Simulation-based inference in gravitational wave science

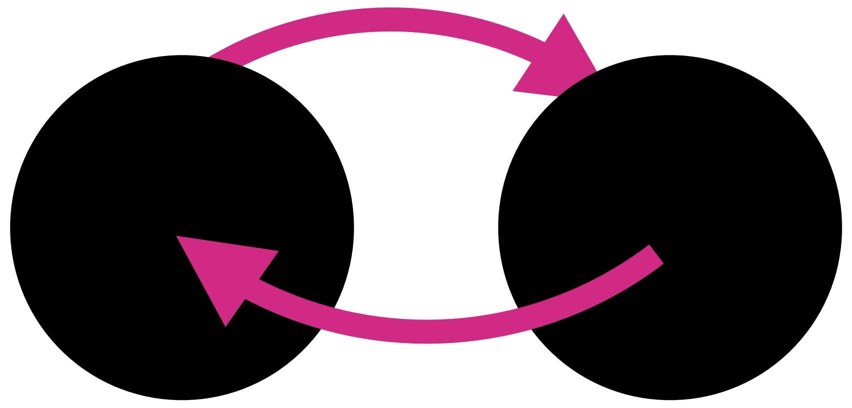
An overview

Annalena Kofler, 05.09.2025

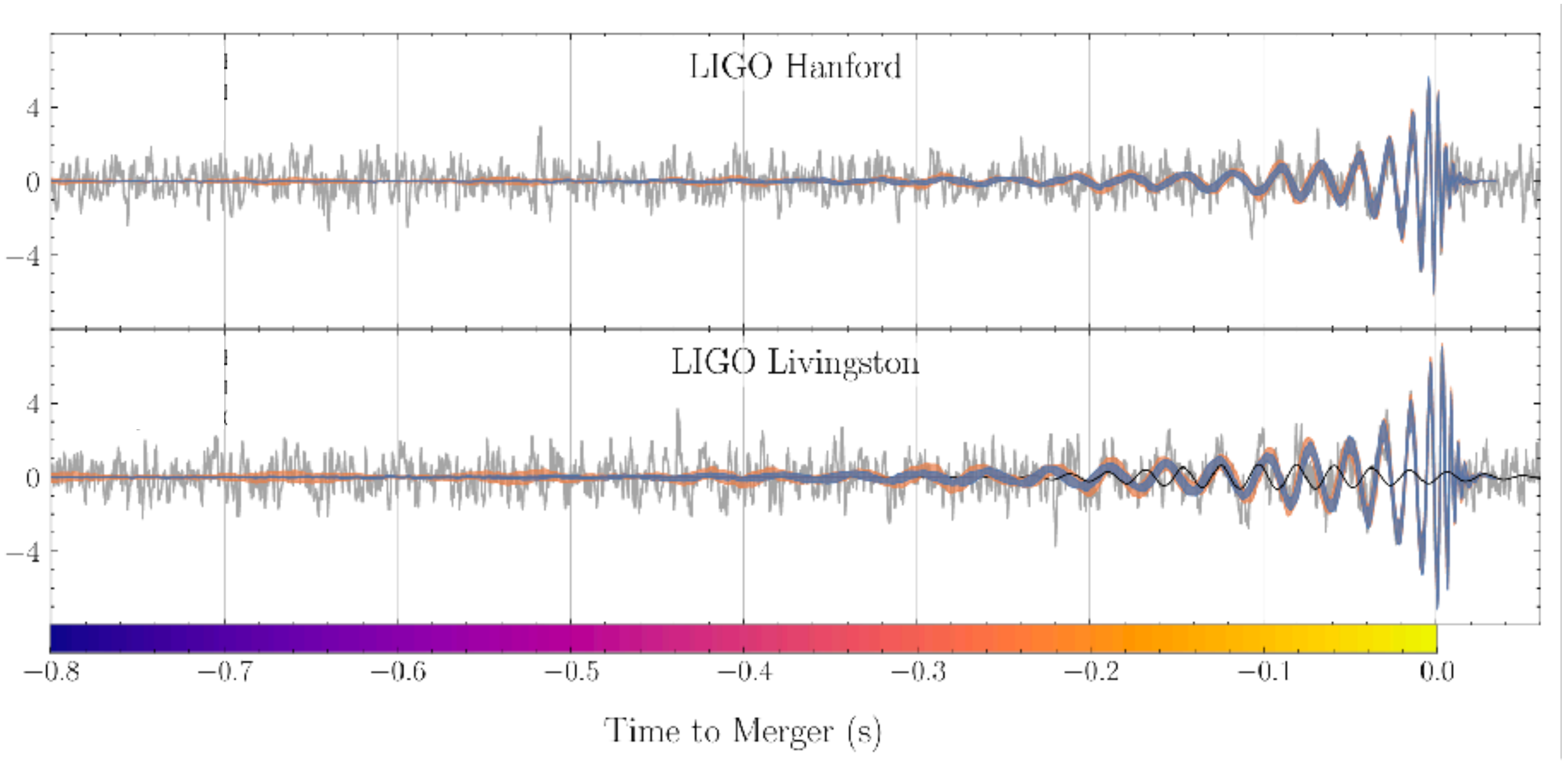
MIAPbP workshop “Build big or build smart”

What are gravitational waves?

Black holes merge
(or neutron stars) → Emit gravitational wave → Measured in detectors



Universe



Described by physics parameters

$$\theta \in \mathbb{R}^{15}$$

Masses, spins, sky position, orientation

Measured data d

Gupte+, arXiV 2024

Why is GW science exciting?

New window into the universe!

Direct observation of two merging black holes

Physical Review Letters

FEATURED IN PHYSICS | EDITORS' SUGGESTION | OPEN ACCESS

Observation of Gravitational Waves from a Binary Black Hole Merger

B. P. Abbott[1], R. Abbott[1], T. D. Abbott[2], M. R. Abernathy[1], F. Acernese[3,4], K. Ackley[5], C. Adams[6], T. Adams[7], P. Addesso[3] et al. (LIGO Scientific Collaboration and Virgo Collaboration)

Phys. Rev. Lett. 116, 061102 – Published 11 February, 2016

nature

Article | Published: 23 October 2019

Identification of strontium in the merger of two neutron stars

Science

Light curves of the neutron star merger GW170817/SSS17a: Implications for r-process nucleosynthesis

Heavy elements (gold, platinum, uranium, …) on Earth likely created by neutron star mergers

Estimating the expansion rate of the Universe

nature

Letter | Published: 16 October 2017

A gravitational-wave standard siren measurement of the Hubble constant

The LIGO Scientific Collaboration and The Virgo Collaboration, The 1M2H Collaboration, The Dark Energy Camera GW-EM Collaboration and the DES Collaboration, The DLT40 Collaboration, The Las Cumbres Observatory Collaboration, The VINROUGE Collaboration & The MASTER Collaboration

Nature 551, 85–88 (2017) | Cite this article

Testing matter under extreme conditions

Physical Review Letters

EDITORS' SUGGESTION

GW170817: Measurements of Neutron Star Radii and Equation of State

B. P. Abbott[1], R. Abbott[1], T. D. Abbott[2], F. Acernese[3,4], K. Ackley[5], C. Adams[6], T. Adams[7], P. Addesso[8], R. X. Adhikari[1] et al. (The LIGO Scientific Collaboration and the Virgo Collaboration)

Phys. Rev. Lett. 121, 161101 – Published 15 October, 2018

THE ASTROPHYSICAL JOURNAL LETTERS

INTEGRAL Detection of the First Prompt Gamma-Ray Signal Coincident with the Gravitational-wave Event GW170817

V. Savchenko, C. Ferrigno, E. Kuulkers, A. Bazzano, E. Bozzo, S. Brandt, J. Chenevez,

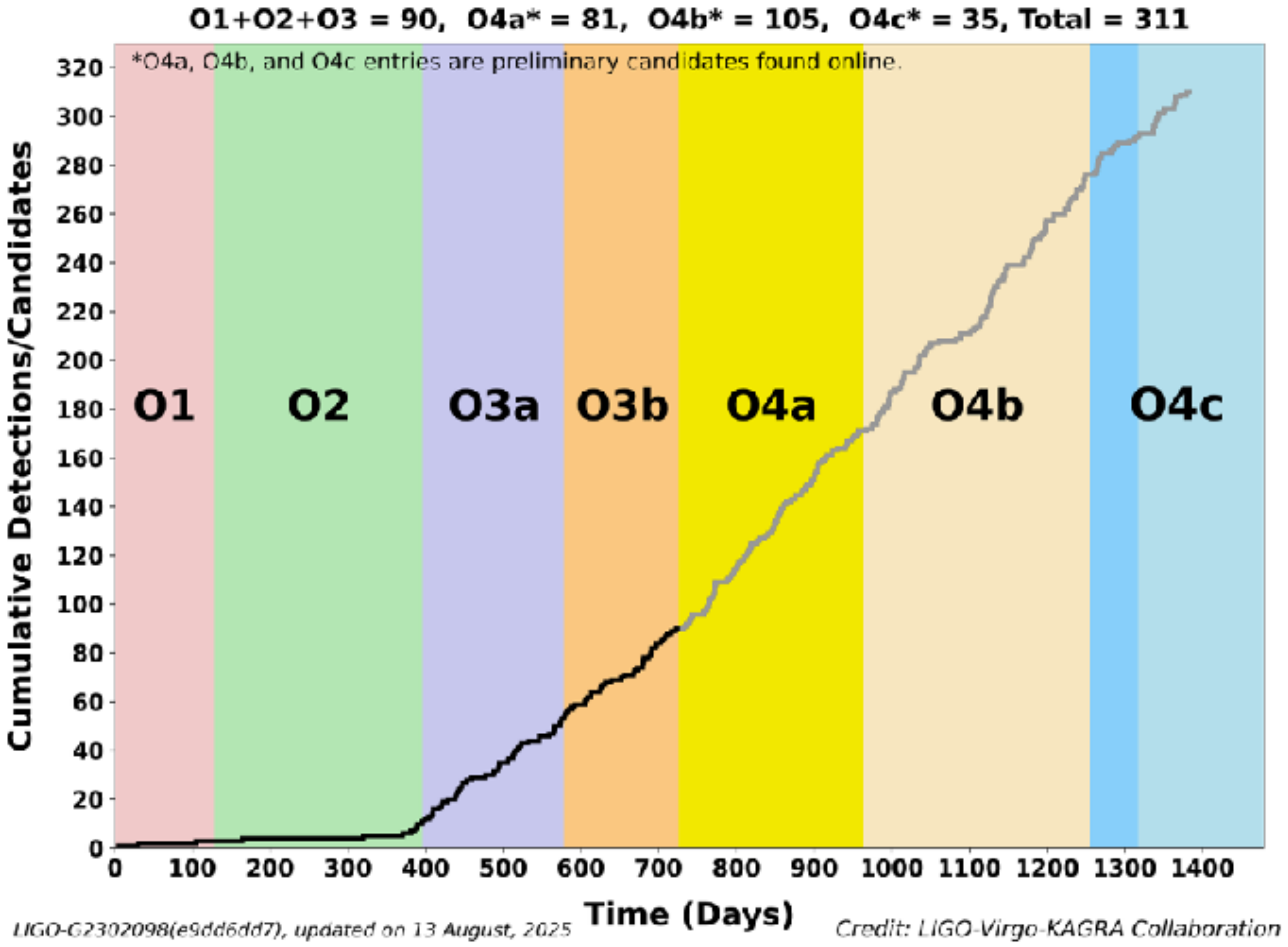
Investigating fundamental concepts:
Lorentz invariance, speed of gravity, equivalence principle

Machine Learning in GW science

- Increasing number of detections: $\sim 1000 \times$ until 2030s

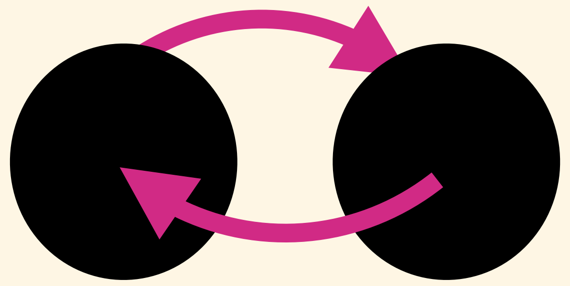
- We need efficient analysis methods for:
  - Signal detection
  - Glitch classification, glitch removal
  - Parameter estimation

    $\rightarrow$ **Standard methods need hours - days for a single event**



Build smart vs. build big
for the example of GW parameter estimation

What is parameter estimation?



Physics parameters

$\theta \in \mathbb{R}^{15}$

Masses, spins, sky position, orientation

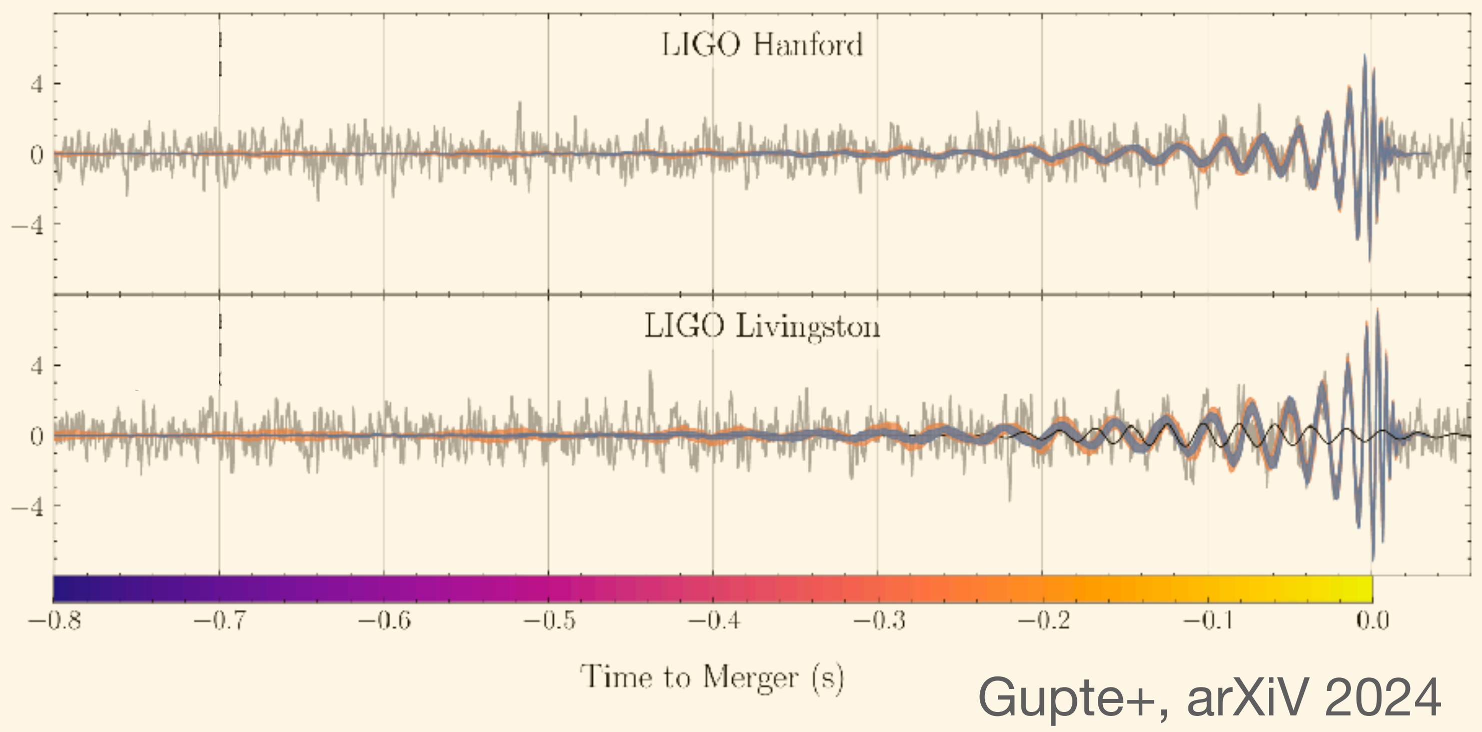
Universe

Simulator

Posterior Estimation

$p(\theta | d)$

Measured data d



Gupte+, arXiV 2024

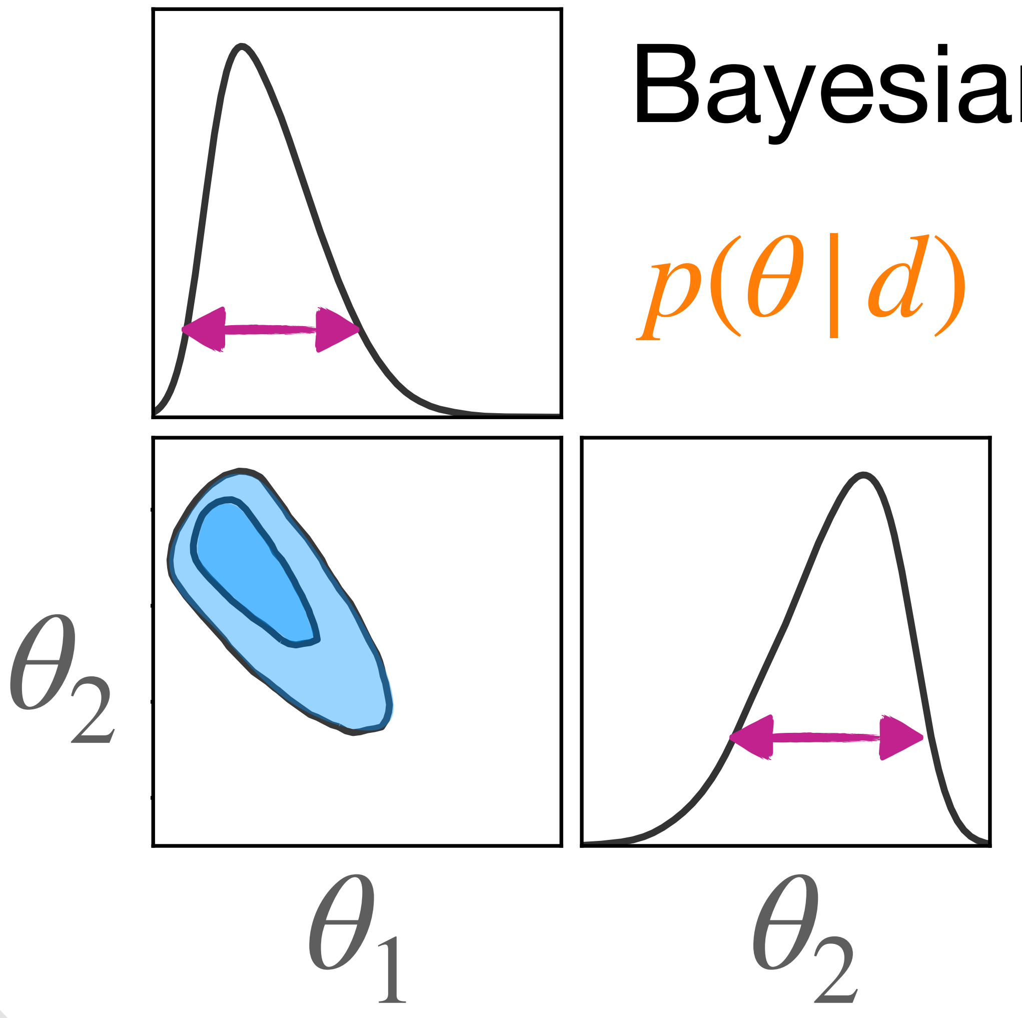
Inverse direction with Bayesian inference:

Likelihood

Prior belief

$$p(\theta | d) = \frac{p(d | \theta)}{p(d)} p(\theta)$$

Posterior



Bayesian posterior

$p(\theta | d)$

Uncertainties

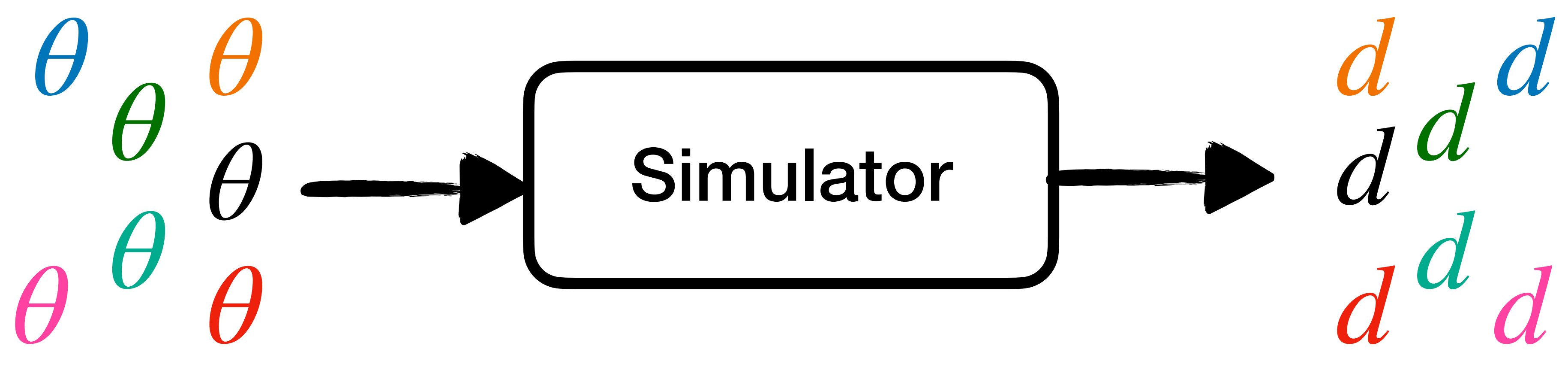
Correlations

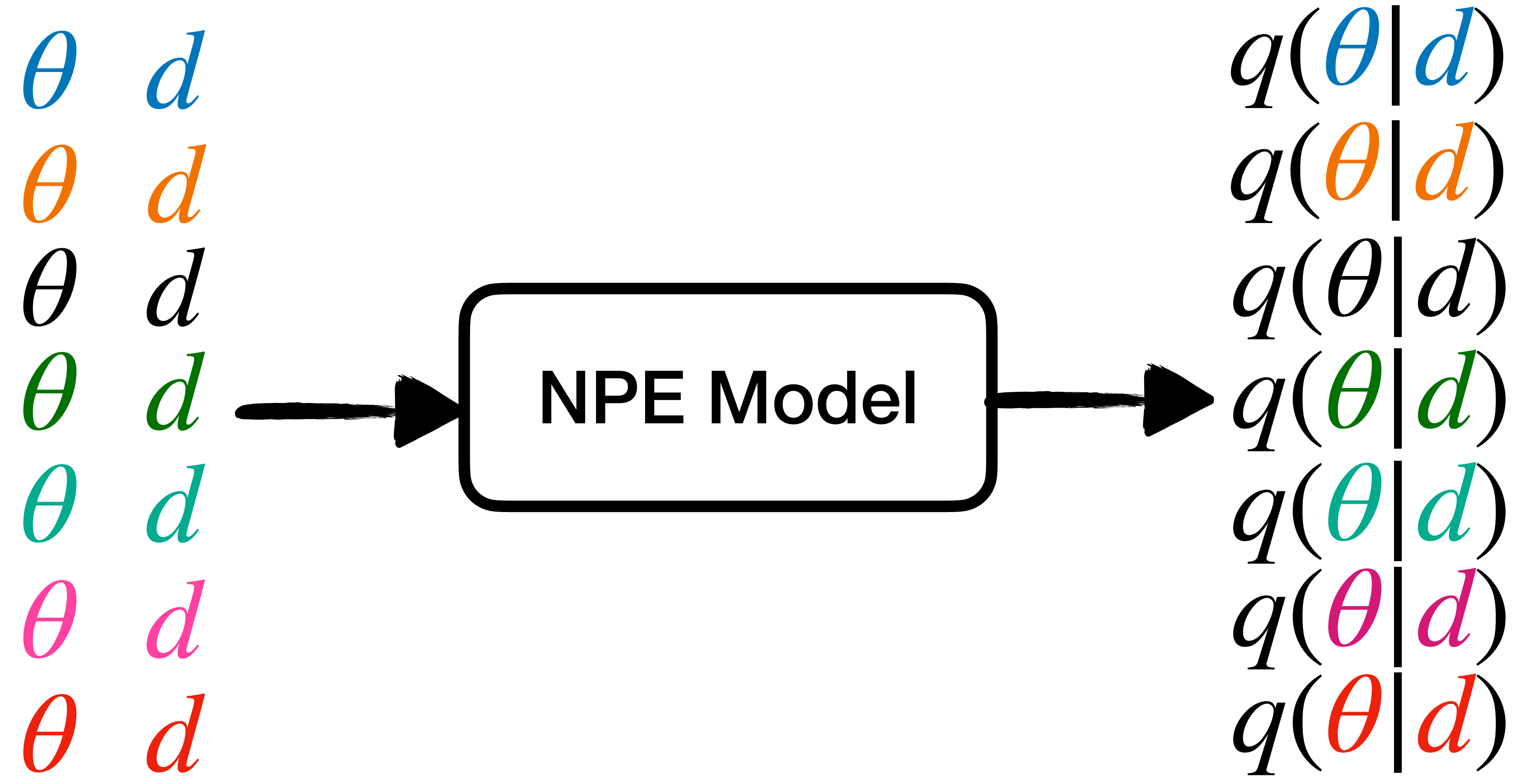
Simulation-based inference (SBI) to the rescue

***specifically: Neural Posterior Estimation (NPE)**

Idea: Train model to invert simulator

1. Generate large training data set with simulator

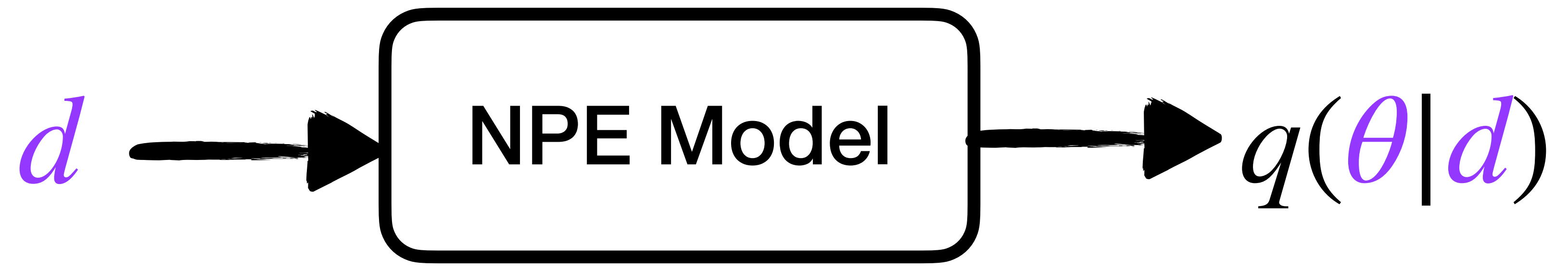




2. Train machine learning model q to approximate posterior

$$p(\theta \,|\, d) \approx q(\theta \,|\, d)$$

3. Evaluate model on measurement d to obtain $q(\theta|d)$

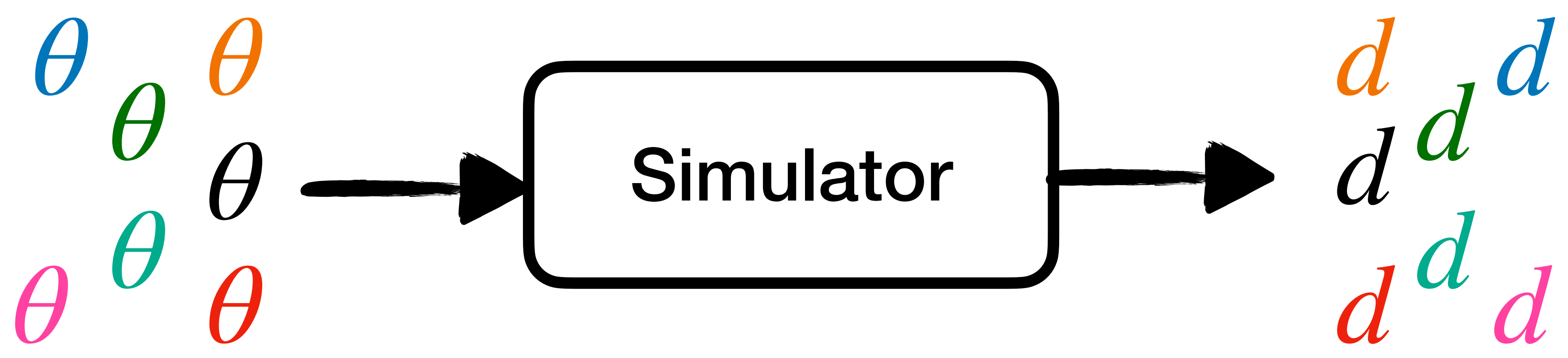


Simulation-based inference (SBI) to the rescue

***specifically: Neural Posterior Estimation (NPE)**

Idea: Train model to invert simulator

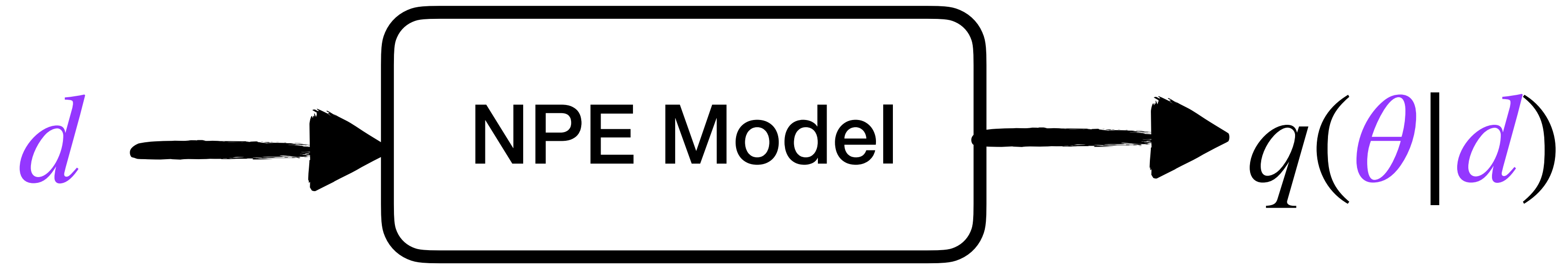
1. Generate large training data set with simulator



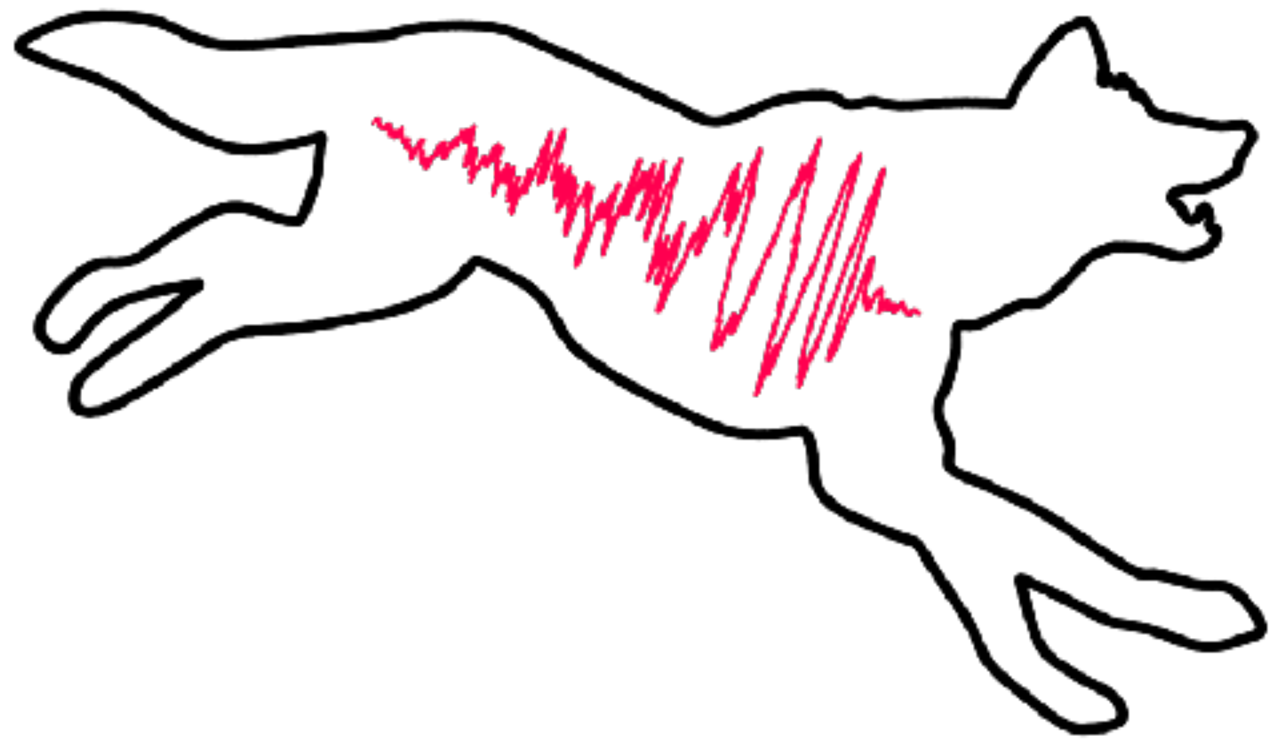
2. Train machine learning model q to approximate posterior

$$p(\theta \mid d) \approx q(\theta \mid d)$$

3. Evaluate model on measurement d to obtain $q(\theta|d)$



Amortization with NPE:
- Train once
- Evaluate on many events

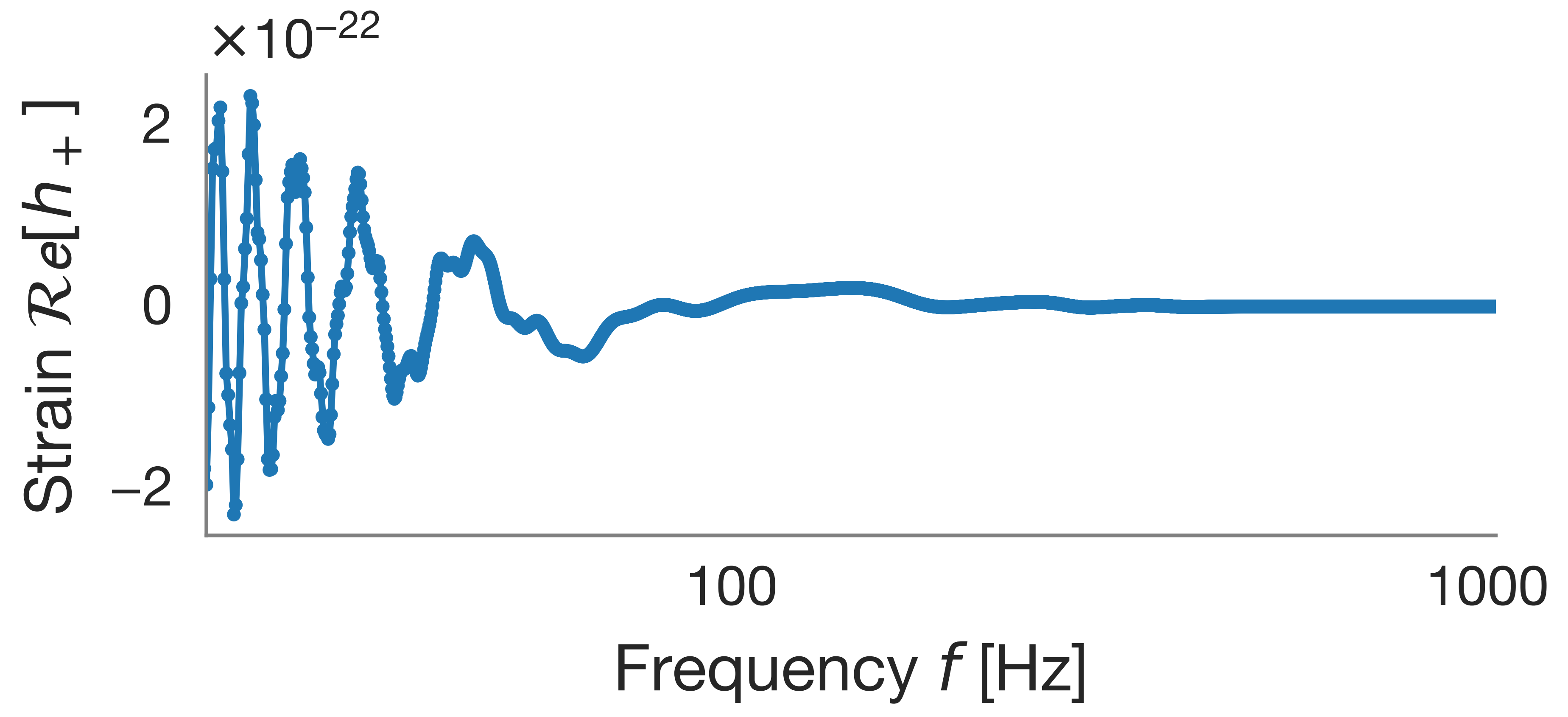


NPE for gravitational waves

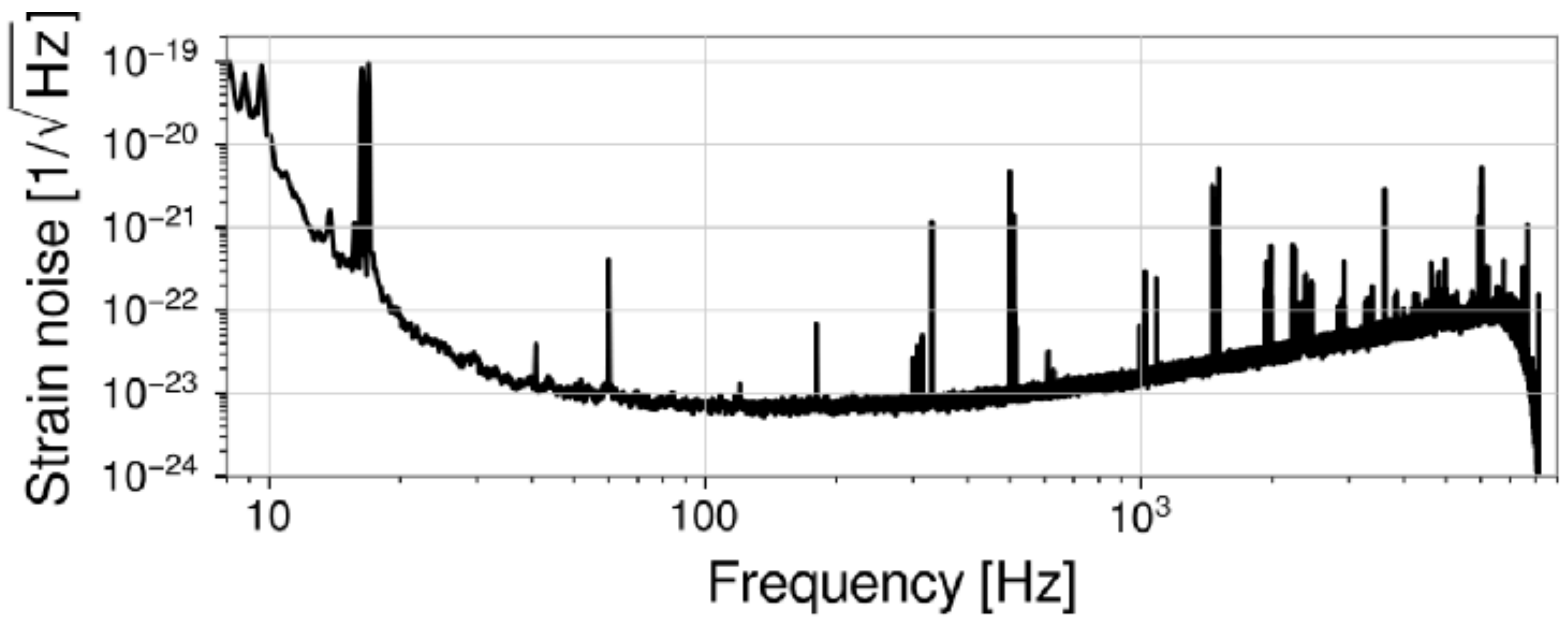
DINGO = Deep INference for Gravitational wave Observations

NPE for gravitational waves: Overview

- Generate simulated waveforms: $\theta \sim p(\theta), \;\; h = \text{simulator}(\theta) \;\rightarrow \{\theta, h\}$



- Add realistic noise to the waveform
  $S_n(f) \rightarrow \{S_n^{(i)}(f)\}$
  1. Sample noise $n^{(i)} \sim \mathcal{N}(0, S_n^{(i)})$
  2. Add to waveform $d^{(i)} = h(\theta^{(i)}) + n^{(i)}$



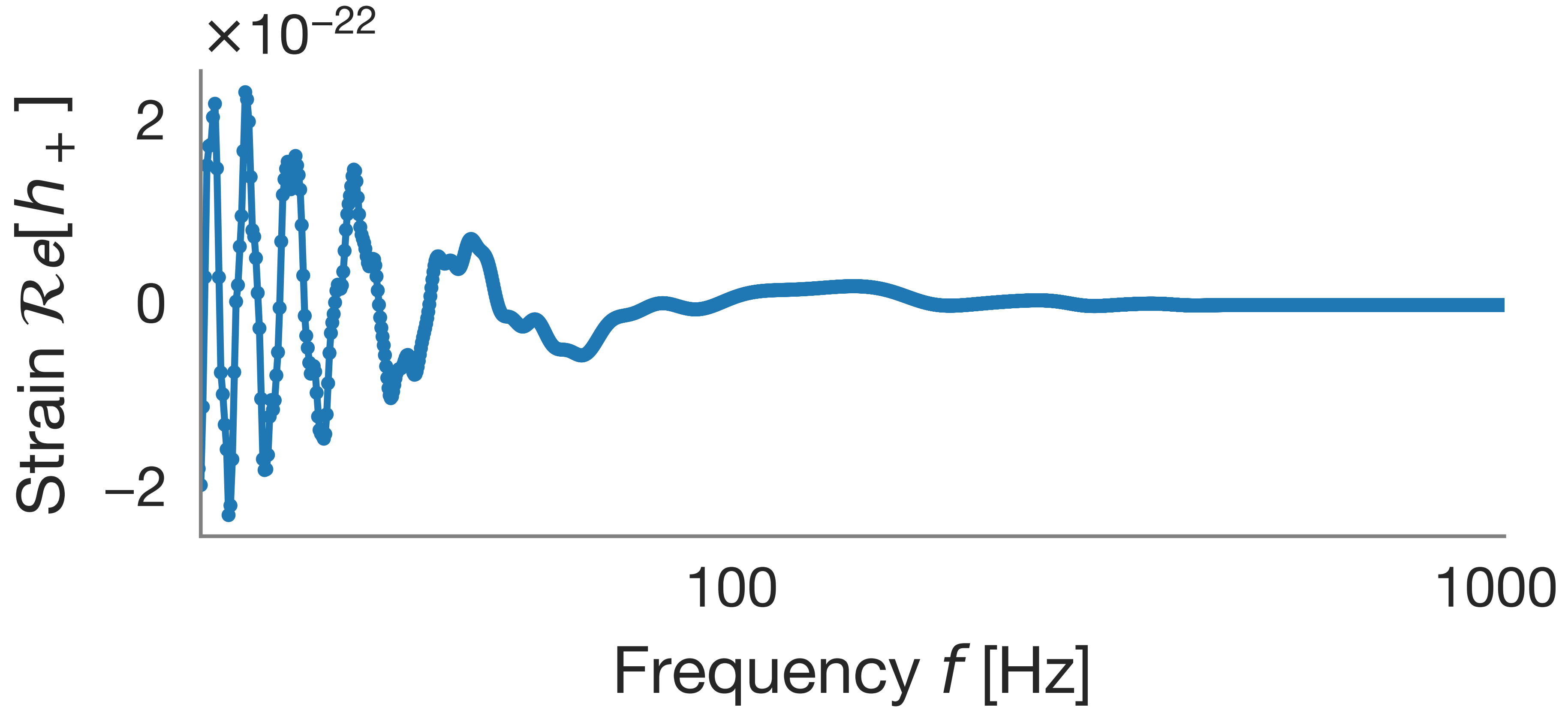
- Train density estimator

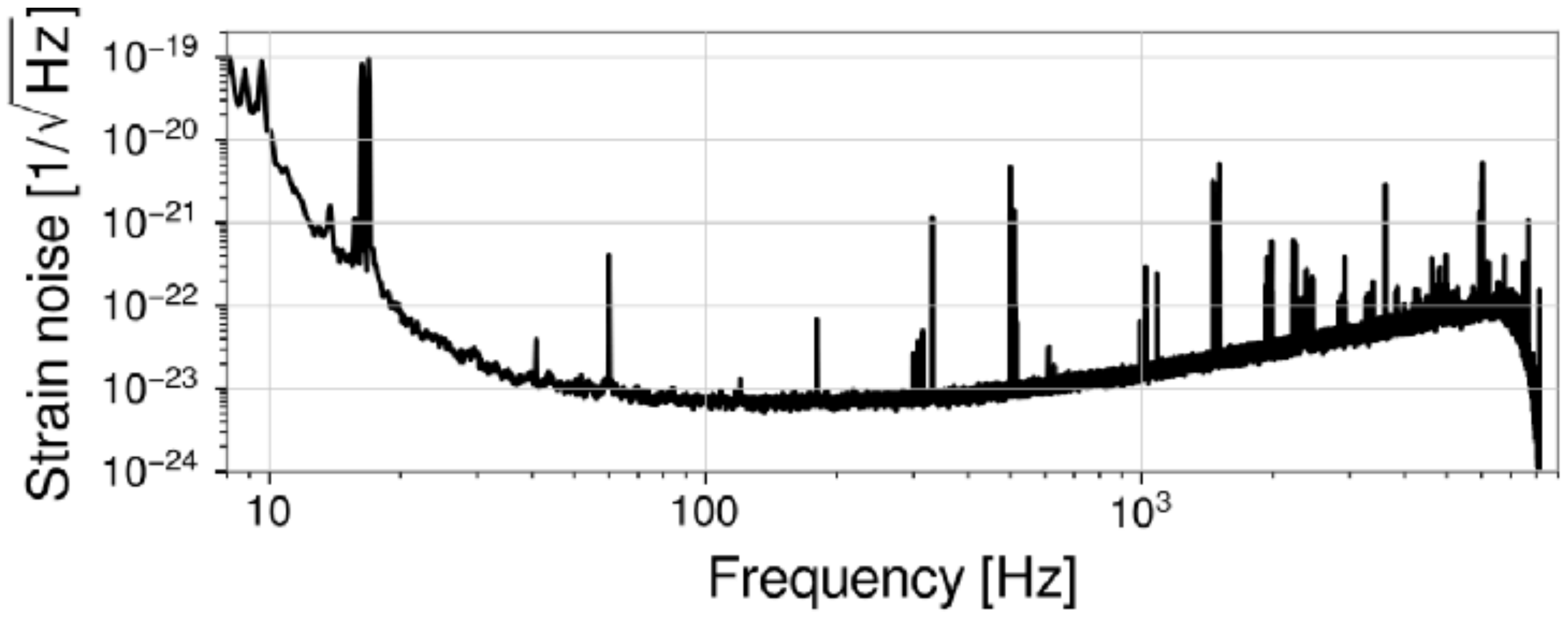
Dax+, PRL 2021

Changing PSDs

- Detector noise $S_n(f)$ varies from event to event
  $\rightarrow$ augment training to include collection of PSDs $S_n(f) \rightarrow \{S_n^{(i)}(f)\}$

1. **Sample PSD** $\quad S_n^{(i)} \sim p(S_n)$
2. Generate noise $\quad n^{(i)} \sim \mathcal{N}(0,\, S_n^{(i)})$
3. Add signal $\quad d^{(i)} = h(\theta^{(i)}) + n^{(i)}$



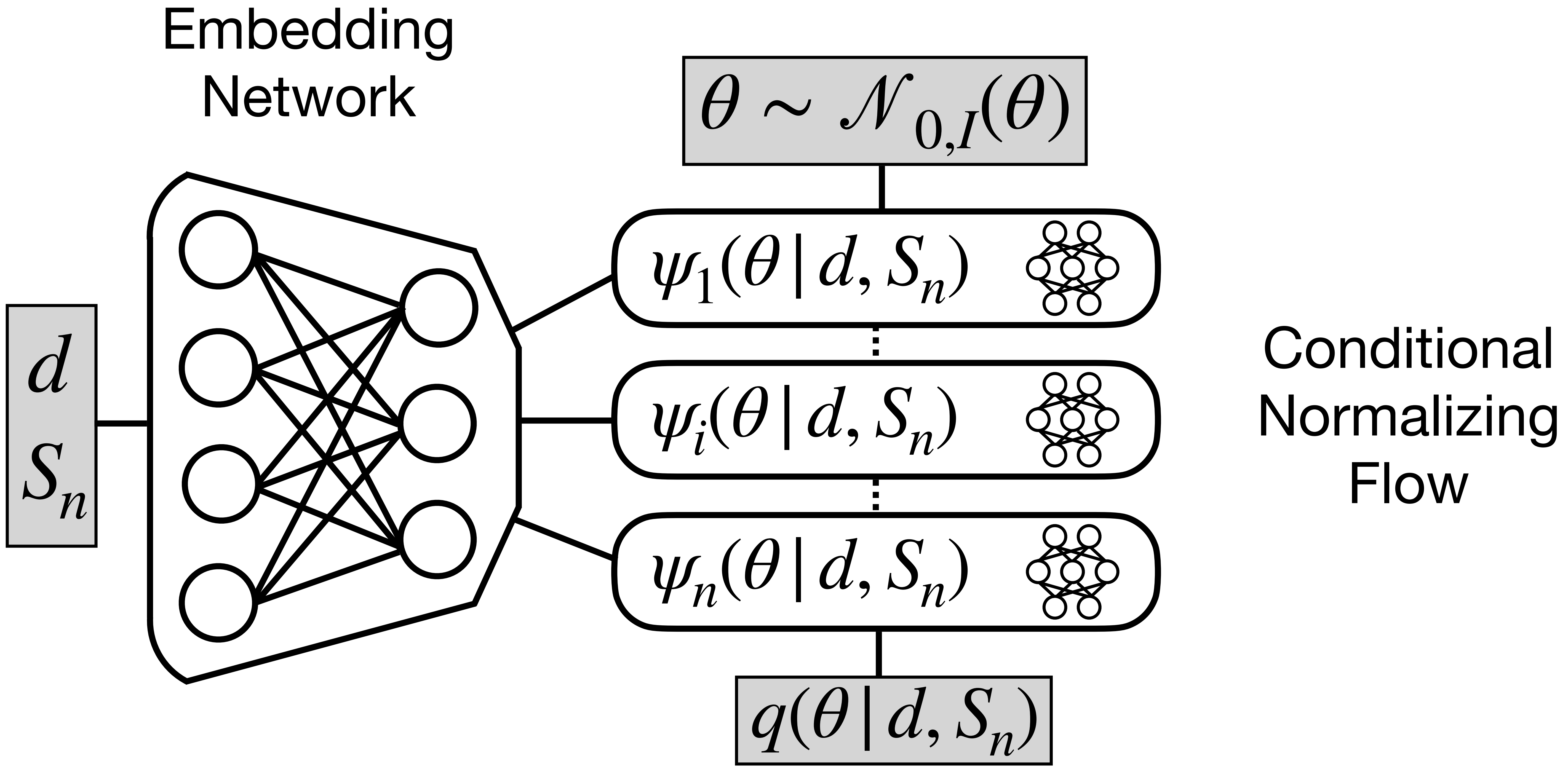


Dax+, PRL 2021

Training the model

- Provide data d and noise curve S_n to embedding network
- Train with negative log-likelihood loss

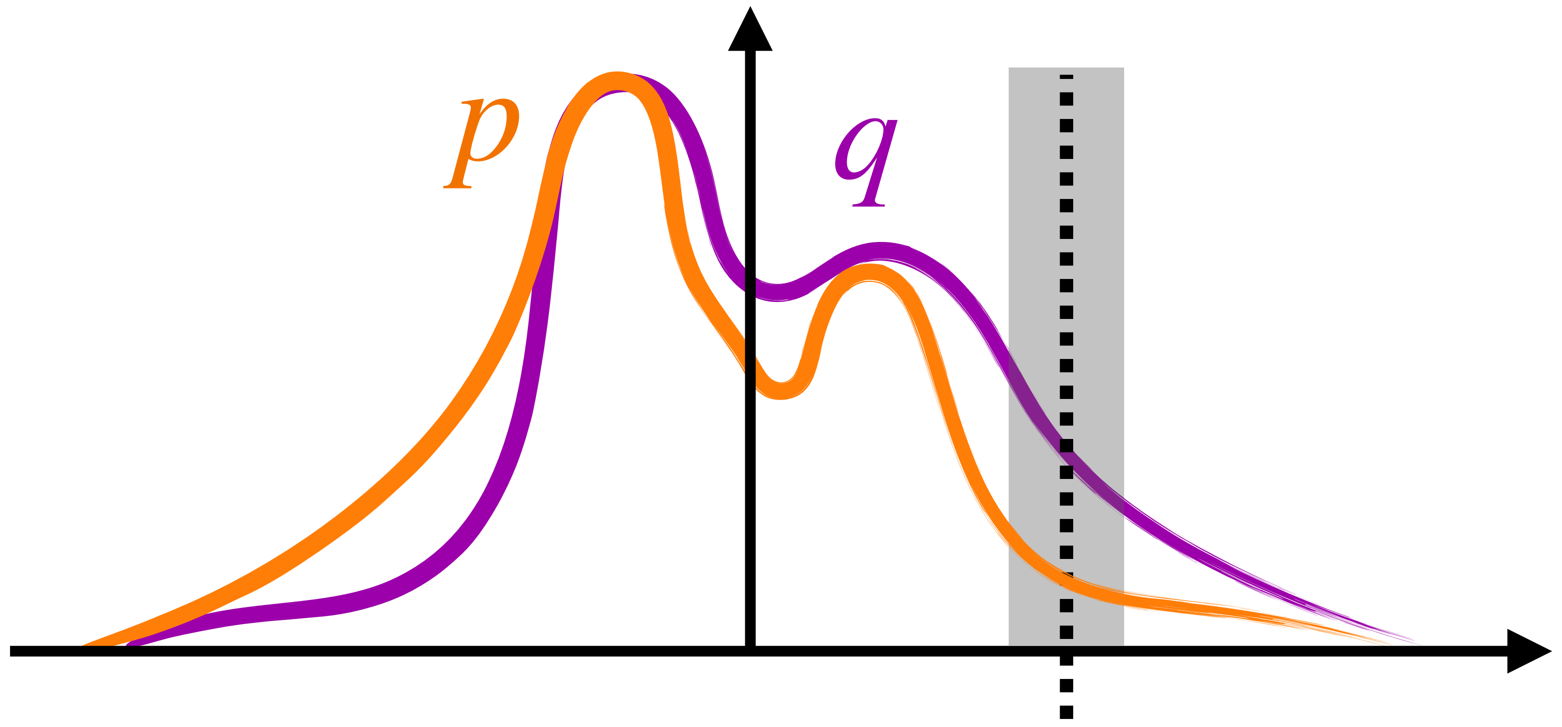
$$\mathscr{L} = -\frac{1}{N} \sum_{\substack{\theta^{(i)} \sim p(\theta) \\ d^{(i)} \sim p(d\,|\,\theta^{(i)})}} \log q(\theta^{(i)}\,|\,d^{(i)})$$



Dax+, PRL 2021

Importance Sampling (DINGO-IS)

- Importance sampling to validate model & reweigh samples towards true posterior



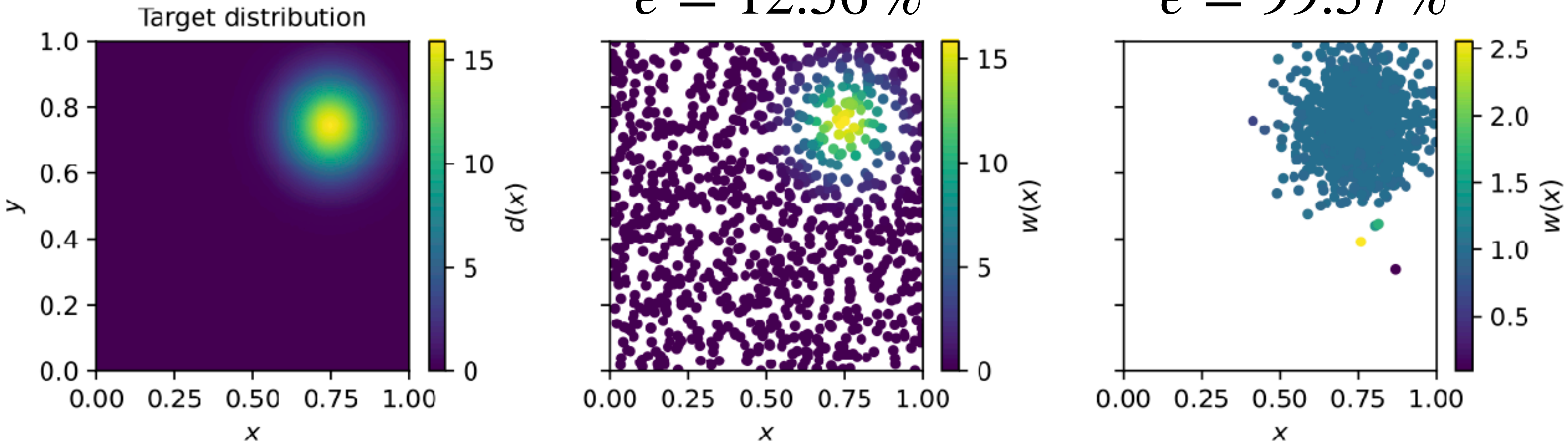
$$\frac{p(\theta|d)}{q(\theta|d)} \propto w = \frac{p(d|\theta)\,p(\theta)}{q(\theta|d)}$$

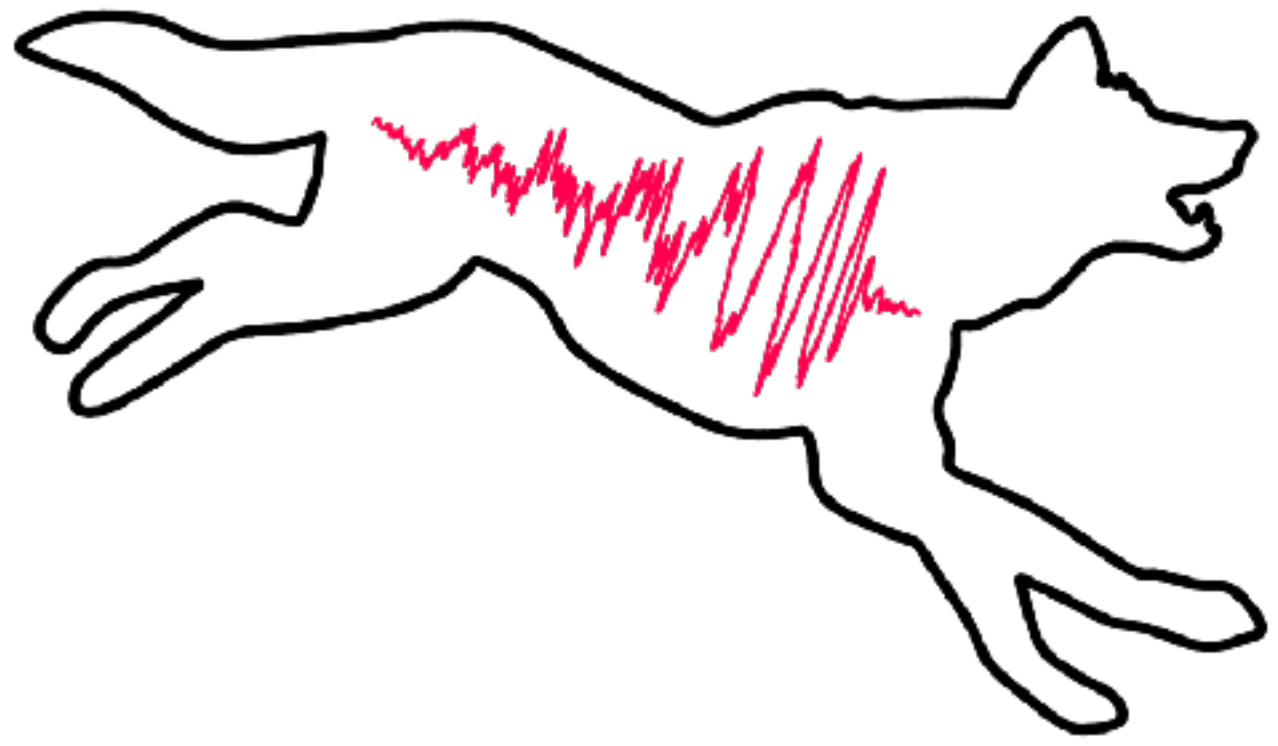
$q(\theta|d) > p(\theta|d)$
Low w

- Sample efficiency

$$\epsilon = \frac{1}{N}\frac{\left(\sum_{i=1}^{N} w_i\right)^2}{\sum_{i=1}^{N} w_i^2}$$

$\epsilon = 12.56\,\%$ $\epsilon = 99.57\,\%$



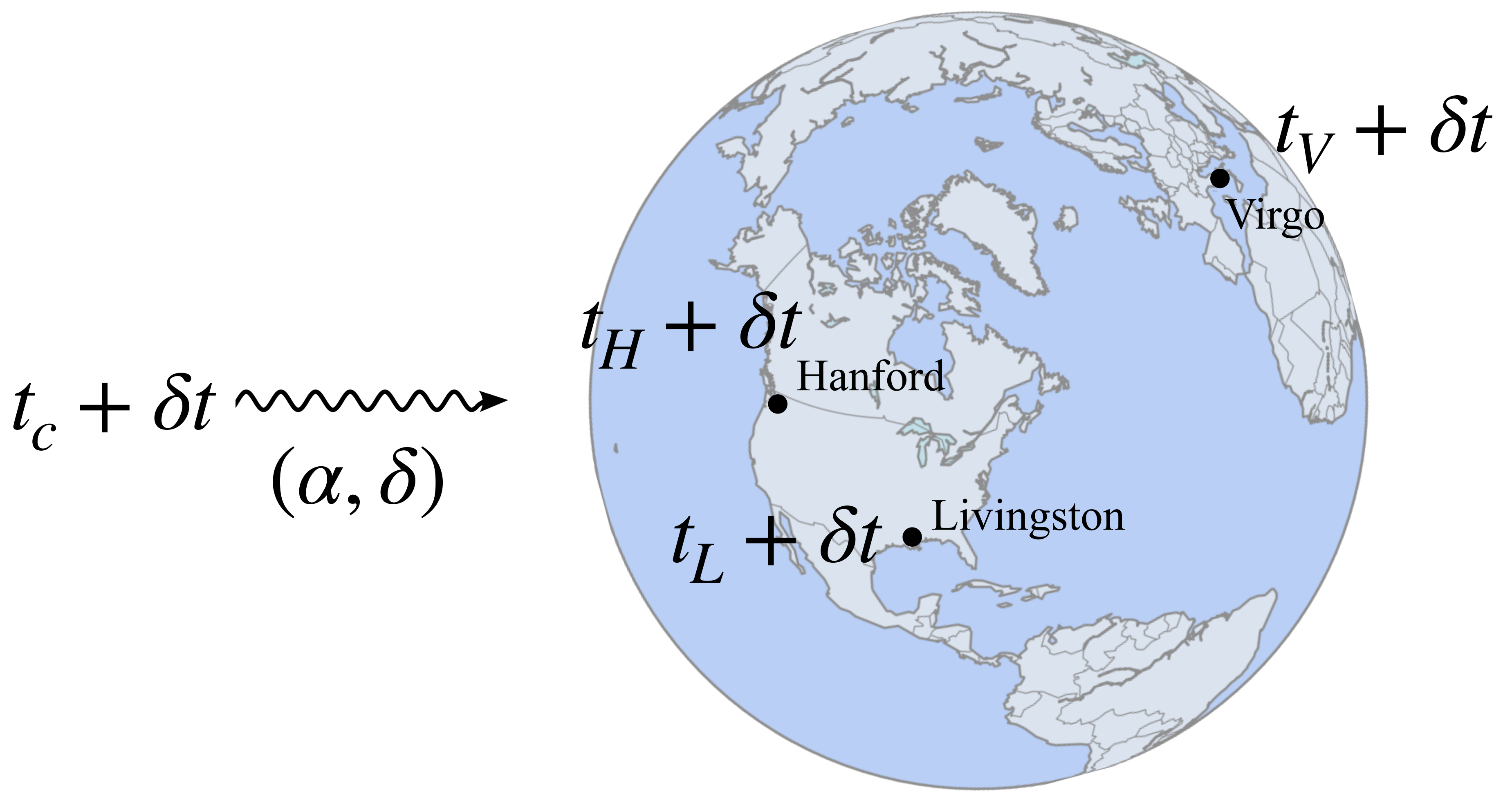


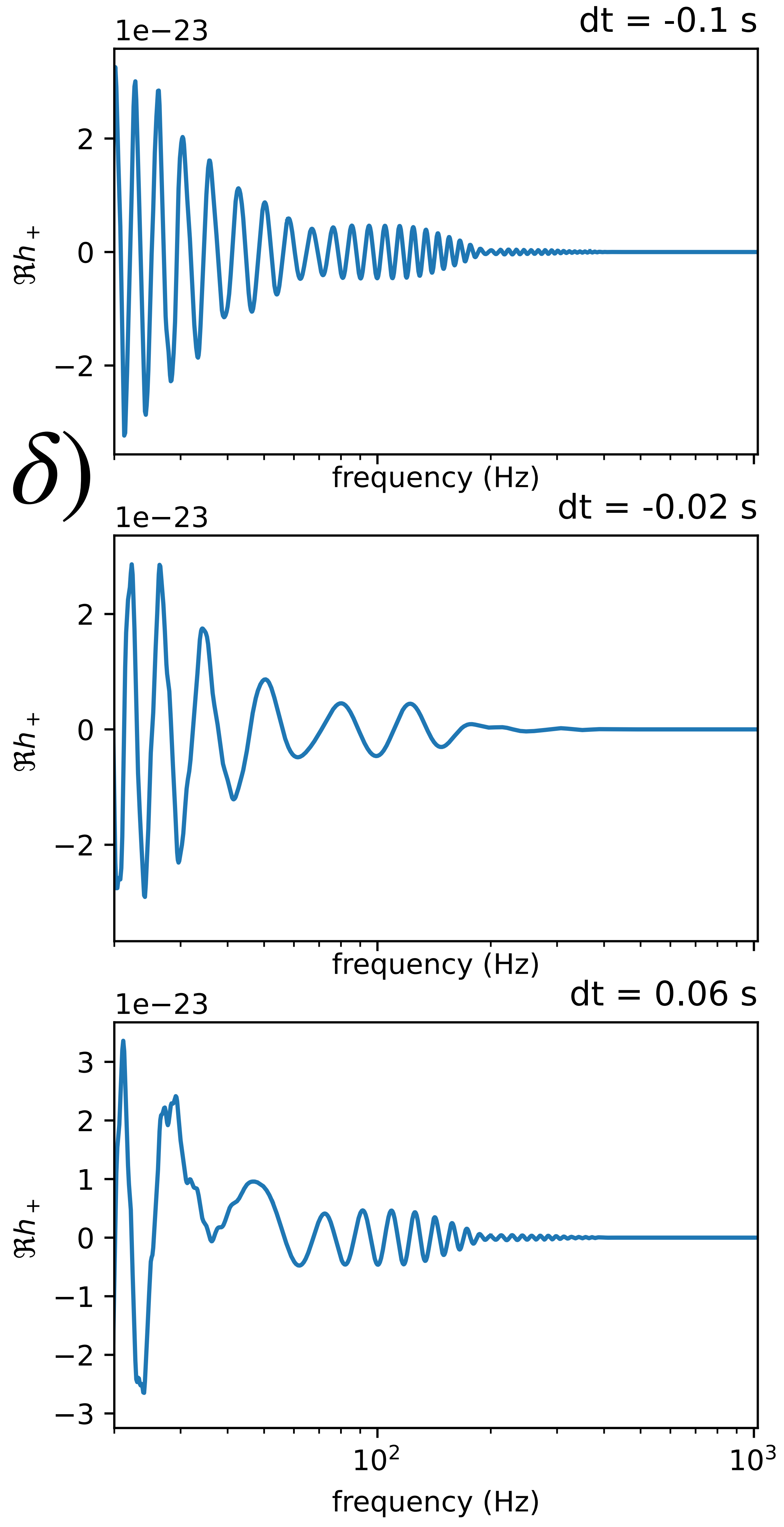
Building smart: Group-equivariant NPE

Symmetries in the data

- Posterior should be …
  … equivariant wrt. overall coalescence time t_c
  … approximately equivariant wrt. changes in sky position (α, δ)

- Frequency-domain data
  $\rightarrow$ time shift = multiplication by $e^{2\pi i f \cdot \delta t}$





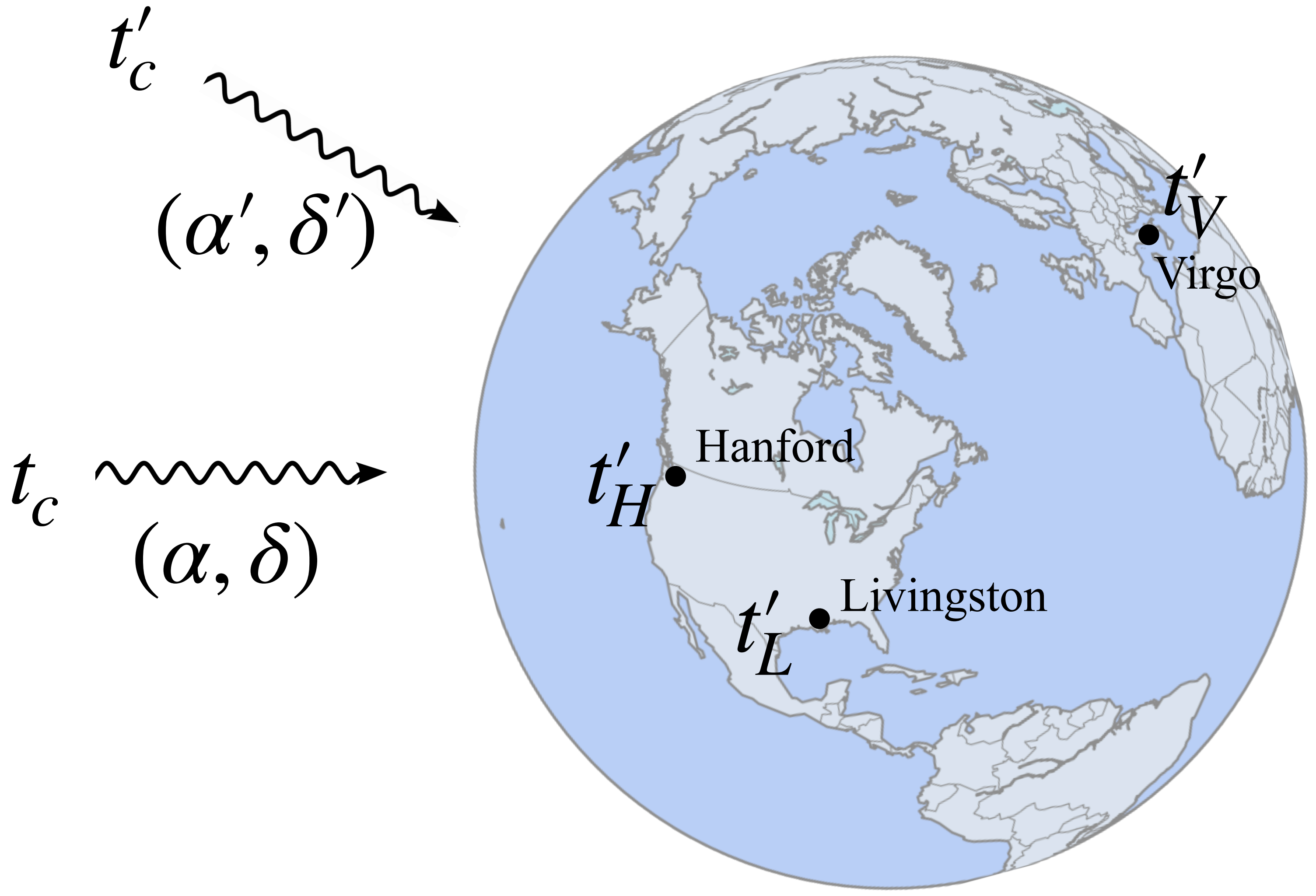
Dax+, ICLR 2022

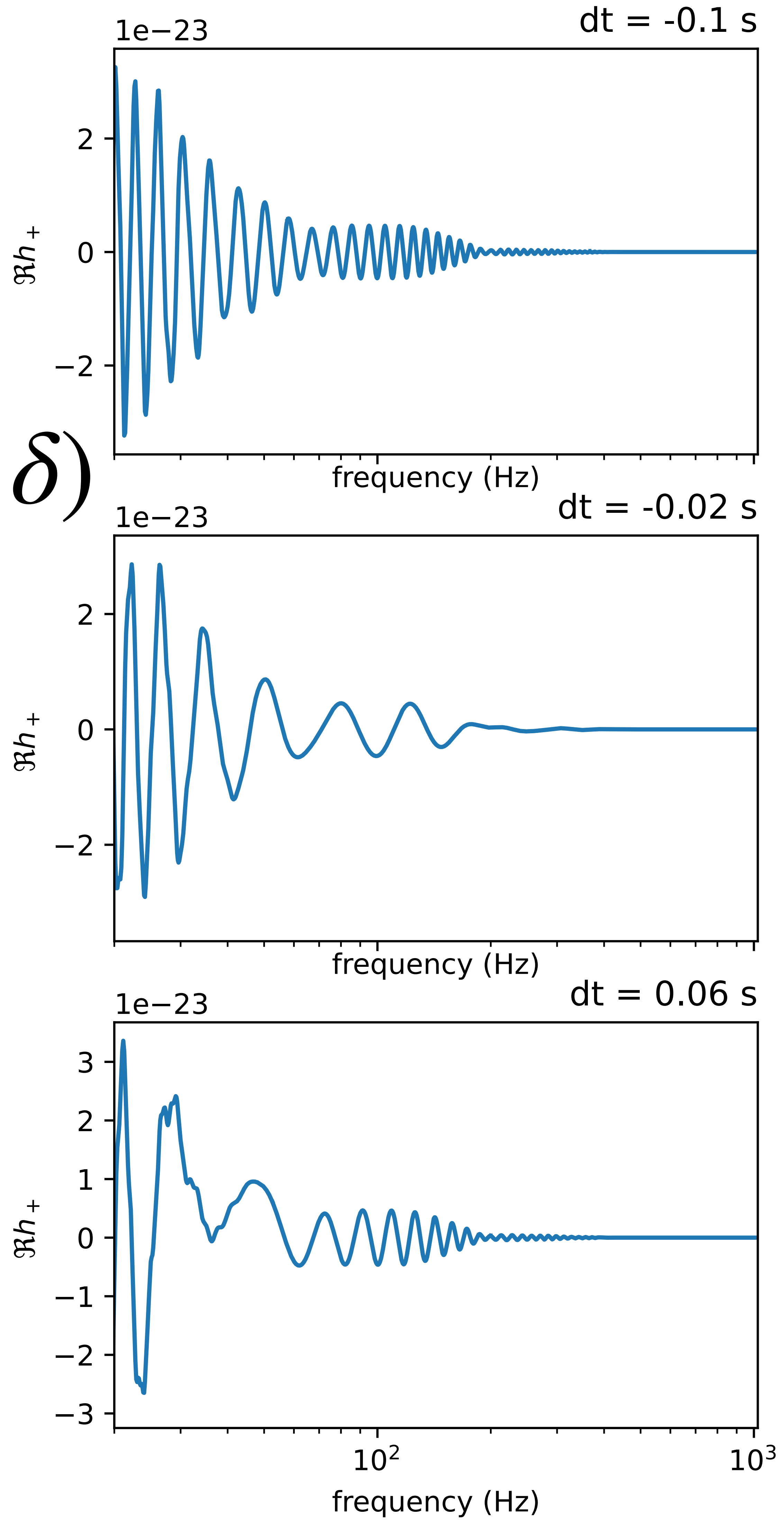
Symmetries in the data

- Posterior should be …
  … equivariant wrt. overall coalescence time t_c
  … approximately equivariant wrt. changes in sky position (α, δ)

- Frequency-domain data
  $\rightarrow$ time shift = multiplication by $e^{2\pi i f \cdot \delta t}$

- Hard to learn just from data





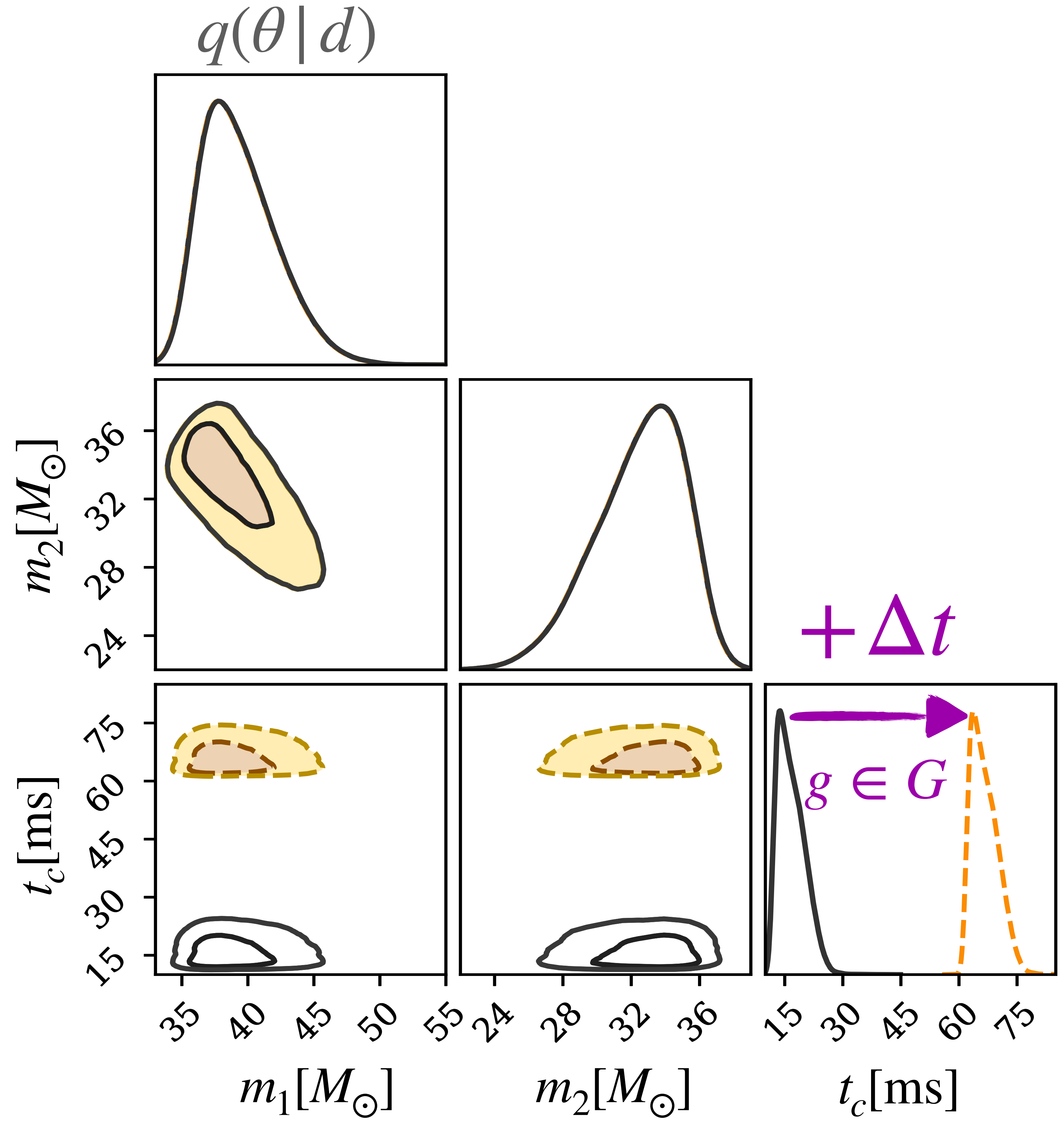
Dax+, ICLR 2022

NPE with symmetries: Group-equivariant NPE

- Challenging problem:

$$p(\theta\,|\,d) = p(g\theta\,|\,T_g d)\,|\det T_g\,|$$

- Model can learn symmetries via training data
- Enforce these symmetries in the model
- **Group-equivariant NPE (GNPE):**
  - Standardize data to remove time shifts
  - Requires iterative inference to find correct standardization (→ slow)



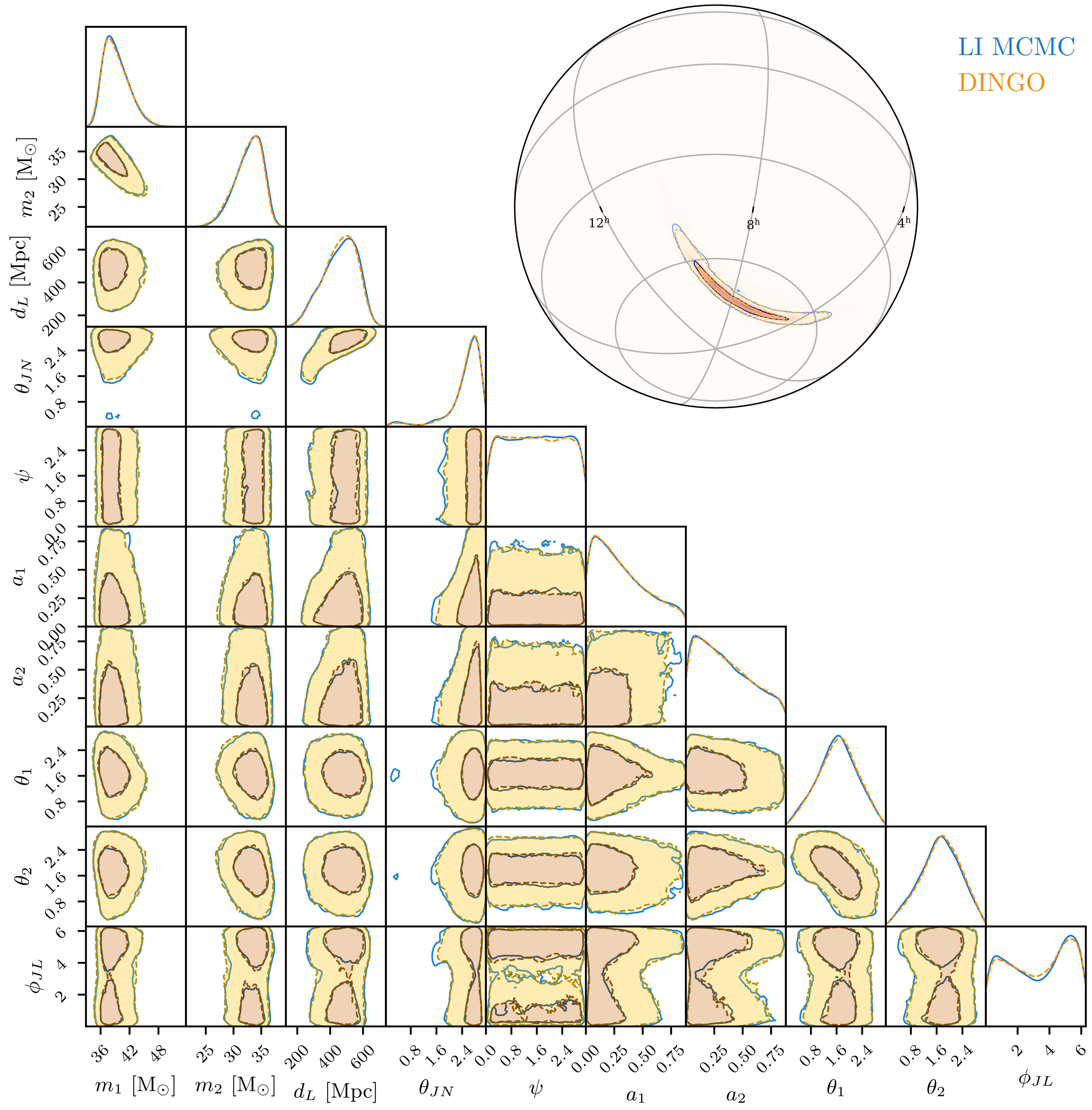
Dax+, ICLR 2022

Results with GNPE

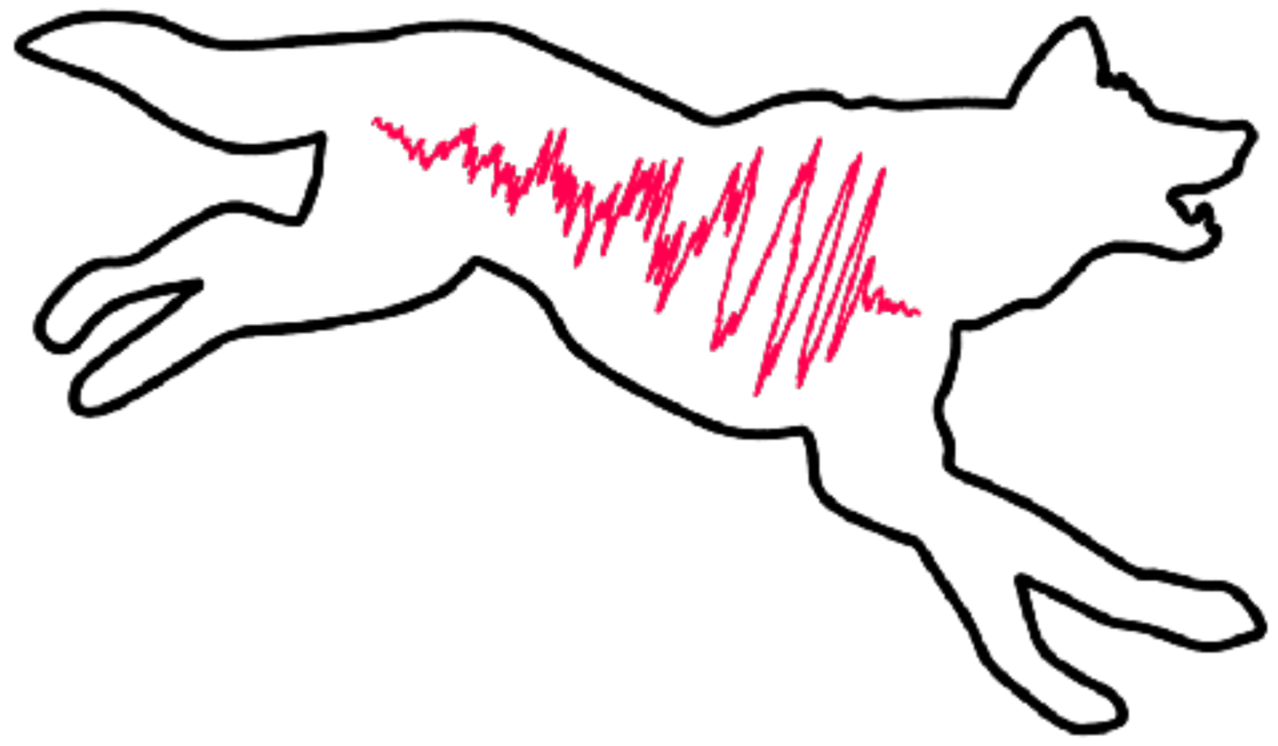
- Very good agreement with standard techniques

- $\sim 10^7$ training examples
  $\sim 10^8$ network parameters

- Training: $\sim$ few days
  Inference: $\sim$ few minutes



Dax+, PRL 2021

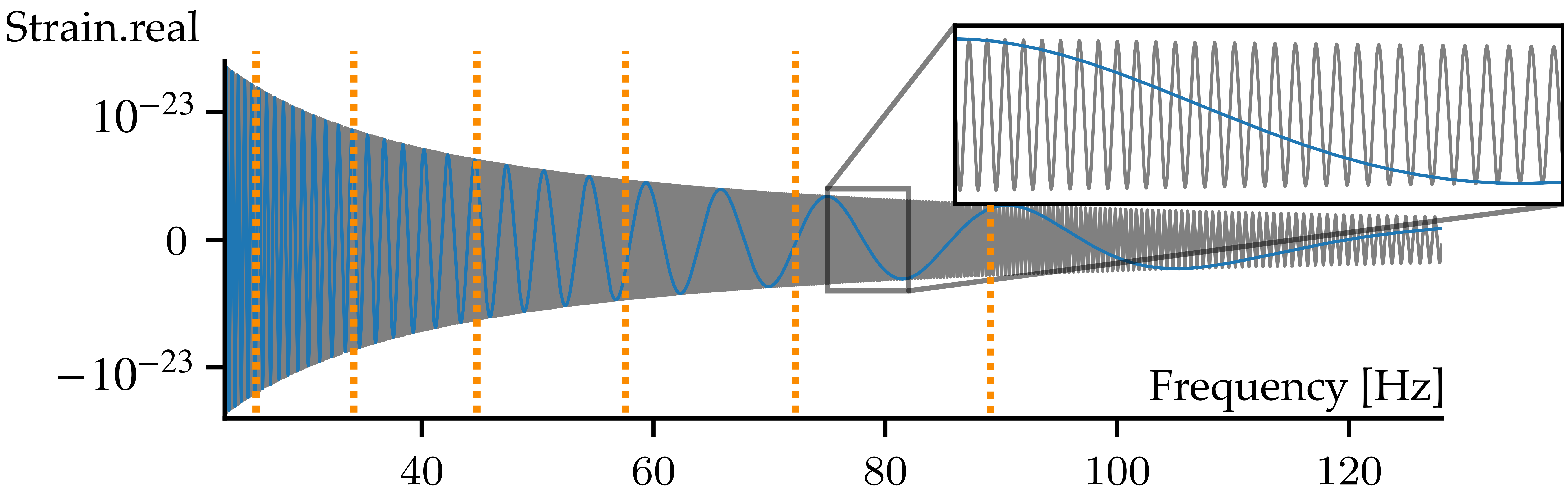


Building smart:
NPE for Binary Neutron Stars

Challenge: BNS data is high-dimensional

- BNS signals are very long → high-dimensional (10M+ frequency bins)
- Compress signal (assuming we know the chirp mass $\mathscr{M}_c$)
  1. Heterodyning: factor out overall phase $\propto (M_c^{\mathrm{est}} f)^{-5/3}$ (Cornish+, 2010)
  2. Multibanding: reduce resolution at higher f (Vinciguerra+, 2017)

$$\mathscr{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}$$



⇒ Loss-free compression by factor $100\times$

Dax+, Nature 2025

Challenge: BNS data is high-dimensional

- BNS signals are very long → high-dimensional (10M+ frequency bins)
- Compress signal **(assuming we know the chirp mass $\mathscr{M}_c$)**
    1. Heterodyning: factor out overall phase $\propto (M_c^{\mathrm{est}} f)^{-5/3}$ (Cornish+, 2010)
    2. Multibanding: reduce resolution at higher f (Vinciguerra+, 2017)

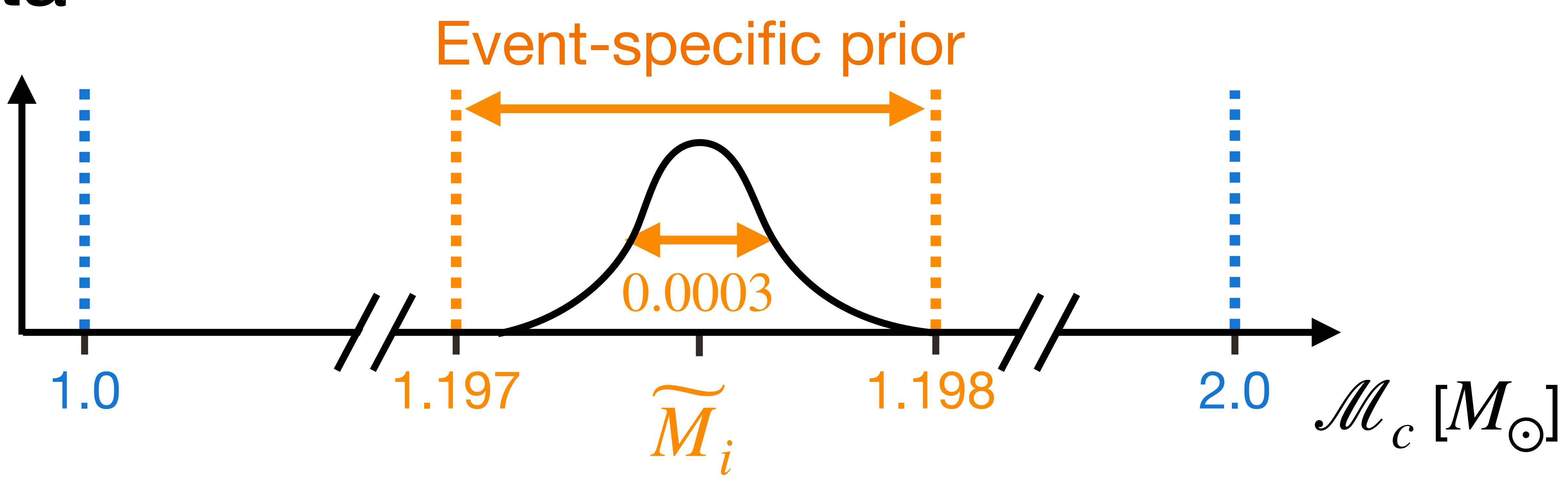
$$\mathscr{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}$$

Problem:
- Chirp mass is an (unknown) inference parameter
- Cannot use it for compression

Dax+, Nature 2025

Prior conditioning

- Chirp mass well constrained in BNS data
  - Between events: large range
  - Specific event: small range
- Prior-tunable network
  1. Sample training prior hierarchically $\widetilde{M}_i \sim \hat{p}(\widetilde{M}),\, \theta_i \sim p_{\widetilde{M}_i}(\theta)$
  2. Condition NPE on choice of prior $p(\theta \,|\, d, \widetilde{M})$
  3. Compress based on prior center $d \rightarrow d_{\widetilde{M}}$



For BNS:

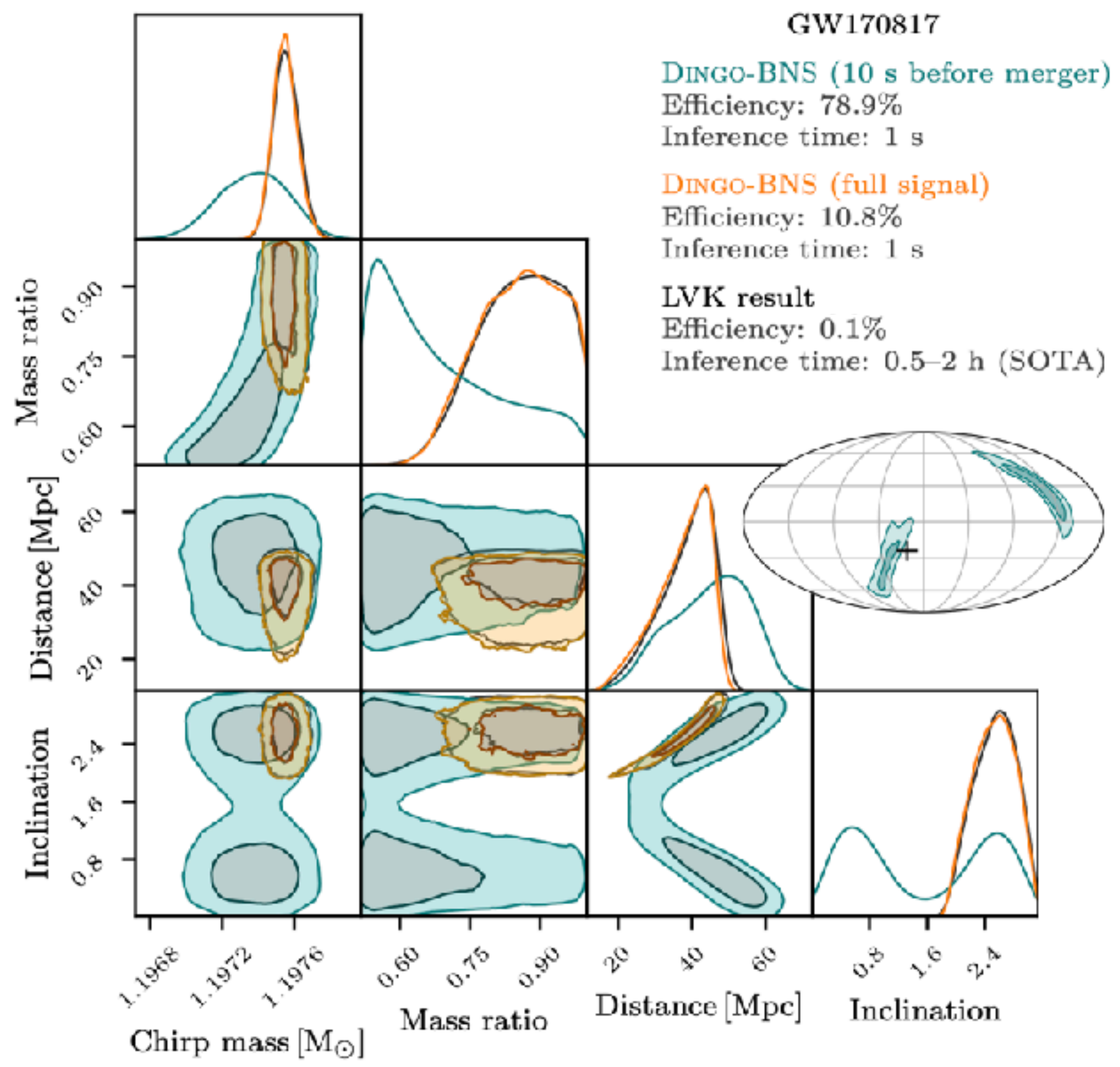
$\hat{p}(\widetilde{M}) = U[1.0, 2.0]\ \mathrm{M}_\odot$

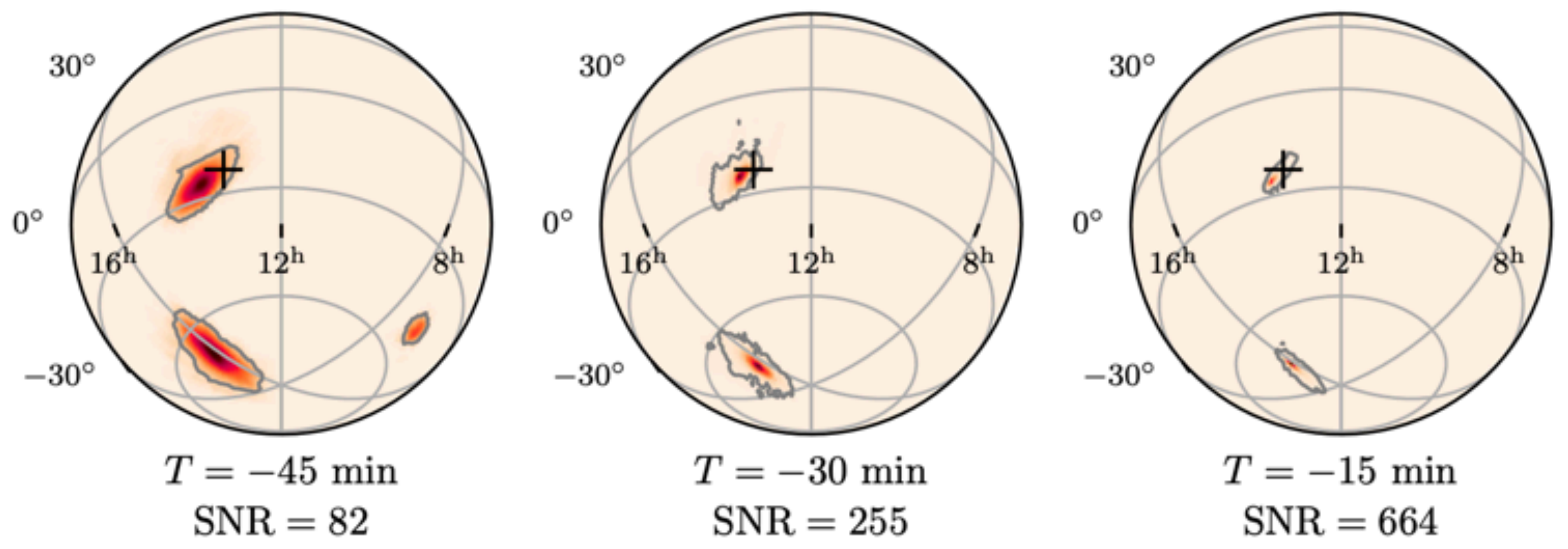
$p_{\widetilde{M}}(M_c) = U[\widetilde{M} - 0.005\, M_\odot, \widetilde{M} + 0.005\, M_\odot]$

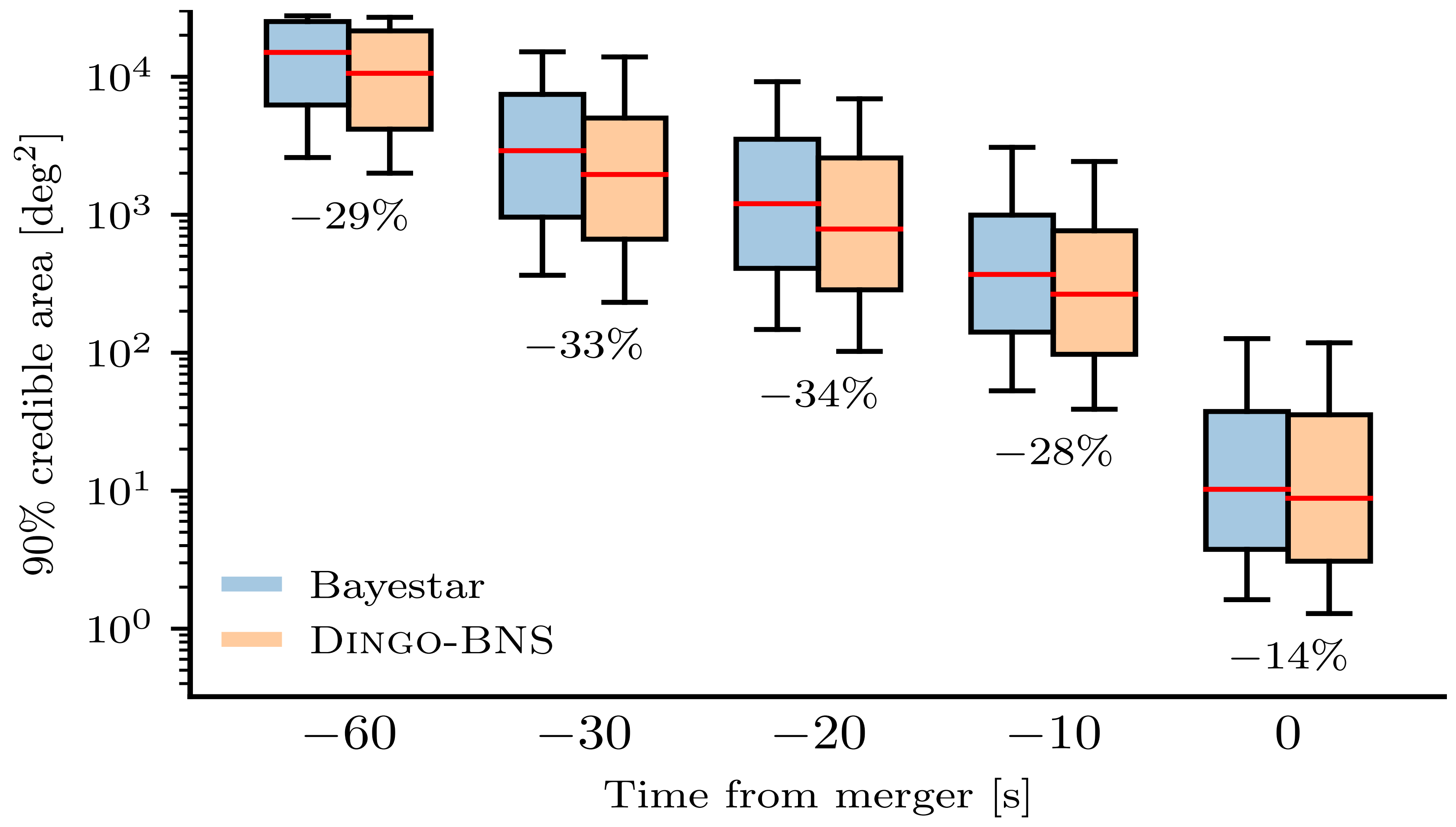
Dax+, Nature 2025

Pre-merger inference

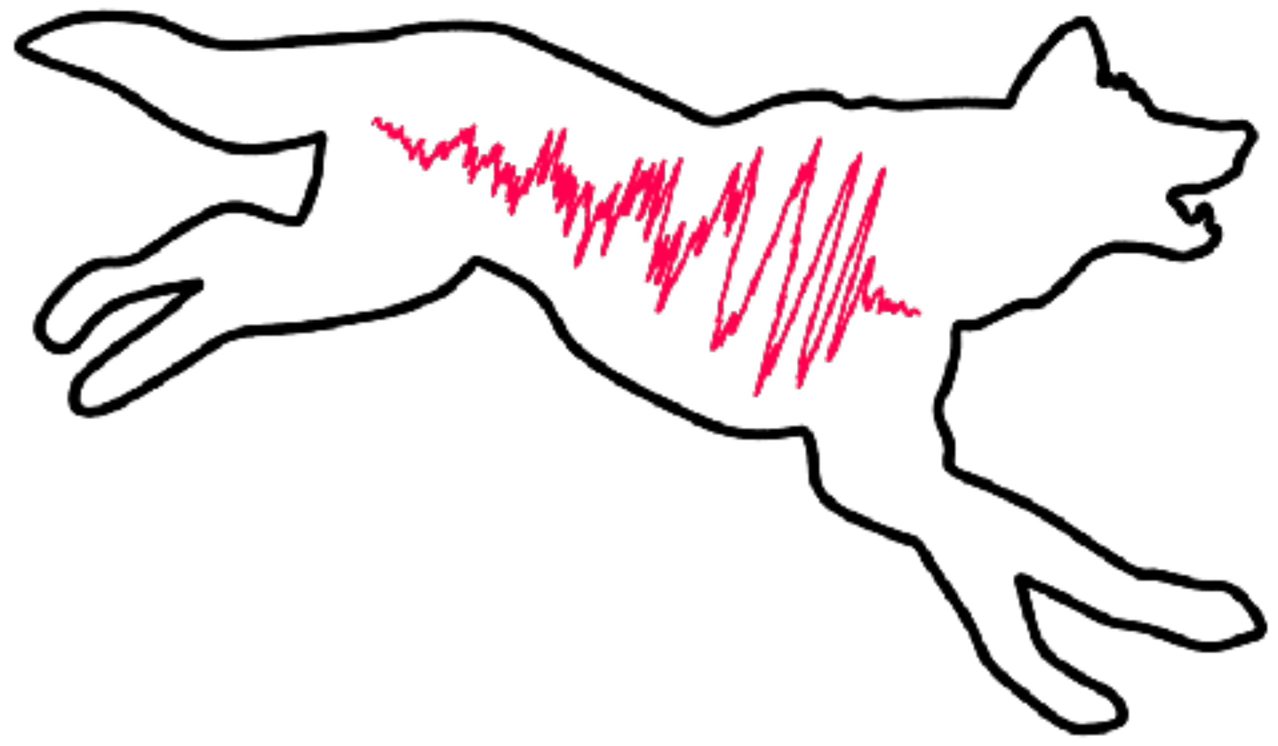
- DINGO-BNS reproduces public LVK results with only 1s inference times
- Inference at arbitrary times before the merger → improves low-latency localization
- Scales to hour-long signals of next-gen detectors







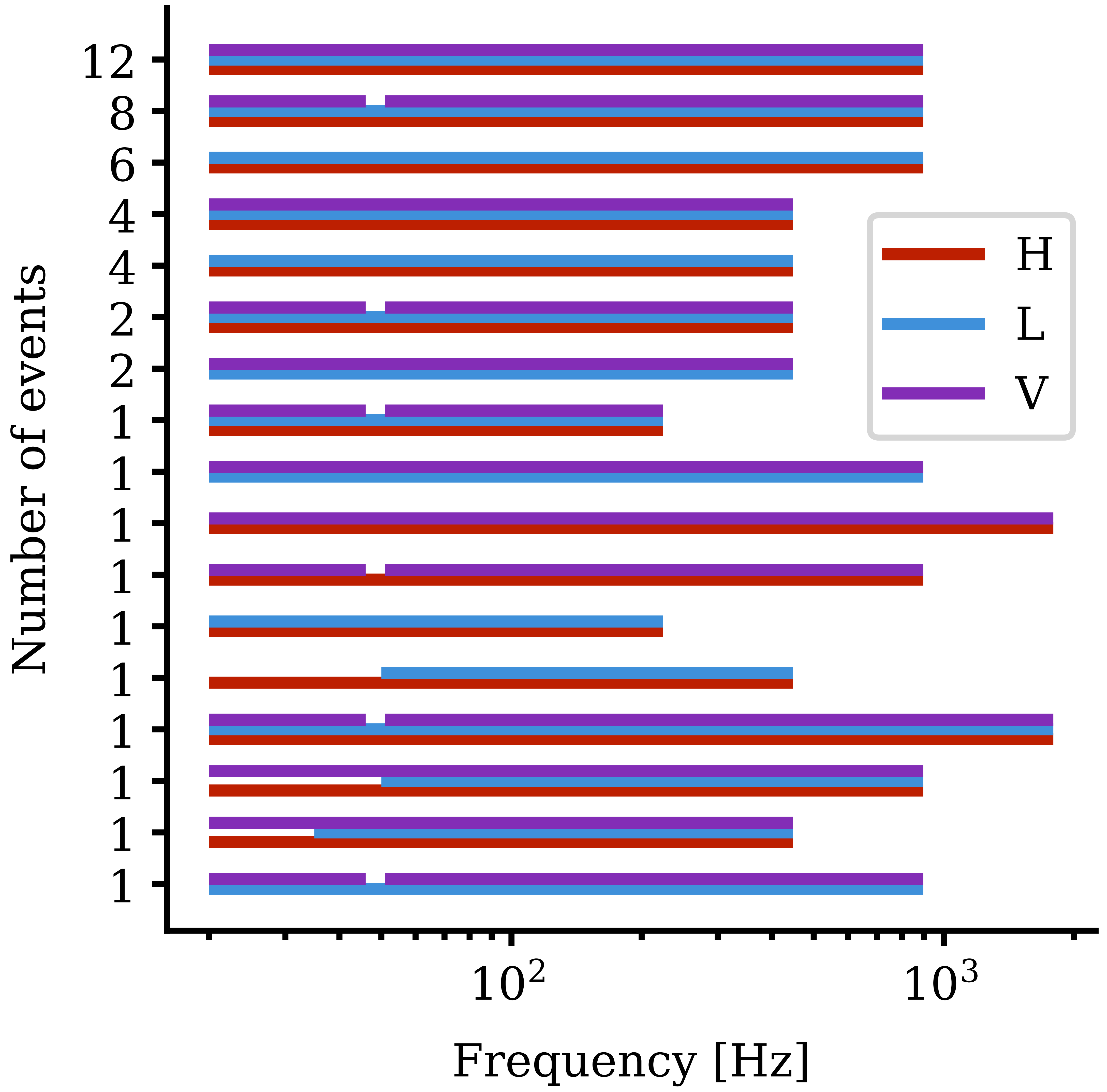
Dax+, Nature 2025



Building big: DINGO-Transformer

Real data is messy

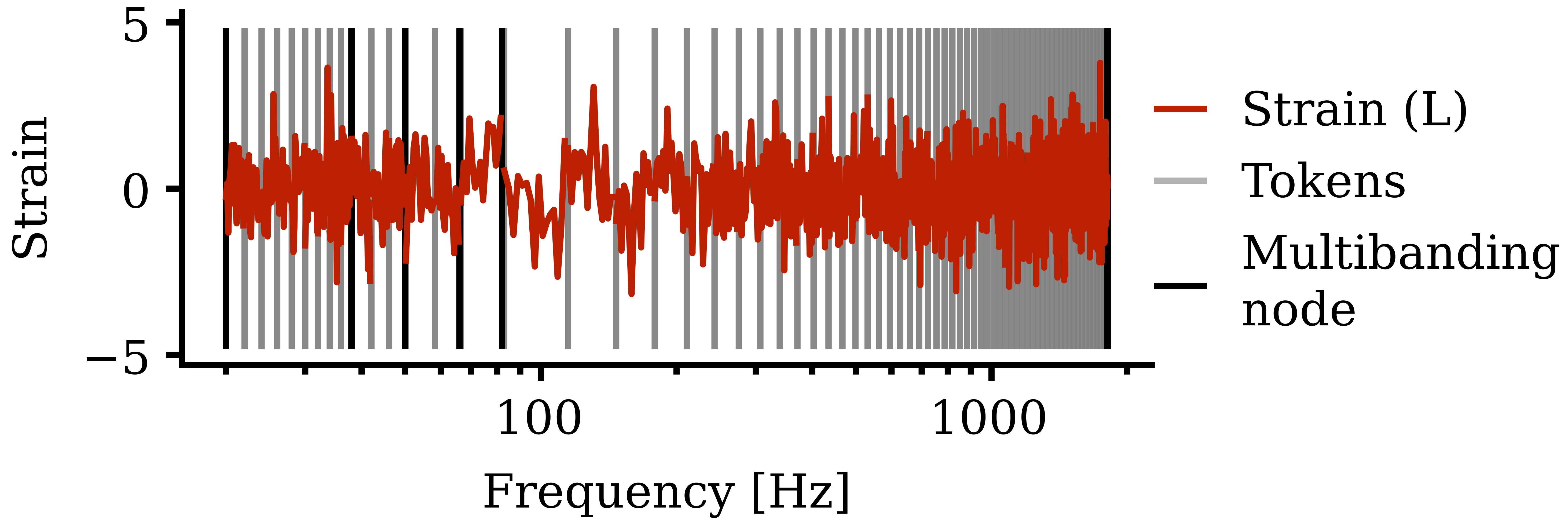
- Problem: DINGO models have fixed input shape
  - Fixed number of detectors
  - Fixed frequency range
  - Official LVK analysis: 48 GW events with 17 different data configurations
- Goal:
  - Amortize different data settings
  - Analyze events with a **single flexible model**

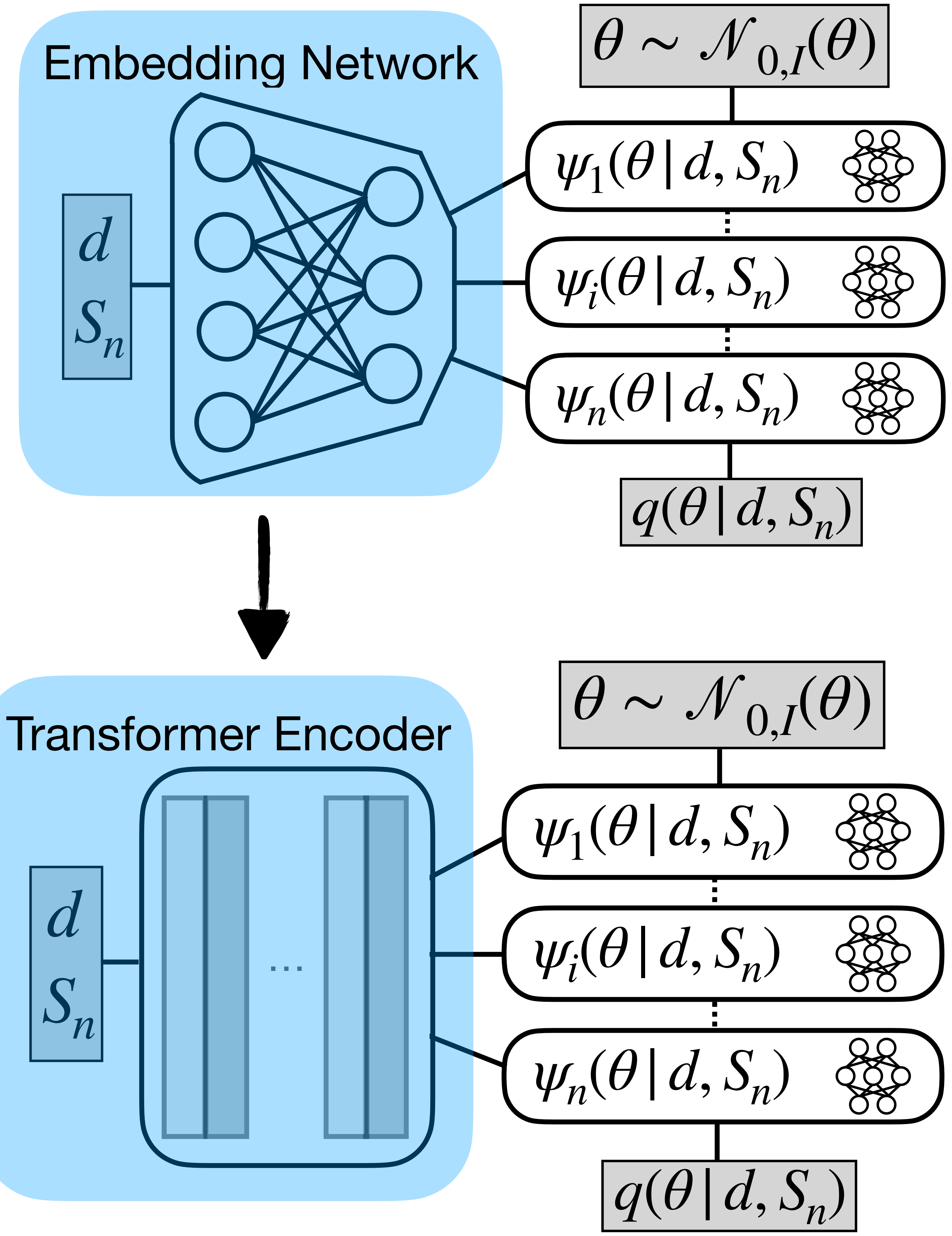


Kofler+, in preparation, 2025

DINGO-Transformer

- Replace fixed embedding network with transformer encoder
- Compress data with multibanding
- Partition into equally sized tokens



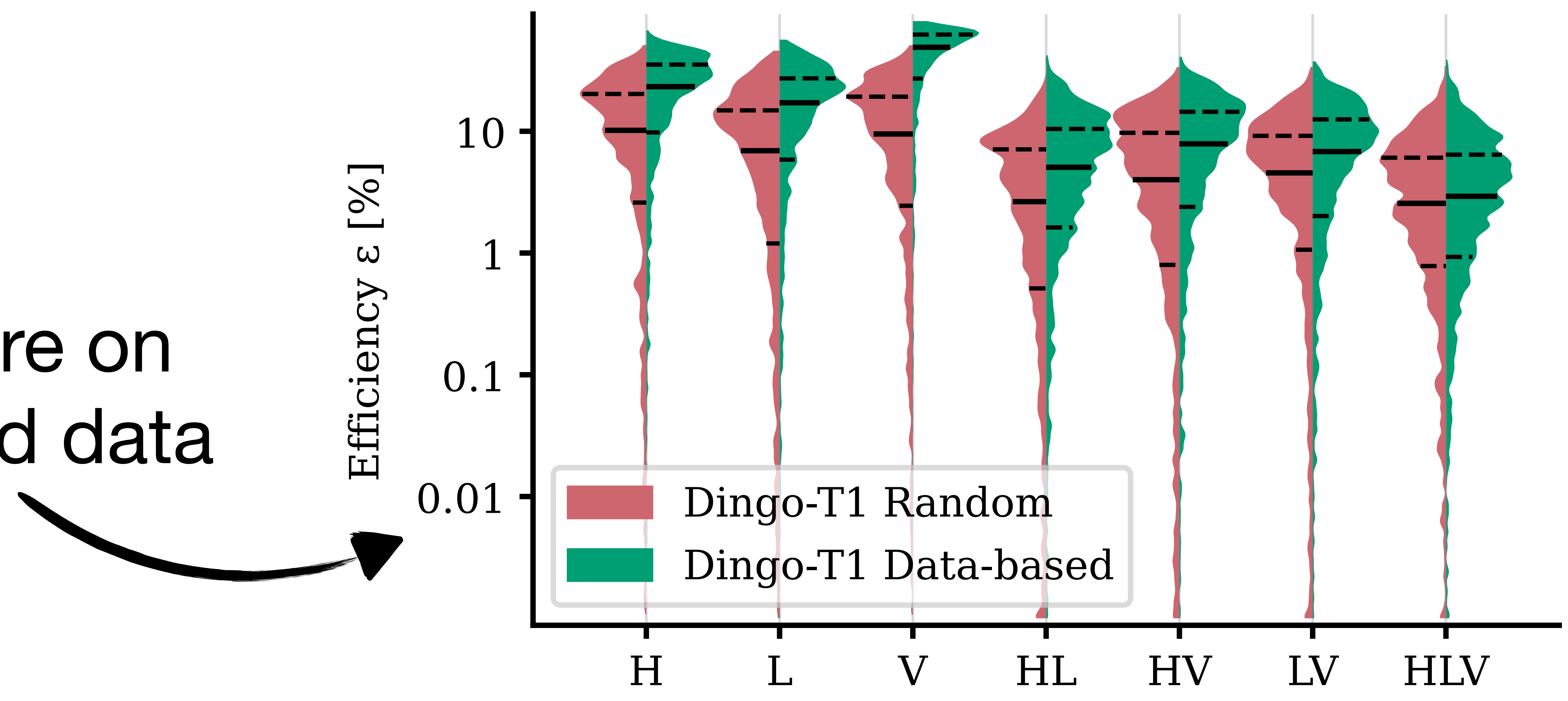


Kofler+, in preparation, 2025

Masking strategies

Random masking
- Mask 40 %
- Sample number of tokens to mask
- Generate random token mask

Compare on simulated data



Data-based masking

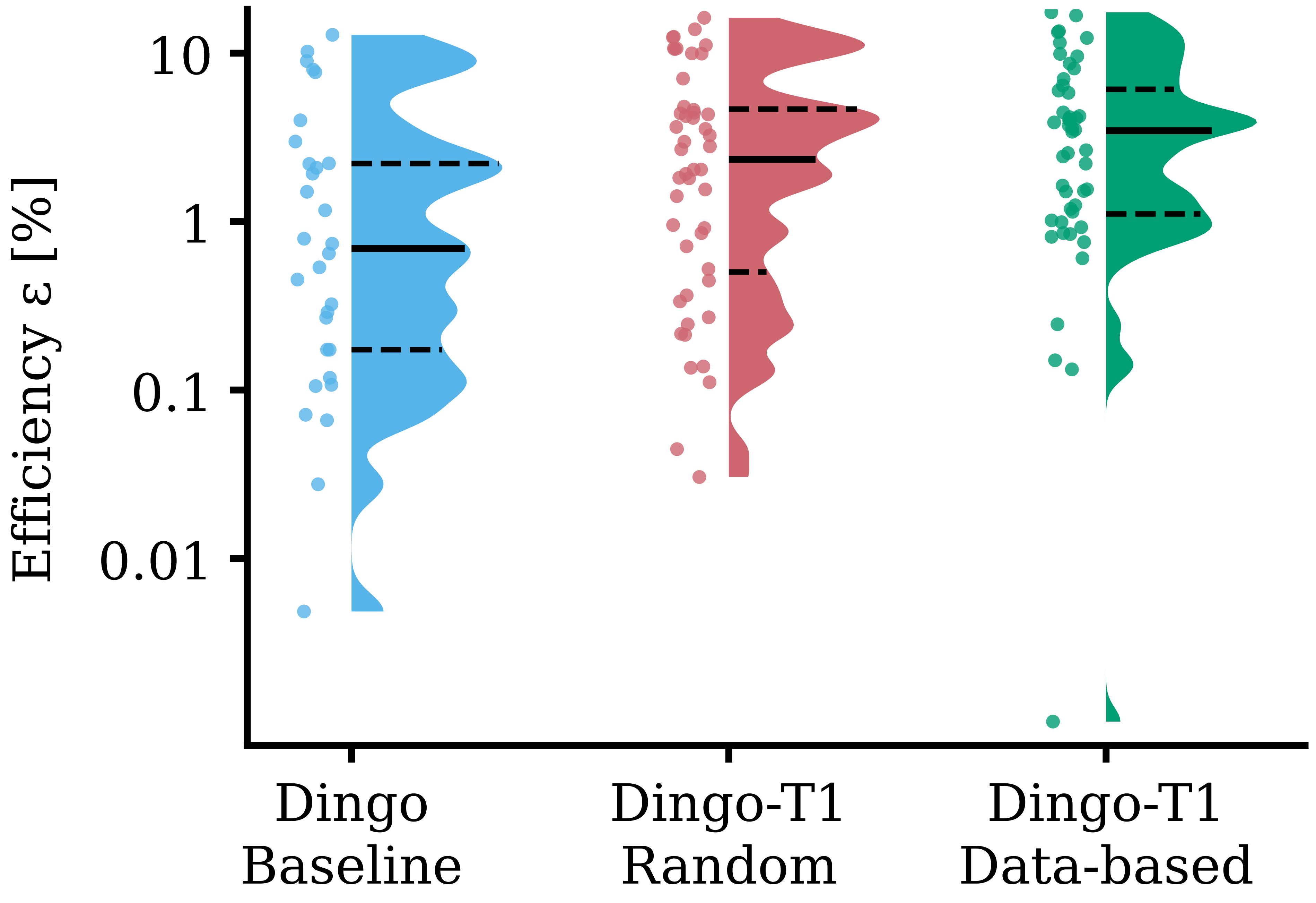
Mask detectors
- How many detectors?
- Which detectors?

Update $f_{\min}/f_{\max}$
- Single detector or all?
- Sample $f_{\min,\mathrm{new}}$ and $f_{\max,\mathrm{new}}$

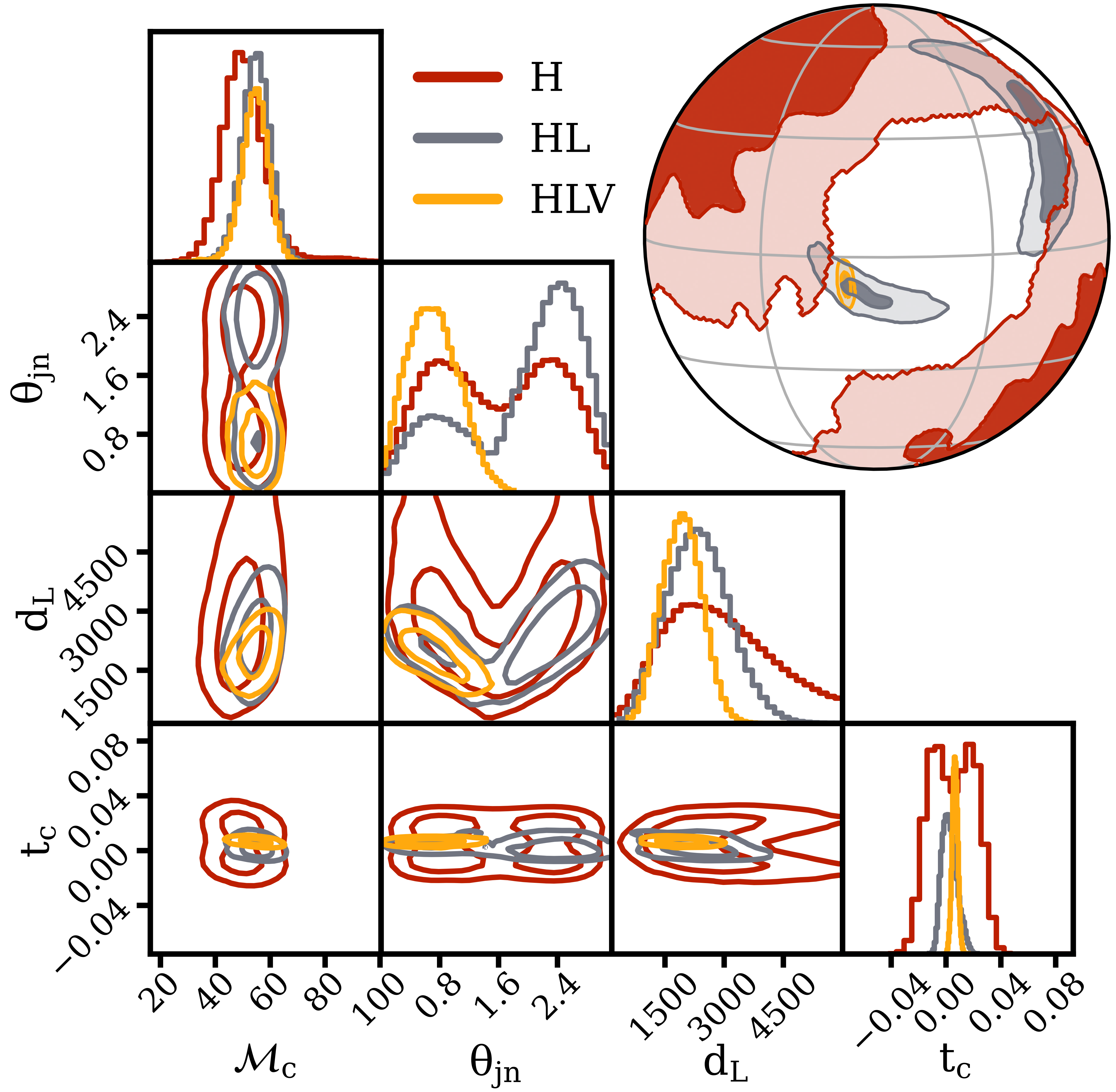
Remove small interval $[f_{\mathrm{low}}, f_{\mathrm{high}}]$
- Sample f_{lower} and f_{high}

Real events

- Evaluated 48 events with 17 different data analysis configurations



GW190701 203306



Kofler+, in preparation, 2025

Towards a DINGO foundation model

- Amortize:
  - **Detector configurations**
  - **Frequency settings**
  - Resolutions
  - Noise of different observing runs
  - BNS & BBH
  - …

Key for "foundation model" in parameter estimation: **Flexibility**

Independent of "big vs. smart"

Where/How is DINGO used?

- Analyze large collection of observations
  - Learn population posterior directly — Leyde+, PRD, 2023
  - Evidence for eccentricity — Gupte+, ArXiv, 2024
- DINGO code reviewed within LIGO
  → Will be used in production analysis!

Package: `pip install dingo-gw`

Resources:
- Website and documentation
- Google Colab Tutorial

Wildberger+, PRD 2023

The Dingo Pack


Maximilian Dax


Stephen Green


Annalena Kofler


Nihar Gupte


Michael Pürrer


Alex Roussopoulos


Samuel Clyne


Ashwin Girish


Cecilia Fabbri


Jonas Wildberger


Vincent Berenz


Jonathan Gair

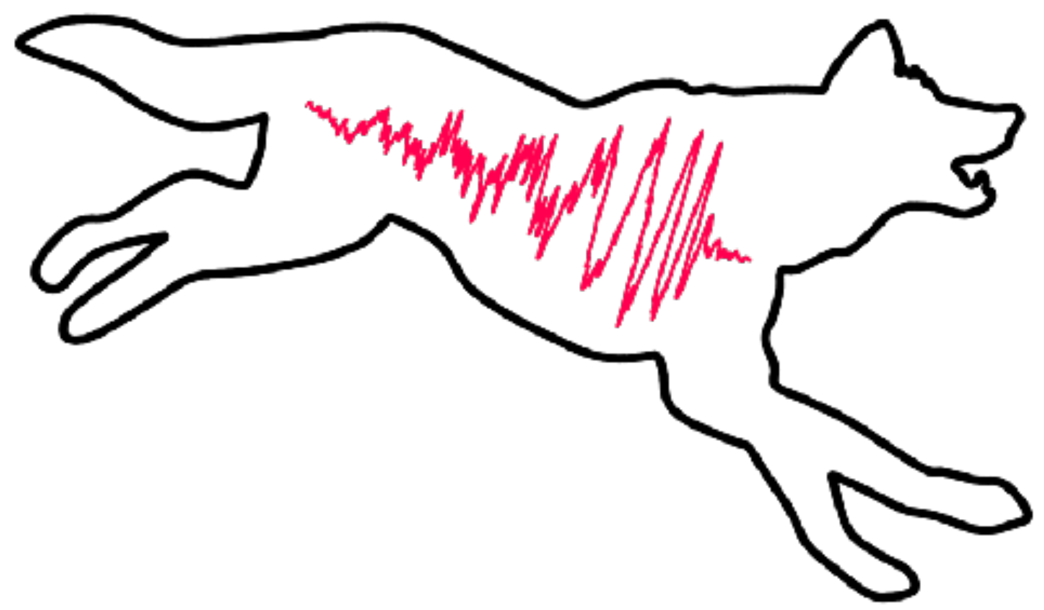

Jakob Macke


Bernhard Schölkopf


Alessandra Buonanno

Thank you!

Do you have any questions?



References

- Dax+, Real-Time Gravitational Wave Science with Neural Posterior Estimation. PRL 127, 2021
- Dax+, Group Equivariant Neural Posterior Estimation, ICLR 2022
- Dax+, Neural Importance Sampling for Rapid and Reliable Gravitational Wave Inference, PRL 130, 2023
- Wildberger+, Adapting to noise distribution shifts in flow-based gravitational-wave inference, PRD 107, 2023
- Wildberger+, Flow Matching for Scalable Simulation-Based Inference, NeurIPS 2023
- Gupte+, Evidence for eccentricity in the population of binary black holes observed by LIGO-Virgo-KAGRA, arXiv:2404.14286v1, 2024
- Dax+, Real-time Gravitational-Wave Inference for Binary Neutron Stars using Machine Learning, Nature, 2025
- Kofler+, in preparation, 2025