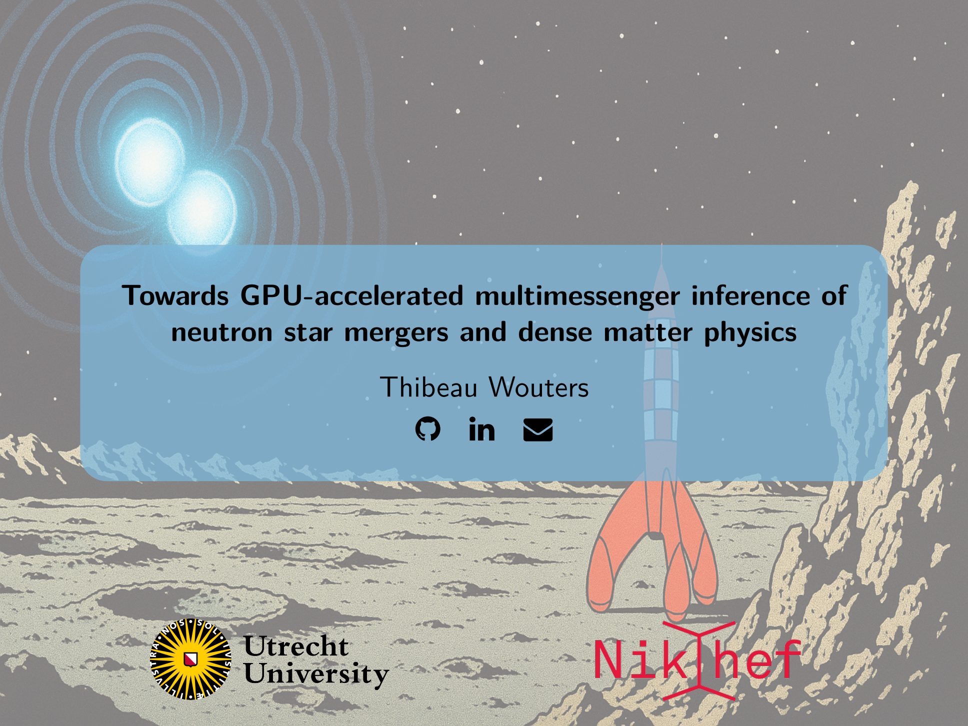


Towards GPU-accelerated multimessenger inference of neutron star mergers and dense matter physics

Thibeau Wouters

Utrecht University

Nikhef

Table of Contents

1. Introduction
2. Methods
3. Applications
4. Outlook and conclusion

Contents

1. Introduction
   - Neutron stars and equation of state
   - Multimessenger astrophysics
   - Future detectors and challenges
2. Methods
3. Applications
4. Outlook and conclusion

Contents

1 Introduction
Neutron stars and equation of state
Multimessenger astrophysics
Future detectors and challenges

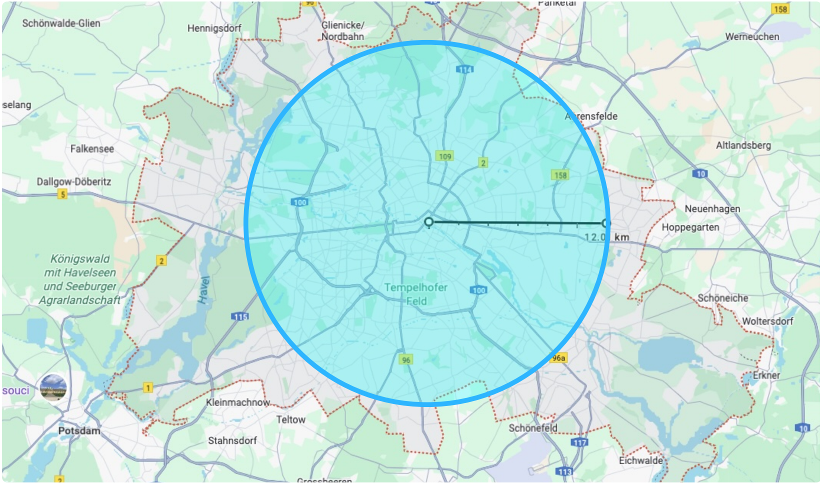
2 Methods

3 Applications

4 Outlook and conclusion

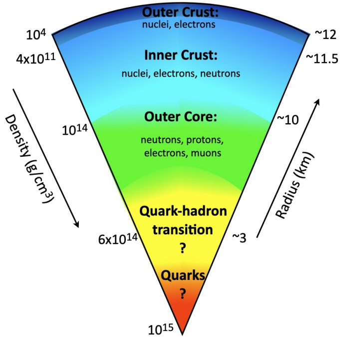
Neutron stars

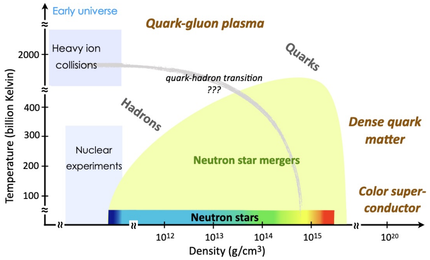
- Neutron stars: supernova remnants, densest matter in the universe
- $m \sim 1.2 - 2.3 M_{\odot}$, $R \sim 10 - 13$ km



Neutron stars

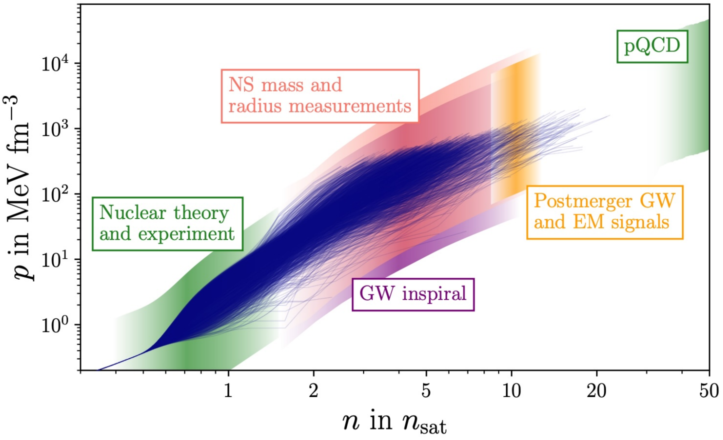
- Neutron stars: supernova remnants, densest matter in the universe
- $m \sim 1.2 - 2.3 M_\odot$, $R \sim 10 - 13$ km
- Cosmic laboratories for fundamental physics





Equation of state

- Neutron stars uniquely probe the cold equation of state (EOS) of dense nuclear matter [1]
- How to determine the EOS from neutron stars?



Contents

1 Introduction
Neutron stars and equation of state
Multimessenger astrophysics
Future detectors and challenges

2 Methods

3 Applications

4 Outlook and conclusion

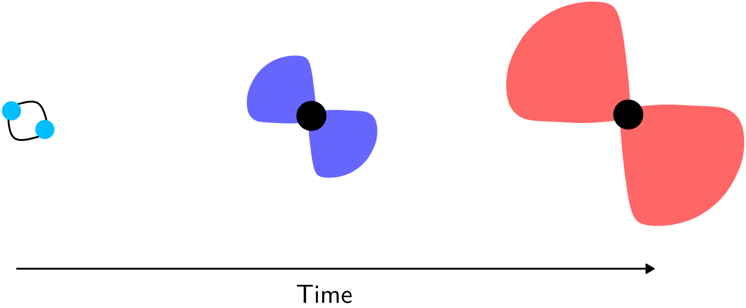
Multimessenger astrophysics

Observe neutron star mergers through gravitational waves and electromagnetic radiation: GW170817 [2, 3]

TODO: Snapshots

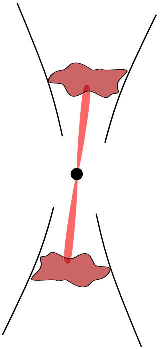
Kilonovae

- Electromagnetic counterpart of a binary neutron star merger [4]
- Lasts days-weeks, isotropic
- Powered by radioactive decay of heavy elements: r-process
- Ejecta depend on EOS TODO: how and put in the slide



Gamma-ray bursts afterglow

- Short gamma-ray burst: relativistic jet launched after merger (~ 1 s)
- Afterglow: broadband emission from radio to X-rays
- Lasts weeks-years



Contents

1 Introduction
Neutron stars and equation of state
Multimessenger astrophysics
Future detectors and challenges

2 Methods

3 Applications

4 Outlook and conclusion

Future GW detectors: Einstein Telescope

Einstein Telescope: Third-generation ground-based GW detector [5, 6]

Future GW detectors: Einstein Telescope

Einstein Telescope: Third-generation ground-based GW detector [5, 6]

- Increased sensitivity
  - Louder signals
  - $10^5 - 10^6$ black hole mergers/year (now: $\sim 200/10$ years)
  - $10^4 - 10^5$ binary neutron star mergers/year (now: $2/10$ years)
  - $10^2 - 10^3$ multimessenger events/year (now: $1/10$ years)

Future GW detectors: Einstein Telescope

Einstein Telescope: Third-generation ground-based GW detector [5, 6]

- Increased sensitivity
  - Louder signals
  - $10^5 - 10^6$ black hole mergers/year (now: $\sim 200/10$ years)
  - $10^4 - 10^5$ binary neutron star mergers/year (now: $2/10$ years)
  - $10^2 - 10^3$ multimessenger events/year (now: $1/10$ years)
- Observe from 5 Hz (now: 20 Hz)
  - Longer signals: binary neutron star: 2 hours vs 2 minutes
  - Signals will overlap

Future GW detectors: Einstein Telescope

Einstein Telescope: Third-generation ground-based GW detector [5, 6]

- Increased sensitivity
  - Louder signals
  - $10^5 - 10^6$ black hole mergers/year (now: $\sim 200/10$ years)
  - $10^4 - 10^5$ binary neutron star mergers/year (now: $2/10$ years)
  - $10^2 - 10^3$ multimessenger events/year (now: $1/10$ years)
- Observe from 5 Hz (now: 20 Hz)
  - Longer signals: binary neutron star: 2 hours vs 2 minutes
  - Signals will overlap
- Can we handle this computationally?

Computational burden of Einstein Telescope

Costs of Bayesian Parameter Estimation in Third-Generation Gravitational Wave Detectors: a Review of Acceleration Methods

Qian Hu[1,*] and John Veitch[1,†]

[1] *Institute for Gravitational Research, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, United Kingdom*

(Dated: December 4, 2024)

Bayesian inference with stochastic sampling has been widely used to obtain the properties of gravitational wave (GW) sources. Although computationally intensive, its cost remains manageable for current second-generation GW detectors because of the relatively low event rate and signal-to-noise ratio (SNR). The third-generation (3G) GW detectors are expected to detect hundreds of thousands of compact binary coalescence events every year with substantially higher SNR and longer signal duration, presenting significant computational challenges. In this study, we systematically evaluate the computational costs of source parameter estimation (PE) in the 3G era by modeling the PE time cost as a function of SNR and signal duration. We examine the standard PE method alongside acceleration methods including relative binning, multibanding, and reduced order quadrature. We predict that PE for a one-month-observation catalog with 3G detectors could require billions to quadrillions of CPU core hours with the standard PE method, whereas acceleration techniques can reduce this demand to millions of core hours. These findings highlight the necessity for more efficient PE methods to enable cost-effective and environmentally sustainable data analysis for 3G detectors. In addition, we assess the accuracy of accelerated PE methods, emphasizing the need for careful treatment in high-SNR scenarios.

- Billions – quadrillions of CPU hours for 1 month of data [7]
- Neglecting binary neutron stars & multimessenger, overlapping signals,...
- Current software does not scale

My research focus: why – how – what

Why?

To make inference of multimessenger astrophysics scalable

- Prepare for future detectors
- Understand systematic effects through simulated data

My research focus: why – how – what

Why?

To make inference of multimessenger astrophysics scalable

- Prepare for future detectors
- Understand systematic effects through simulated data

How?

Without compromises to flexibility and accuracy

- Accelerate with GPU hardware, differentiable programming
- Use machine learning to assist inference, not replace it

My research focus: why – how – what

Why?

To make inference of multimessenger astrophysics scalable

- Prepare for future detectors
- Understand systematic effects through simulated data

How?

Without compromises to flexibility and accuracy

- Accelerate with GPU hardware, differentiable programming
- Use machine learning to assist inference, not replace it

What?

GPU-accelerated Bayesian joint inference framework of neutron star mergers and dense matter physics

Contents

1 Introduction

2 Methods

Bayesian inference and sampling

3 Applications

4 Outlook and conclusion

Contents

1 Introduction

2 Methods
    Bayesian inference and sampling

3 Applications

4 Outlook and conclusion

Parameter estimation: Bayesian inference

How do we "measure" source parameters θ for data d?

Bayesian inference:

$$\text{posterior} = p(\theta|d,\mathcal{M}) = \frac{p(d|\theta,\mathcal{M})p(\theta|\mathcal{M})}{p(d|\mathcal{M})} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}$$

- Prior: specified by users
- Likelihood: often costly
- Posterior: intractable, needs stochastic samplers
- Evidence: model selection

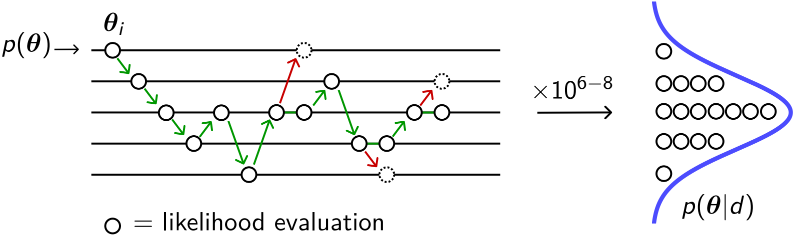
Parameter estimation: Nested sampling

How do we sample the posterior? Nested sampling
TODO: explain nested sampling

Parameter estimation: Markov chain Monte Carlo

How do we sample the posterior? Markov chain Monte Carlo

- N chains $\boldsymbol{\theta}_i$ sample posterior in parallel
- Evolve chains to new position: proposal
- Compute likelihood $\rightarrow$ accept/reject



Computational aspects

$$\text{Total runtime} \approx N_{\text{likelihood}} \times \tau_{\text{likelihood}}$$

Computational aspects

$$\text{Total runtime} \approx N_{\text{likelihood}} \times \tau_{\text{likelihood}}$$

- $N_{\text{likelihood}}$:
  - Dimensionality of $\boldsymbol{\theta}$
  - Signal-to-noise ratio
  - Multimodality/shape posterior
  - Efficiency of sampler (proposals)

Computational aspects

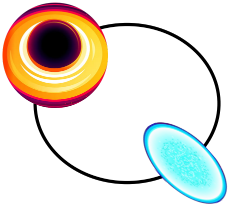
$$\text{Total runtime} \approx N_{\text{likelihood}} \times \tau_{\text{likelihood}}$$

- $N_{\text{likelihood}}$:
  - Dimensionality of $\boldsymbol{\theta}$
  - Signal-to-noise ratio
  - Multimodality/shape posterior
  - Efficiency of sampler (proposals)
- $\tau_{\text{likelihood}}$:
  - Complexity model
  - Parallelization of likelihood evaluations
  - Hardware

Tidal deformability

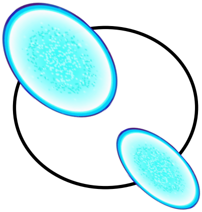
- Neutron stars are tidally deformed in a binary, quantified by tidal deformability Λ
- Depends on the mass and equation of state: $\Lambda = \Lambda(m, \mathrm{EOS})$
- Black holes: $\Lambda = 0$

$\Lambda_1 = 0$



$\Lambda(m_2, \mathrm{EOS})$

$\Lambda(m_1, \mathrm{EOS})$



$\Lambda(m_2, \mathrm{EOS})$

Tidal deformability

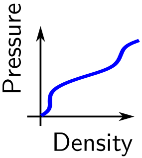
- Neutron stars are tidally deformed in a binary, quantified by tidal deformability Λ
- Depends on the mass and equation of state: $\Lambda = \Lambda(m, \mathrm{EOS})$
- Black holes: $\Lambda = 0$
- Affect phase of GWs

TODO: waveform

Neutron star observables

- Masses, radii, and tidal deformation of neutron stars
- Uniquely determined by the equation of state, have to solve the TOV equations
- How can we probe them in multimessenger astronomy?

Equation of state

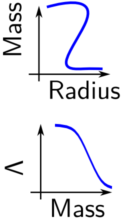


Solve TOV equations

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{\varepsilon(r)m(r)}{r^2}\frac{\left[1+\frac{p(r)}{\varepsilon(r)}\right]\left[1+\frac{4\pi r^3 p(r)}{m(r)}\right]}{\left[1-\frac{2m(r)}{r}\right]}$$

$$\frac{\mathrm{d}m}{\mathrm{d}r} = 4\pi r^2 \varepsilon(r)\,,$$

Neutron stars



Parameter estimation

Estimate parameters θ of a model for data d with Bayesian inference:

$$p(\theta|d) \propto p(d|\theta)p(\theta)$$

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

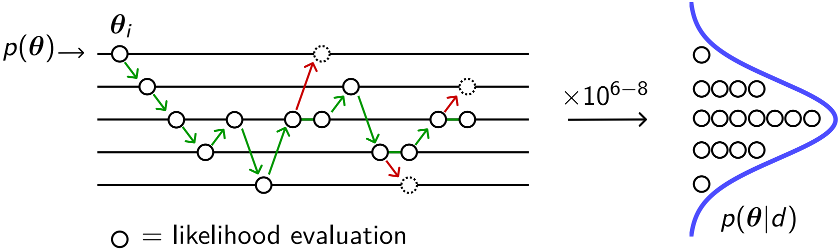
Parameter estimation

Estimate parameters θ of a model for data d with Bayesian inference:

$$p(\theta|d) \propto p(d|\theta)p(\theta)$$

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

- Sample the posterior: MCMC or nested sampling
- Propose samples, accept/reject based on likelihood
- $\mathcal{O}(10^6)$ likelihood evaluations: computational bottleneck



Future gravitational wave detectors

Future GW detectors: $10\times$ more sensitive

- $\mathcal{O}(10^5)$ events/year (now: $\mathcal{O}(10^2)$ events/decade)
- Signals are longer, louder, and overlap

Future gravitational wave detectors

Future GW detectors: $10\times$ more sensitive

- $\mathcal{O}(10^5)$ events/year (now: $\mathcal{O}(10^2)$ events/decade)
- Signals are longer, louder, and overlap

Premise: Current software does not meet these demands [7]

Future gravitational wave detectors

Future GW detectors: $10\times$ more sensitive

- $\mathcal{O}(10^5)$ events/year (now: $\mathcal{O}(10^2)$ events/decade)
- Signals are longer, louder, and overlap

Premise: Current software does not meet these demands [7]

How to make parameter estimation scalable?

- Reduce cost of likelihood evaluations
- Improve MCMC proposals

Future gravitational wave detectors

Future GW detectors: $10\times$ more sensitive

- $\mathcal{O}(10^5)$ events/year (now: $\mathcal{O}(10^2)$ events/decade)
- Signals are longer, louder, and overlap

Premise: Current software does not meet these demands [7]

How to make parameter estimation scalable?

- Reduce cost of likelihood evaluations
- Improve MCMC proposals

Goal: Fast sampling with minimal pretraining: flexible alternative to simulation-based inference [8–12]

Accelerate Python with JAX:

- GPUs
- Automatic differentiation:
  - Gradient-based samplers
  - Optimization

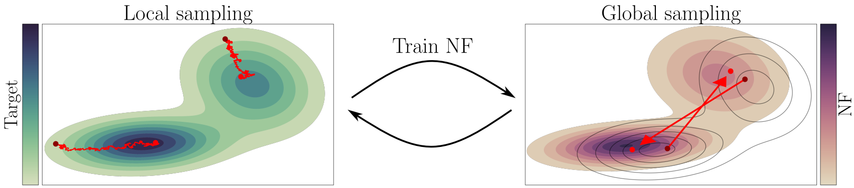
JAX & FLOWMC (Kaze Wong)

Accelerate Python with JAX:

- GPUs
- Automatic differentiation:
  - Gradient-based samplers
  - Optimization

FLOWMC [13, 14]:

- MCMC + normalizing flow proposals in JAX
- Training data: MCMC chains → no pre-training



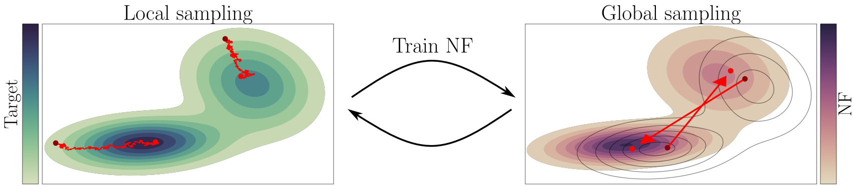
JAX & FLOWMC (Kaze Wong)

Accelerate Python with JAX:

- GPUs
- Automatic differentiation:
  - Gradient-based samplers
  - Optimization

FLOWMC [13, 14]:

- MCMC + normalizing flow proposals in JAX
- Training data: MCMC chains → no pre-training
- Also see NESSAI [15, 16], POCOMC [17]



Contents

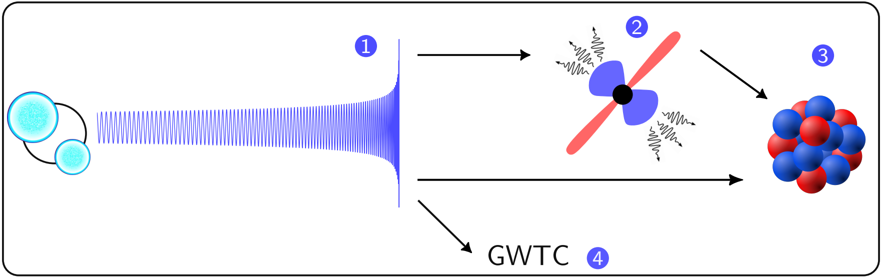
1. Introduction
2. Methods
3. Applications
4. Outlook and conclusion

Overview

Analyzing a multi-messenger **binary neutron star** signal:

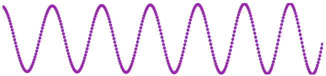
1. **Gravitational waves**
2. Electromagnetic counterparts
3. Nuclear equation of state
4. Gravitational wave transient catalogue

+ flowMC



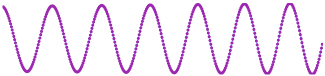
Gravitational waves

- Waveforms on GPU: $\mathcal{O}(10^3)$ faster
- From LALSUITE to JAX: RIPPLE [18]
  - Also see SFTS [19]



Gravitational waves

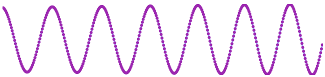
- Waveforms on GPU: $\mathcal{O}(10^3)$ faster
- From LALSUITE to JAX: RIPPLE [18]
  - Also see SFTS [19]



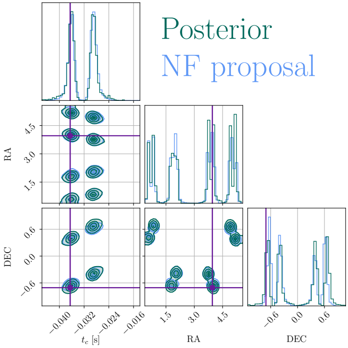
- Parameter estimation: JIM [20, 21]
- ✓ Current detectors
  - Hours → minutes

Gravitational waves

- Waveforms on GPU: $\mathcal{O}(10^3)$ faster
- From LALSUITE to JAX: RIPPLE [18]
  - Also see SFTS [19]



- Parameter estimation: JIM [20, 21]
- ✓ Current detectors
  - Hours → minutes
- Ongoing work for future detectors:
  - Binary neutron star: 13D
  - Einstein Telescope
  - 30 mins on H100 GPU

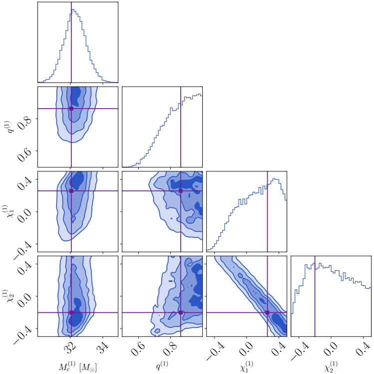


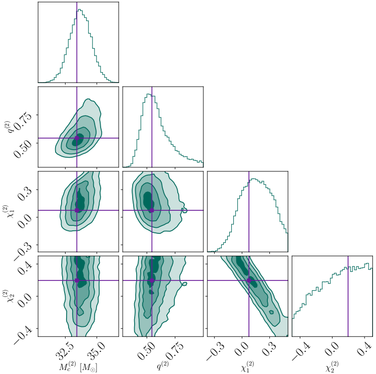
Overlapping signals (Luca Negri, Justin Janquart, James Alvey, Uddipta Bhardwaj)

- Assess scaling of JIM: BBH+BBH with LIGO-Virgo
  - 2 binary black hole mergers: 22 parameters
  - $M_c^{(1)} = 32 M_\odot$, $M_c^{(2)} = 33 M_\odot$, $\Delta t = 70$ ms
  - $\mathrm{SNR}^{(1)} = 25.76$, $\mathrm{SNR}^{(2)} = 25.24$

Overlapping signals (Luca Negri, Justin Janquart, James Alvey, Uddipta Bhardwaj)

- Assess scaling of JIM: BBH+BBH with LIGO-Virgo
  - 2 binary black hole mergers: 22 parameters
  - $\mathcal{M}_c^{(1)} = 32 M_\odot$, $\mathcal{M}_c^{(2)} = 33 M_\odot$, $\Delta t = 70$ ms
  - $\mathrm{SNR}^{(1)} = 25.76$, $\mathrm{SNR}^{(2)} = 25.24$
  - 1h28m on H100 (vs 23 days on 16 CPUs [22])



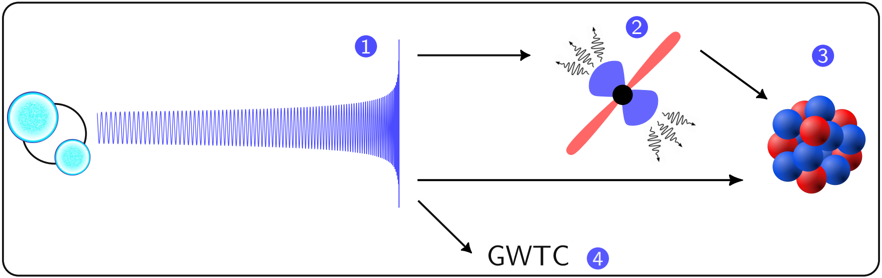


Overview

Analyzing a multi-messenger **binary neutron star** signal:

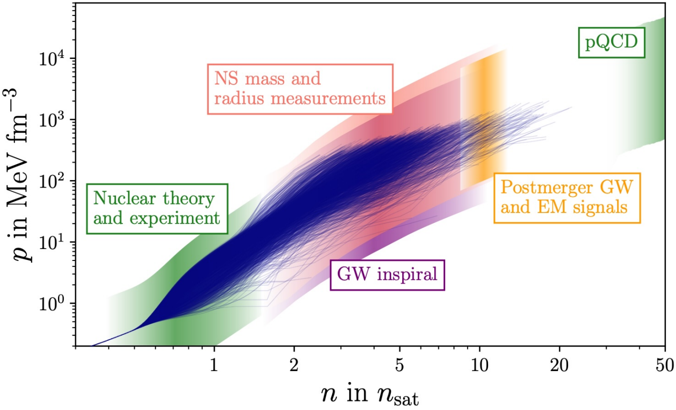
1. Gravitational waves
2. Electromagnetic counterparts
3. **Nuclear equation of state**
4. Gravitational wave transient catalogue

+ flowMC



The nuclear equation of state

- The equation of state of dense nuclear matter is uncertain [1]
- Neutron stars probe its high density regime
- Solve inverse problem with Bayesian inference

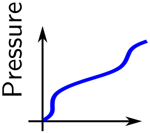


Equation of state

- Parametrization θ_{EOS}: constrain with Bayesian inference

- Parametrization θ_{EOS}: constrain with Bayesian inference
- To predict neutron star properties, we solve the TOV equations: ordinary differential equations (ODEs)

Equation of state

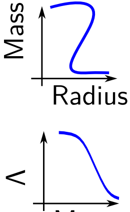


Solve TOV equations

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{\varepsilon(r)m(r)}{r^2}\frac{\left[1+\frac{p(r)}{\varepsilon(r)}\right]\left[1+\frac{4\pi r^3 p(r)}{m(r)}\right]}{\left[1-\frac{2m(r)}{r}\right]}$$

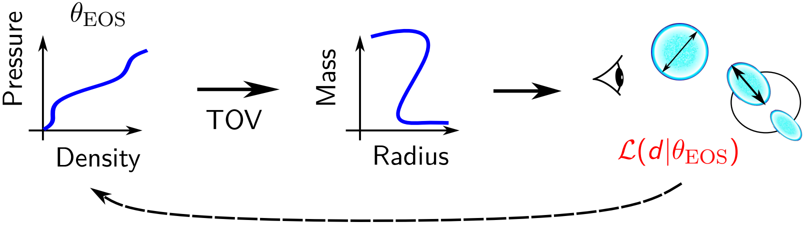
$$\frac{\mathrm{d}m}{\mathrm{d}r} = 4\pi r^2 \varepsilon(r)\,,$$

Neutron stars



Equation of state

- Parametrization θ_{EOS}: constrain with Bayesian inference
- To predict neutron star properties, we solve the TOV equations: ordinary differential equations (ODEs)
- Done for each sample θ_{EOS}: costly likelihood

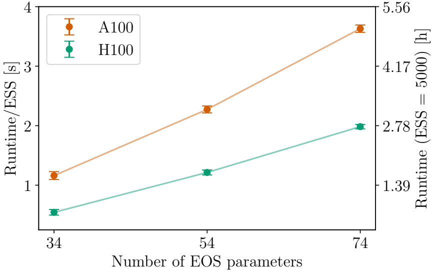


JESTER (Peter T.H. Pang)

- Solving TOV equations (EOS $\rightarrow$ NS) is slow

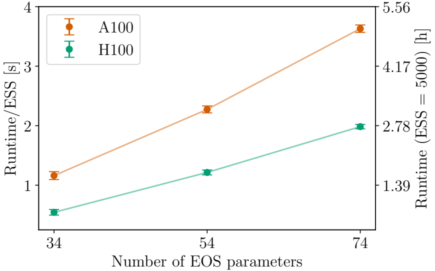
JESTER (Peter T.H. Pang)

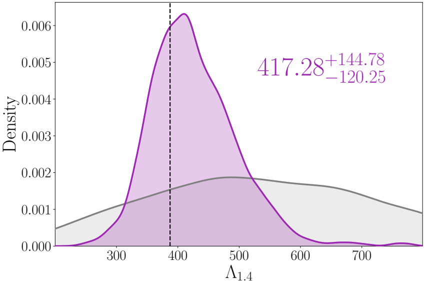
- Solving TOV equations (EOS $\rightarrow$ NS) is slow
- JESTER [23]: JAX-based TOV solver
  - Full inference in $\sim$hours
  - No need for ML emulators



JESTER (Peter T.H. Pang)

- Solving TOV equations (EOS $\rightarrow$ NS) is slow
- JESTER [23]: JAX-based TOV solver
  - Full inference in $\sim$hours
  - No need for ML emulators
- End-to-end analysis: from gravitational waves of neutron star mergers to the equation of state
  - Example: 20 BNS in O5





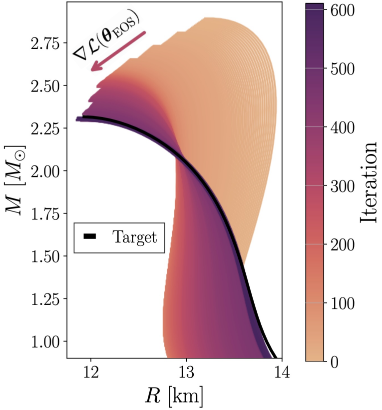
Auto-differentiable ODE solvers

- ODE solvers in JAX are auto-differentiable (DIFFRAX)
- Frame inference as optimization problem:
  - Gradient descent on loss function $\mathcal{L}(\theta_{\mathrm{EOS}})$

Auto-differentiable ODE solvers

- ODE solvers in JAX are auto-differentiable (DIFFRAX)
- Frame inference as optimization problem:
  - Gradient descent on loss function $\mathcal{L}(\theta_{\mathrm{EOS}})$

$$\mathcal{L}(\boldsymbol{\theta}_{\mathrm{EOS}}) = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{R_i(\boldsymbol{\theta}_{\mathrm{EOS}}) - \hat{R}_i}{\hat{R}_i} \right|$$



Contents

1. Introduction
2. Methods
3. Applications
4. Outlook and conclusion

Conclusion

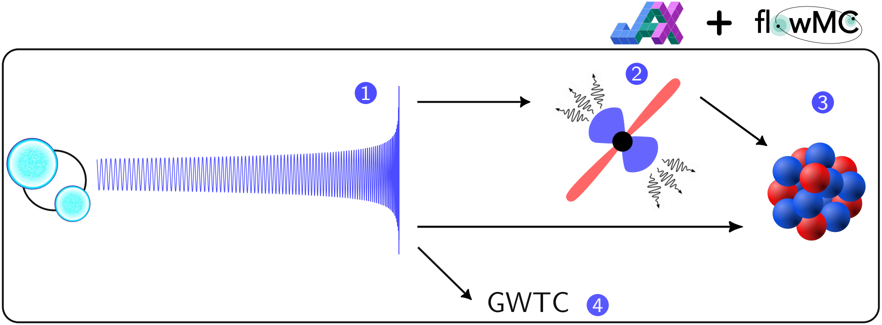
- Progress on scalable Bayesian inference, with minimal pre-training
- Hybrid acceleration: GPUs + normalizing flow proposals
  - JAX/GPU: faster likelihoods
  - FLOWMC: sampling converges faster
- Simulators in JAX can remove the need for emulators (GW, TOV)
- Auto-differentiable ODE solvers: inference as optimization problem

Let's talk!

Thank you for your attention!

Software written in JAX:

- FLOWMC [13, 14]
- JIM [20, 21] 1 2 4
- FIESTA 2
- JESTER [23] (built with DIFFRAX) 3
- HARMONIC [24–26]



References I

[1] Hauke Koehn et al. "From existing and new nuclear and astrophysical constraints to stringent limits on the equation of state of neutron-rich dense matter". In: (Feb. 2024). arXiv: 2402.04172 [astro-ph.HE].

[2] B. P. Abbott et al. "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral". In: *Phys. Rev. Lett.* 119.16 (2017), p. 161101. DOI: 10.1103/PhysRevLett.119.161101. arXiv: 1710.05832 [gr-qc].

[3] B. P. Abbott et al. "Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A". In: *Astrophys. J. Lett.* 848.2 (2017), p. L13. DOI: 10.3847/2041-8213/aa920c. arXiv: 1710.05834 [astro-ph.HE].

[4] Brian D. Metzger. "Kilonovae". In: *Living Rev. Rel.* 23.1 (2020), p. 1. DOI: 10.1007/s41114-019-0024-0. arXiv: 1910.01617 [astro-ph.HE].

[5] Michele Maggiore et al. "Science Case for the Einstein Telescope". In: *JCAP* 03 (2020), p. 050. DOI: 10.1088/1475-7516/2020/03/050. arXiv: 1912.02622 [astro-ph.CO].

[6] Adrian Abac et al. "The Science of the Einstein Telescope". In: (Mar. 2025). arXiv: 2503.12263 [gr-qc].

[7] Qian Hu and John Veitch. "Costs of Bayesian Parameter Estimation in Third-Generation Gravitational Wave Detectors: a Review of Acceleration Methods". In: (Dec. 2024). arXiv: 2412.02651 [gr-qc].

References II

[8] Jurriaan Langendorff et al. "Normalizing Flows as an Avenue to Studying Overlapping Gravitational Wave Signals". In: *Phys. Rev. Lett.* 130.17 (2023), p. 171402. DOI: 10.1103/PhysRevLett.130.171402. arXiv: 2211.15097 [gr-qc].

[9] Uddipta Bhardwaj et al. "Sequential simulation-based inference for gravitational wave signals". In: *Phys. Rev. D* 108.4 (2023), p. 042004. DOI: 10.1103/PhysRevD.108.042004. arXiv: 2304.02035 [gr-qc].

[10] Maximilian Dax et al. "Real-time inference for binary neutron star mergers using machine learning". In: *Nature* 639.8053 (2025), pp. 49–53. DOI: 10.1038/s41586-025-08593-z. arXiv: 2407.09602 [gr-qc].

[11] Qian Hu et al. "Decoding Long-duration Gravitational Waves from Binary Neutron Stars with Machine Learning: Parameter Estimation and Equations of State". In: (Dec. 2024). arXiv: 2412.03454 [gr-qc].

[12] Filippo Santoliquido et al. "Fast and accurate parameter estimation of high-redshift sources with the Einstein Telescope". In: (Apr. 2025). arXiv: 2504.21087 [astro-ph.HE].

[13] 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].

References III

[14] 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].

[15] Michael J. Williams, John Veitch, and Chris Messenger. "Nested sampling with normalizing flows for gravitational-wave inference". In: *Phys. Rev. D* 103.10 (2021), p. 103006. DOI: 10.1103/PhysRevD.103.103006. arXiv: 2102.11056 [gr-qc].

[16] Michael J. Williams, John Veitch, and Chris Messenger. "Importance nested sampling with normalising flows". In: *Mach. Learn. Sci. Tech.* 4.3 (2023), p. 035011. DOI: 10.1088/2632-2153/acd5aa. arXiv: 2302.08526 [astro-ph.IM].

[17] Minas Karamanis et al. "pocoMC: A Python package for accelerated Bayesian inference in astronomy and cosmology". In: *J. Open Source Softw.* 7.79 (2022), p. 4634. DOI: 10.21105/joss.04634. arXiv: 2207.05660 [astro-ph.IM].

[18] Thomas D. P. Edwards et al. "Differentiable and hardware-accelerated waveforms for gravitational wave data analysis". In: *Phys. Rev. D* 110.6 (2024), p. 064028. DOI: 10.1103/PhysRevD.110.064028. arXiv: 2302.05329 [astro-ph.IM].

References IV

[19] Rodrigo Tenorio and Davide Gerosa. "Scalable data-analysis framework for long-duration gravitational waves from compact binaries using short Fourier transforms". In: *Phys. Rev. D* 111.10 (2025), p. 104044. DOI: 10.1103/PhysRevD.111.104044. arXiv: 2502.11823 [gr-qc].

[20] Kaze W. K. Wong, Maximiliano Isi, and Thomas D. P. Edwards. "Fast Gravitational-wave Parameter Estimation without Compromises". In: *Astrophys. J.* 958.2 (2023), p. 129. DOI: 10.3847/1538-4357/acf5cd. arXiv: 2302.05333 [astro-ph.IM].

[21] Thibeau Wouters et al. "Robust parameter estimation within minutes on gravitational wave signals from binary neutron star inspirals". In: *Phys. Rev. D* 110.8 (2024), p. 083033. DOI: 10.1103/PhysRevD.110.083033. arXiv: 2404.11397 [astro-ph.IM].

[22] Justin Janquart et al. "Analyses of overlapping gravitational wave signals using hierarchical subtraction and joint parameter estimation". In: *Mon. Not. Roy. Astron. Soc.* 523.2 (2023), pp. 1699–1710. DOI: 10.1093/mnras/stad1542. arXiv: 2211.01304 [gr-qc].

[23] Thibeau Wouters et al. "Leveraging differentiable programming in the inverse problem of neutron stars". In: (Apr. 2025). arXiv: 2504.15893 [astro-ph.HE].

[24] Jason D. McEwen et al. *Machine learning assisted Bayesian model comparison: learnt harmonic mean estimator*. 2023. arXiv: 2111.12720 [stat.ME]. URL: https://arxiv.org/abs/2111.12720.

References V

[25] Alicja Polanska et al. *Learned harmonic mean estimation of the marginal likelihood with normalizing flows*. 2024. arXiv: 2307.00048 [stat.ME]. URL: https://arxiv.org/abs/2307.00048.

[26] Alicja Polanska et al. "Accelerated Bayesian parameter estimation and model selection for gravitational waves with normalizing flows". In: *38th conference on Neural Information Processing Systems*. Oct. 2024. arXiv: 2410.21076 [astro-ph.IM].

[27] Kurzgesagt. *Figures taken from "Neutron Stars - The Most Extreme Things that are not Black Holes"*. Accessed on May 14, 2025. 2019. URL: https://www.youtube.com/watch?v=udFxKZRyQt4.

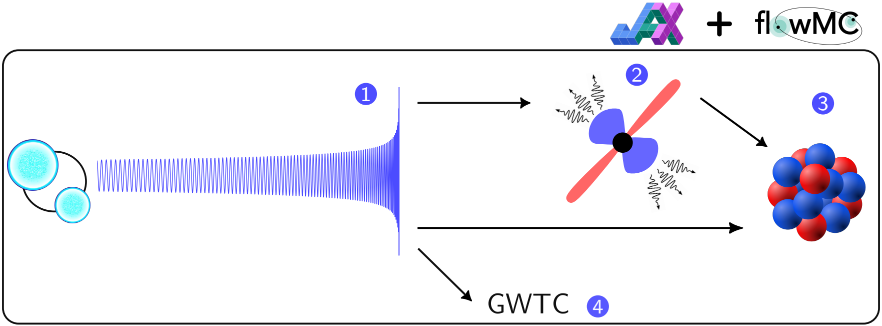
[28] Hergé. *Cover figure created with ChatGPT using this input figure from the comic Destination Moon*. Accessed on May 14, 2025. 2019. URL: https://www.youtube.com/watch?v=udFxKZRyQt4.

[29] Geoffrey Ryan et al. "Gamma-Ray Burst Afterglows in the Multimessenger Era: Numerical Models and Closure Relations". In: *Astrophys. J.* 896.2 (2020), p. 166. DOI: 10.3847/1538-4357/ab93cf. arXiv: 1909.11691 [astro-ph.HE].

Overview

Analyzing a multi-messenger **binary neutron star** signal:

1. Gravitational waves
2. **Electromagnetic counterparts**
3. Nuclear equation of state
4. Gravitational wave transient catalogue

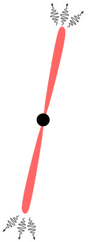


Electromagnetic counterparts (Hauke Koehn, Tim Dietrich)

- BNS mergers lead to kilonovae, gamma-ray bursts (afterglows)
- Numerical models are expensive (e.g. AFTERGLOWPY [29])

Electromagnetic counterparts (Hauke Koehn, Tim Dietrich)

- BNS mergers lead to kilonovae, gamma-ray bursts (afterglows)
- Numerical models are expensive (e.g. AFTERGLOWPY [29])
- Neural network emulators for inference: FIESTA

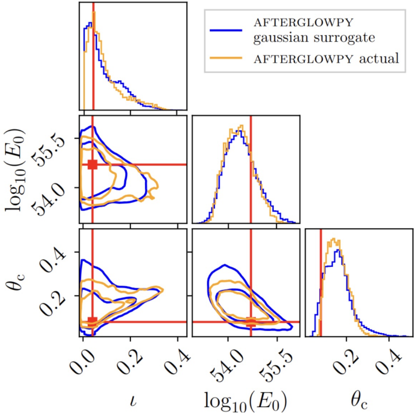


FIESTA

- 1m36s
- 1 H100 GPU

AFTERGLOWPY

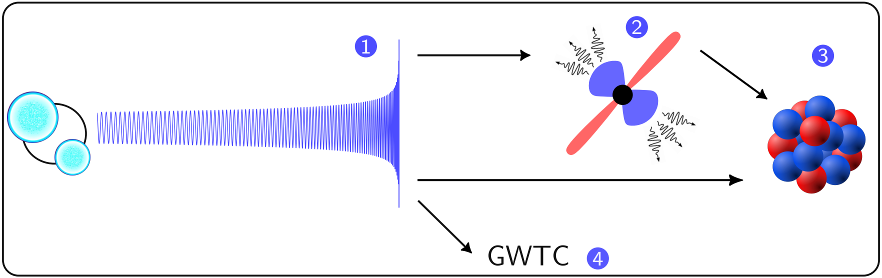
- 4 hours
- 30 CPUs



Overview

Analyzing a multi-messenger **binary neutron star** signal:

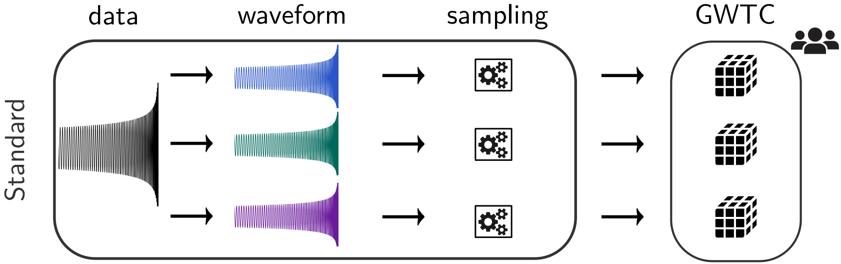
1. Gravitational waves
2. Electromagnetic counterparts
3. Nuclear equation of state
4. **Gravitational wave transient catalogue**



Constructing GWTCs (Thomas Ng, Kaze Wong)

GWTCs do not scale well in memory:

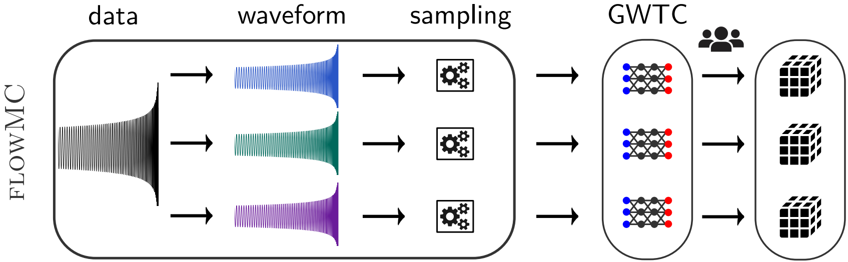
- GWTC stores several samples (different waveforms)
- Standard: fixed sample size, $\sim$ 100 MB



Constructing GWTCs (Thomas Ng, Kaze Wong)

GWTCs do not scale well in memory:

- GWTC stores several samples (different waveforms)
- Standard: fixed sample size, $\sim$ 100 MB
- FLOWMC: generate samples from normalizing flows, $\sim$ 10 MB
  - Also see Michael Williams' talk/poster



Evidence calculation: HARMONIC I

Evidence Z can be computed from posterior samples with HARMONIC [24] with the harmonic mean estimator

$$\begin{aligned}\rho \equiv \mathbb{E}_{P(\theta|d)}\left[\frac{1}{L(\theta)}\right] &= \int \mathrm{d}\theta \frac{1}{\mathcal{L}(\theta)} P(\theta|d) \\ &= \int \mathrm{d}\theta \frac{1}{\mathcal{L}(\theta)} \frac{\mathcal{L}(\theta)\pi(\theta)}{Z} = \frac{1}{Z}\end{aligned}$$

Therefore, estimate ρ with posterior samples:

$$\hat{\rho} = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\mathcal{L}(\theta)}, \quad \theta_i \sim P(\theta|d)$$

Evidence calculation: HARMONIC II

Can be interpreted as importance sampling

$$\rho = \int \mathrm{d}\theta \frac{1}{Z} \frac{\pi(\theta)}{P(\theta|d)} P(\theta|d) \, ,$$

but with target = prior and sampling density = posterior. Therefore, importance sampling is inefficient – how to solve?
New proposal:

$$\begin{aligned} \rho &= \mathbb{E}_{P(\theta|d)} \left[\frac{\varphi(\theta)}{\mathcal{L}(\theta)\pi(\theta)} \right] \\ &= \int \mathrm{d}\theta \, \frac{\varphi(\theta)}{\mathcal{L}(\theta)\pi(\theta)} \, P(\theta|d) \\ &= \int \mathrm{d}\theta \, \frac{\varphi(\theta)}{\mathcal{L}(\theta)\pi(\theta)} \, \frac{\mathcal{L}(\theta)\pi(\theta)}{Z} = \frac{1}{Z} \end{aligned}$$

Evidence calculation: HARMONIC III

Use the following estimator:

$$\hat{\rho} = \frac{1}{N}\sum_{i=1}^{N}\frac{\varphi(\theta_i)}{\mathcal{L}(\theta_i)\pi(\theta_i)}, \quad \theta_i \sim P(\theta|d)$$

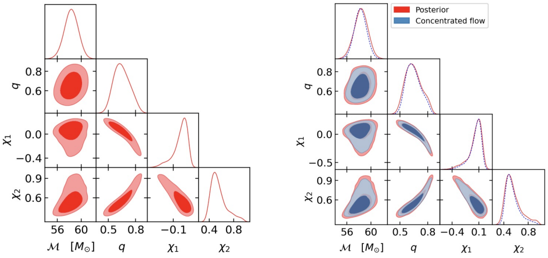
Replace the target distribution π with φ: only requirement is that it is normalized

In practice, this can be achieved with a normalizing flow [25].

This has been verified to give accurate evidences (similar values as nested sampling) when GW posteriors are used [26].

Table 1: Total wall times to compute the evidence estimates for the examples discussed in the main text. We run BILBY on 16 CPU cores and JIM + harmonic on 1 GPU.

Example	Method	$\log(z)$	Sampling time	Evidence estimation time
4D	BILBY	390.33 ± 0.11	31.3 min	–
	JIM + harmonic	$390.360^{+0.006}_{-0.006}$	3.4 min	1.9 min
11D	BILBY	378.29 ± 0.15	3.5 h	–
	JIM + harmonic	$378.420^{+0.09}_{-0.08}$	11.8 min	2.4 min

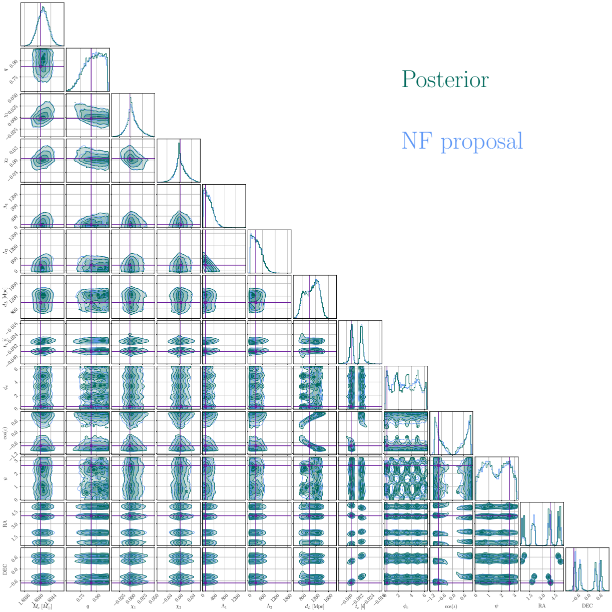


(a) BILBY

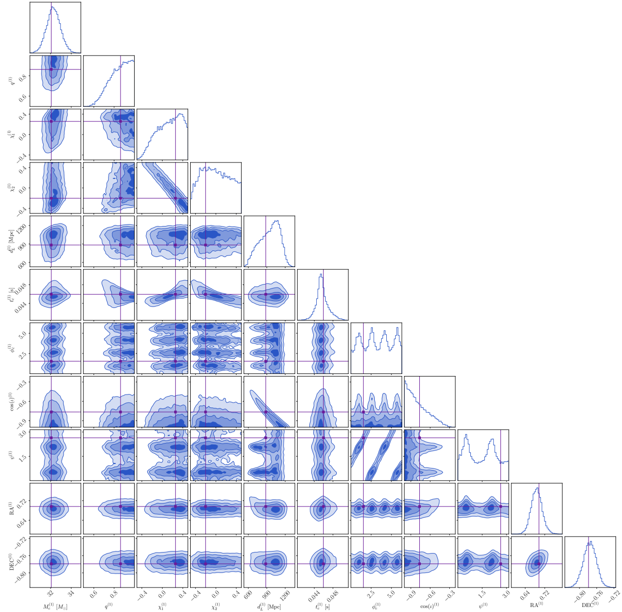
(b) JIM + harmonic

Figure 1: Corner plots for the 4-dimensional posterior samples from (a) BILBY and (b) JIM used for inference (solid red) alongside the concentrated flow at $T = 0.8$ used in the learned harmonic mean (dashed blue).

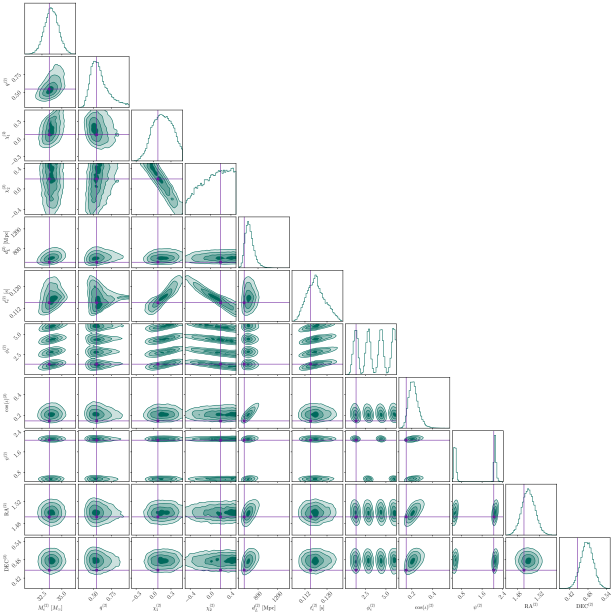
BNS in ET-Δ example: all parameters



Overlapping signals: all parameters signal A

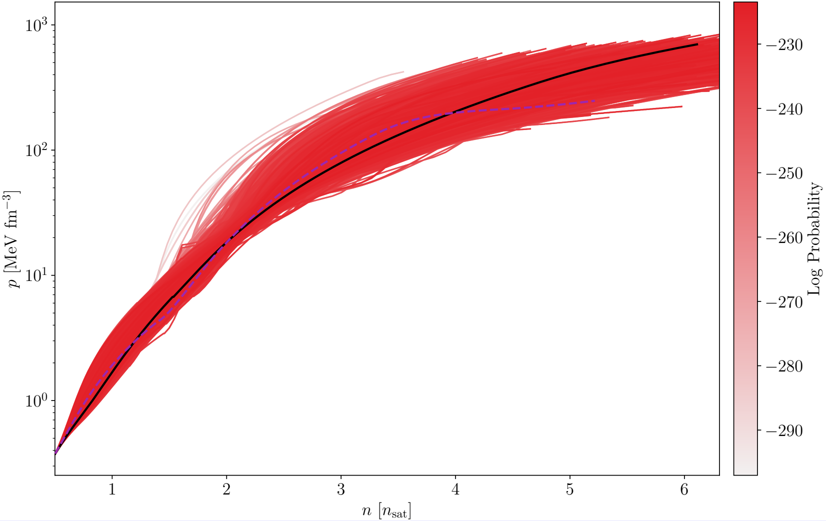


Overlapping signals: all parameters signal B



Equation of state O5 projection with 20 BNS: EOS

- **Purple**: target
- **Red**: posterior EOS samples (**black**: maximum log posterior)



Equation of state O5 projection with 20 BNS: NS