GWSky Seminar Series — February 23, 2026

Real-Time Gravitational-Wave Inference with Probabilistic Machine Learning

Maximilian Dax
ELLIS Institute Tübingen &
MPI for Intelligent Systems

Gravitational-wave (GW) astronomy

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**

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

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

M. R. DROUT, A. L. PIRO, ...

SCIENCE · 16 Oct 2017 ·

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

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

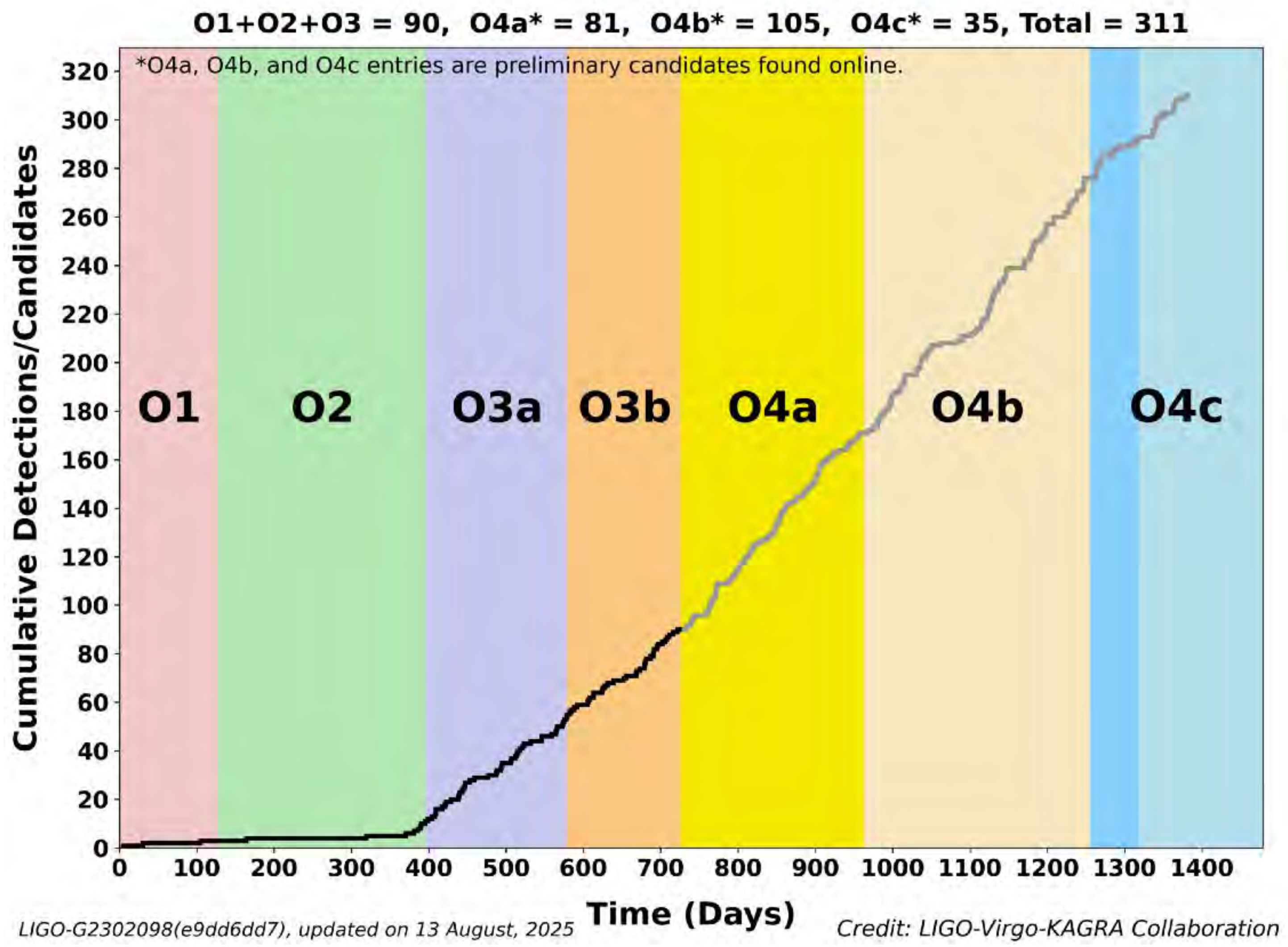
V. Savchenko, C. Ferrigno, E. Kuulkers, A. Bazzano, E. Bozzo, S. Brandt, J. Chenevez,

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

⇒ Such discoveries rely on **GW inference**

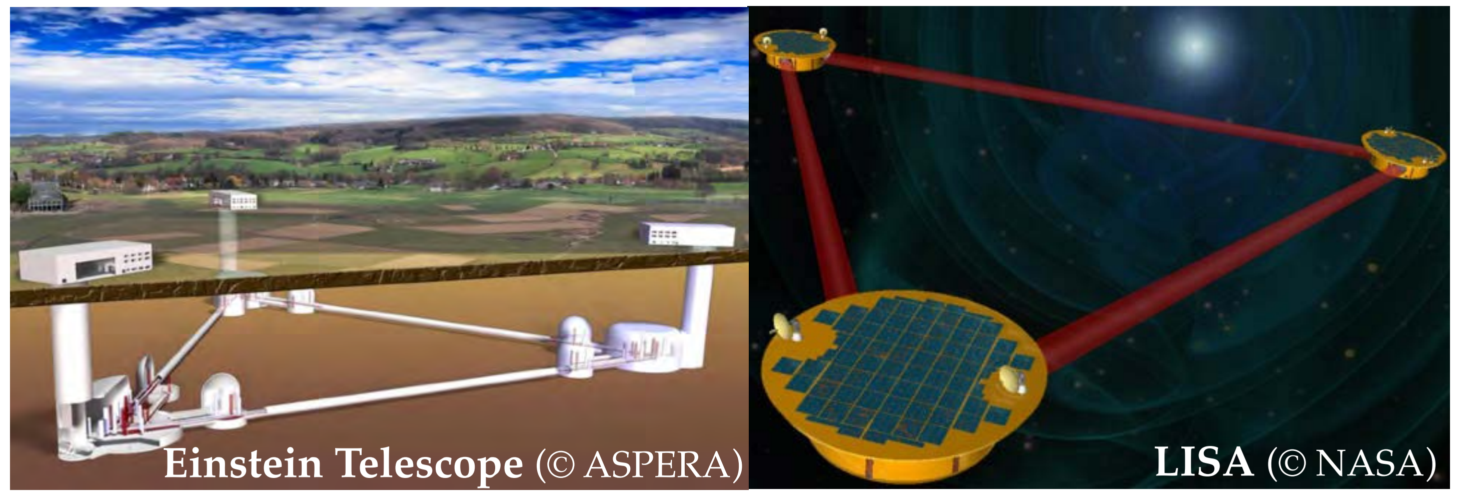
GW inference

- LIGO-Virgo-KAGRA (LVK) have **detected ≈300 GWs**, from mergers of **black holes** or **neutron stars**
  - Binary black holes (**BBH**)
  - Neutron star-black hole (**NSBH**)
  - Binary neutron stars (**BNS**)

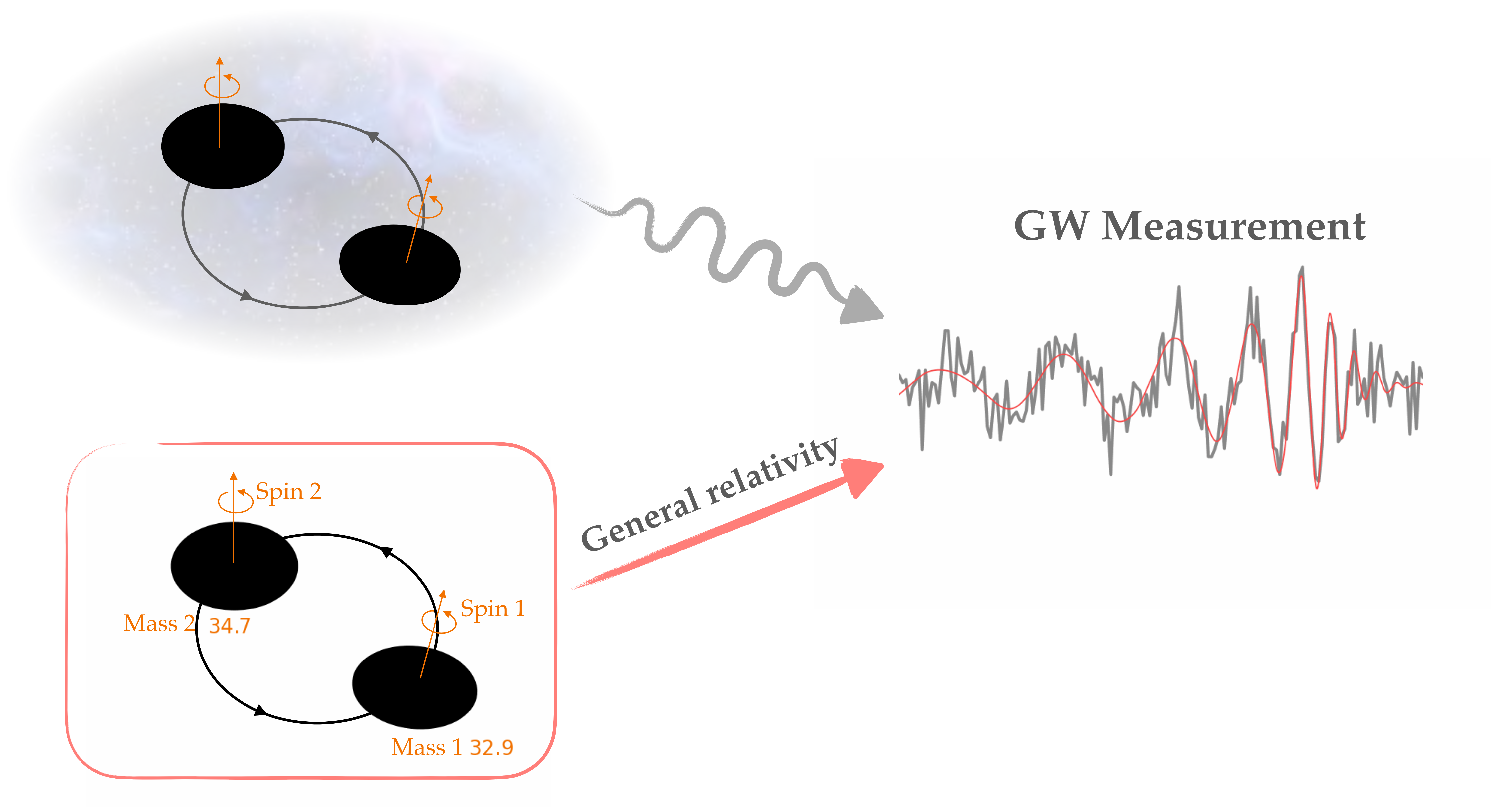


- For each event, **GW inference determines source properties** (component masses, spins, location, orientation, …)
  → takes **hours to days** for a single analysis

- Need for **accelerated analysis**
  - **Efficiency**: 1000x increase in GW detection rates in 2030s
  - **Speed**: detect counterpart radiation for BNS and NSBH



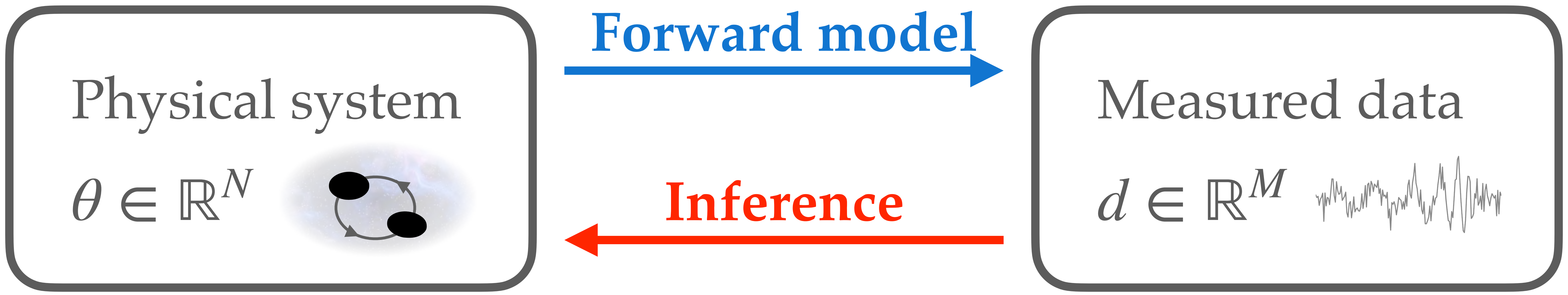
GW inference: comparing data to models



- GWs are spacetime deformations emitted by accelerated masses (e.g., black hole mergers)

- GW shape depends on source (~15 parameters: masses, spins, …)
- GW inference:
  Decode GW information to characterise the astro source

Inverse problems in science

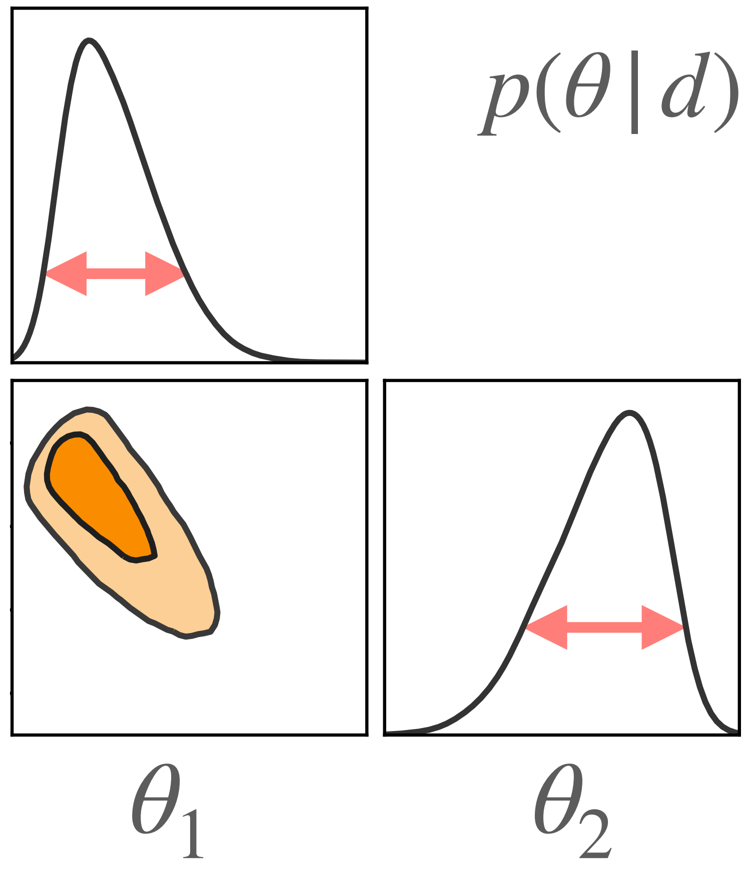


- **Forward direction** $\theta \to d$ is defined by a simulator, $d \sim p(d|\theta)$

- **Inverse direction** with Bayesian inference

forward model — prior belief

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

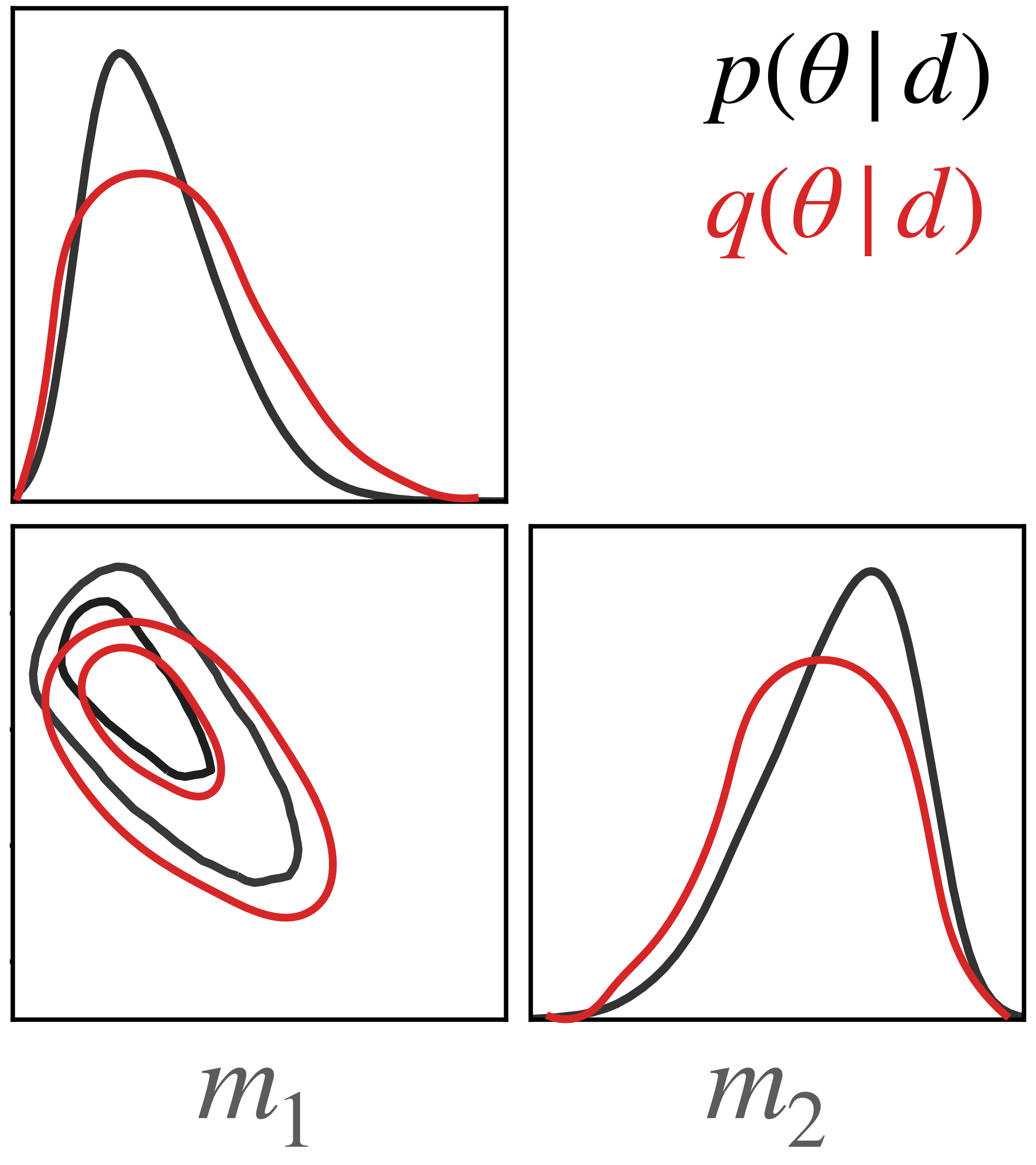


Bayesian posterior

→ uncertainties

→ correlations

Bayesian inference



Goal: **sample the posterior**

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

Idea: train neural emulator

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

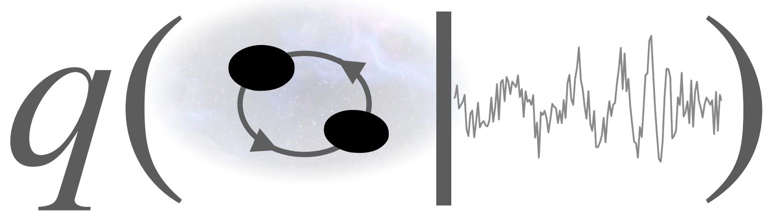
and sample with emulator instead

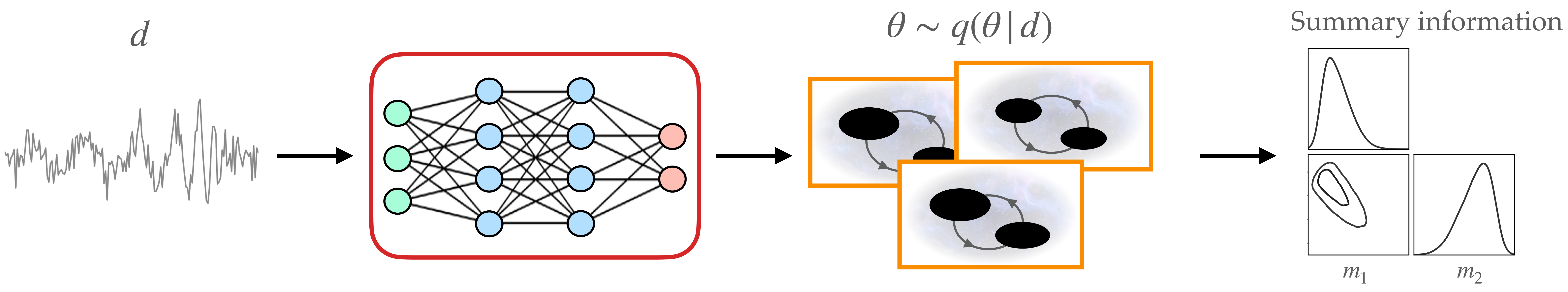
$$\theta \sim q(\theta|d)$$

Papamakarios, Murray. *Fast ε-free inference of simulation models with bayesian conditional density estimation.* **NeurIPS**, 2016.
Cranmer, Brehmer, Louppe. *The frontier of simulation-based inference.* **PNAS**, 2020.

Simulation-based inference (SBI) & Neural posterior estimation (NPE)

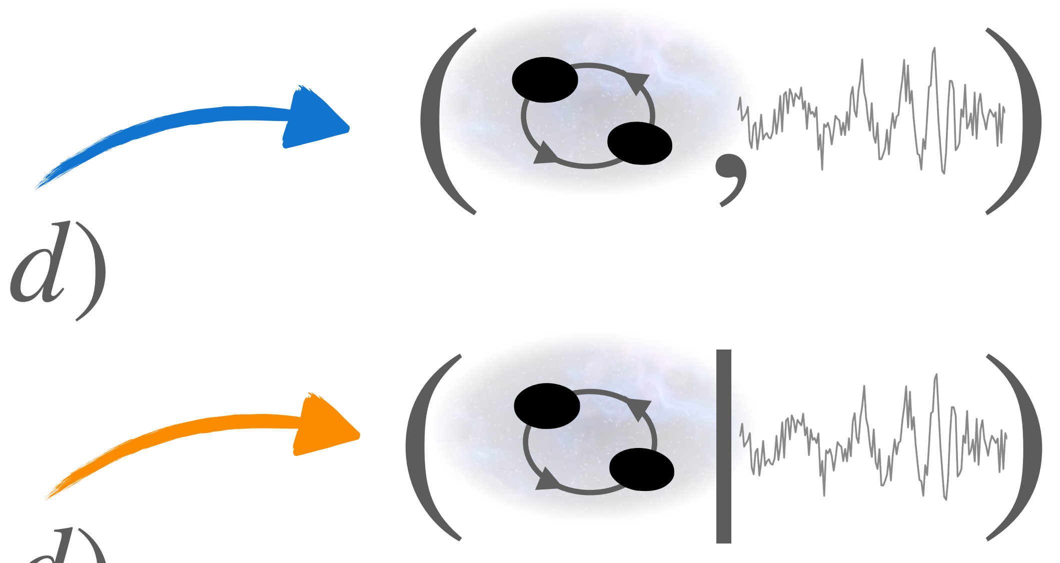
Idea: approximate posterior with an **inference model** $q(\theta \mid d) \approx p(\theta \mid d)$





Training

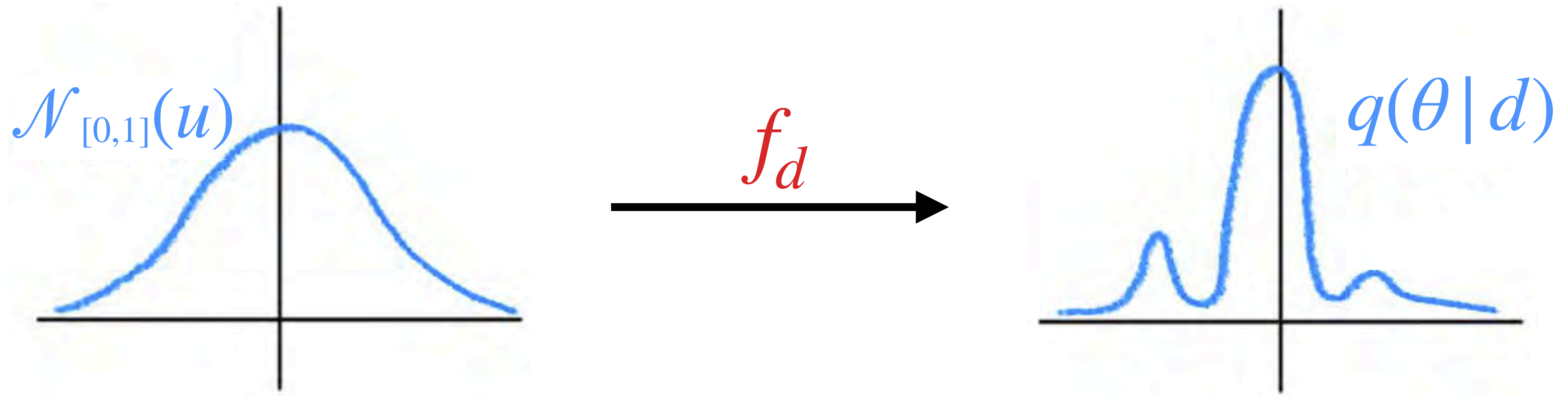
1. Generate samples from the **joint distribution** $(\theta, d) \sim p(\theta, d)$
   sample i generated with forward model $\theta_i \sim p(\theta)$, $d_i \sim p(d \mid \theta_i)$
2. Train neural network to **predict parameters** given data $(\theta \mid d)$
   any density estimator (normalising flows, flow matching, diffusion models, …)



$$L = -\log q(\theta_i \mid d_i),$$
$$\theta_i \sim p(\theta),\ d_i \sim p(d \mid \theta_i)$$

Normalising flows

- Idea: transform base distribution $\mathcal{N}_{[0,1]}$ to $q(\theta \mid d)$ via f_d



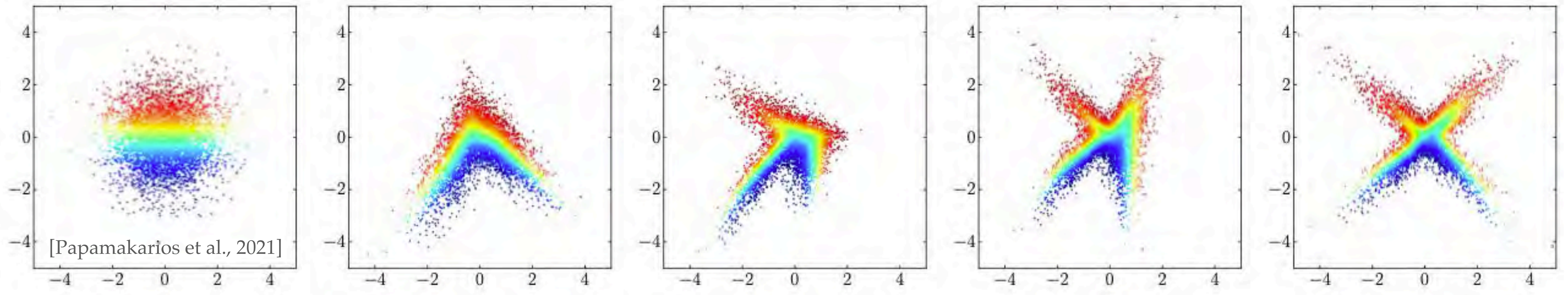
Sampling

$$\theta \sim q(\theta \mid d): \quad \theta = f_d(u),\ u \sim \mathcal{N}(u)$$

Density evaluation

$$q(\theta \mid d) = \mathcal{N}_{[0,1]}\left(f_d^{-1}(\theta)\right) \left|\det J_{f_d}^{-1}\right|$$

- Flexible f_d achieved by **composition of simple transforms**



f_d: NN with learnable parameters ϕ;
with tractable inverse and Jacobian;
parametric dependence on d

- Normalizing flows can be made **arbitrarily expressive**

Neural Posterior Estimation (NPE)

- Minimize

$$D_{\mathrm{KL}}\left(p \,|\, q\right) = -\,\mathbb{E}_{\theta\sim p(\theta)}\mathbb{E}_{d\sim p(d|\theta)}\left[\log q(\theta \,|\, d)\right] \quad + \text{const.}$$

- Monte Carlo approximation: train flow by minimizing loss L across dataset $\mathscr{D}$

$$L = -\log q_\phi(\theta \,|\, d), \qquad \mathscr{D} = \left\{\theta^{(i)}, d^{(i)}\right\}_{i=1}^{N}, \quad \theta^{(i)} \sim p(\theta),\ d^{(i)} \sim p(d \,|\, \theta^{(i)})$$

- Minimization of D_{KL} + arbitrarily expressive ⇒ **perfect recovery of posterior**

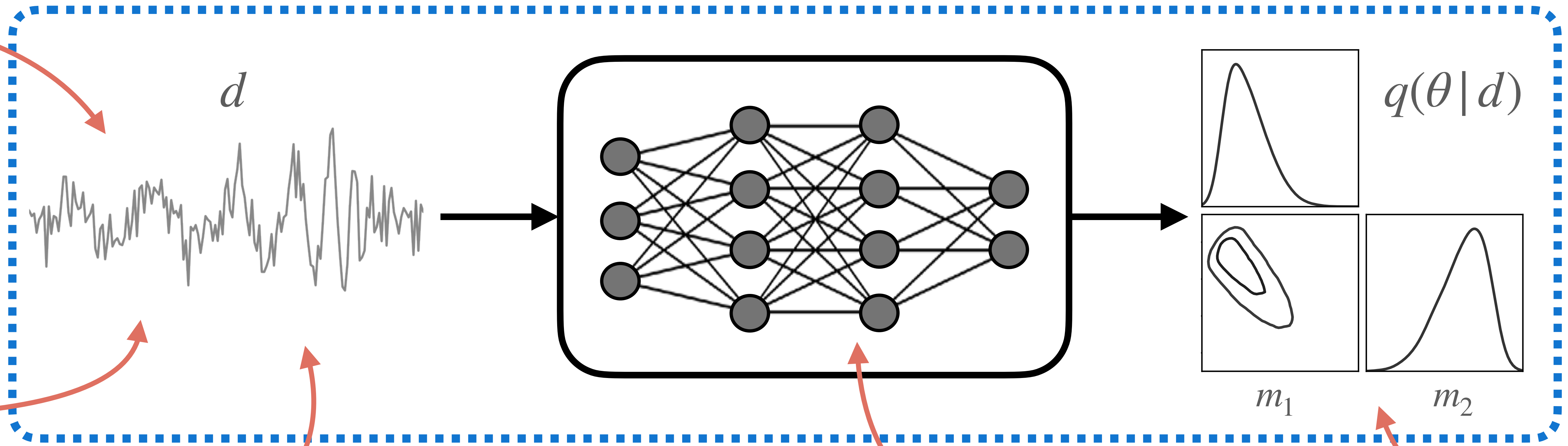
- NPE uses same ingredients as MCMC (prior + likelihood), but **only requires samples**

SBI for Gravitational Waves

SBI for Gravitational Waves

How can we integrate known relations between data and parameters? [1,2]

How can we handle missing data? [8]



How can we handle high dimensional data? [7]

How can we adapt to changing detector noise properties? [1,4]

How can we ensure neural network expressivity? [5]

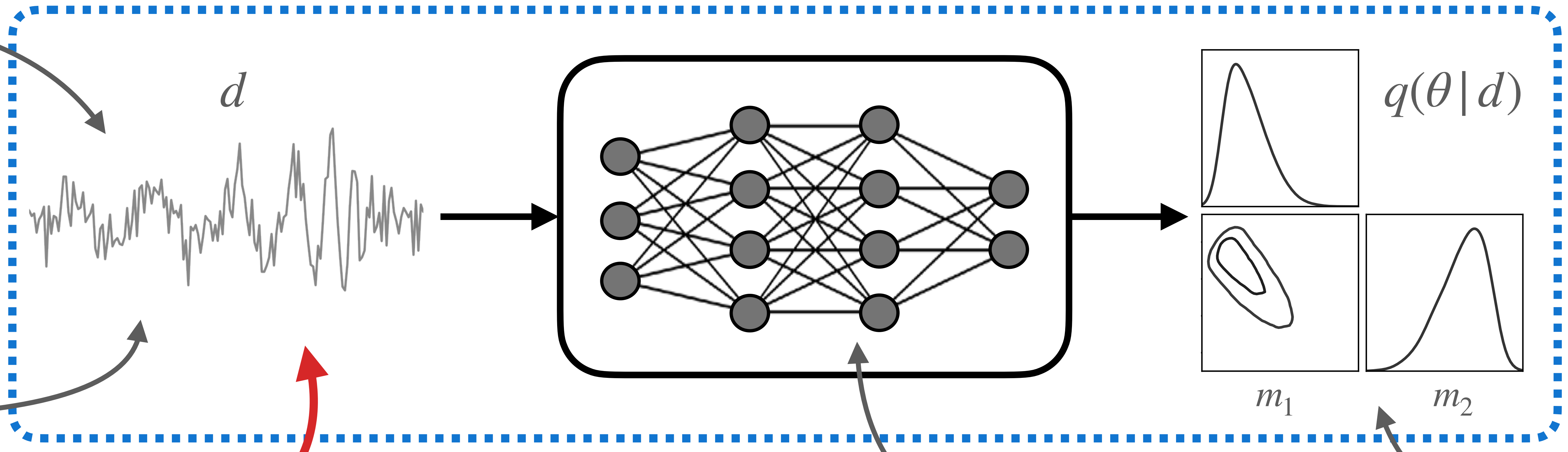
How can we guarantee accurate inference? [3]

[1] Dax, Green, Gair, Macke, Buonanno, Schölkopf. *Real-time gravitational wave science with neural posterior estimation.* **Physical Review Letters**, 2021.
[2] Dax, Green, Gair, Deistler, Schölkopf, Macke. *Group equivariant neural posterior estimation.* **ICLR**, 2022.
[3] Dax, Green, Gair, Pürrer, Wildberger, Macke, Buonanno, Schölkopf. *Neural Importance Sampling for Rapid and Reliable Gravitational-Wave Inference.* **Physical Review Letters**, 2023.
[4] Wildberger, Dax, Green, Gair, Pürrer, Macke, Buonanno, Schölkopf. *Adapting to noise distribution shifts in flow-based gravitational-wave inference*. **Physical Review D**, 2023.
[6] Wildberger, Dax, Buchholz, Green, Macke, Schölkopf. Flow Matching for Scalable Simulation-Based Inference. **NeurIPS**, 2023.
[7] Dax, Green, Gair, Gupte, Pürrer, Raymond, Wildberger, Macke, Buonanno, Schölkopf. *Real-time inference for binary neutron star mergers using machine learning.* **Nature**, 2025.
[8] Kofler, Dax, Green, Wildberger, Gupte, Macke, Gair, Buonanno, Schölkopf. *Flexible Gravitational-Wave Parameter Estimation with Transformers*. **Preprint**, 2025.

Adapting to changing Detector Noise Properties

How can we integrate known relations between data and parameters?

How can we handle missing data?



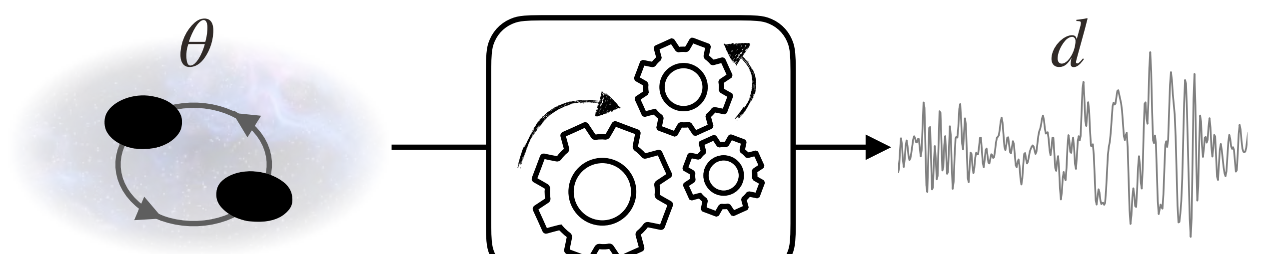
How can we handle high dimensional data?

How can we adapt to changing detector noise properties?

How can we ensure neural network expressivity?

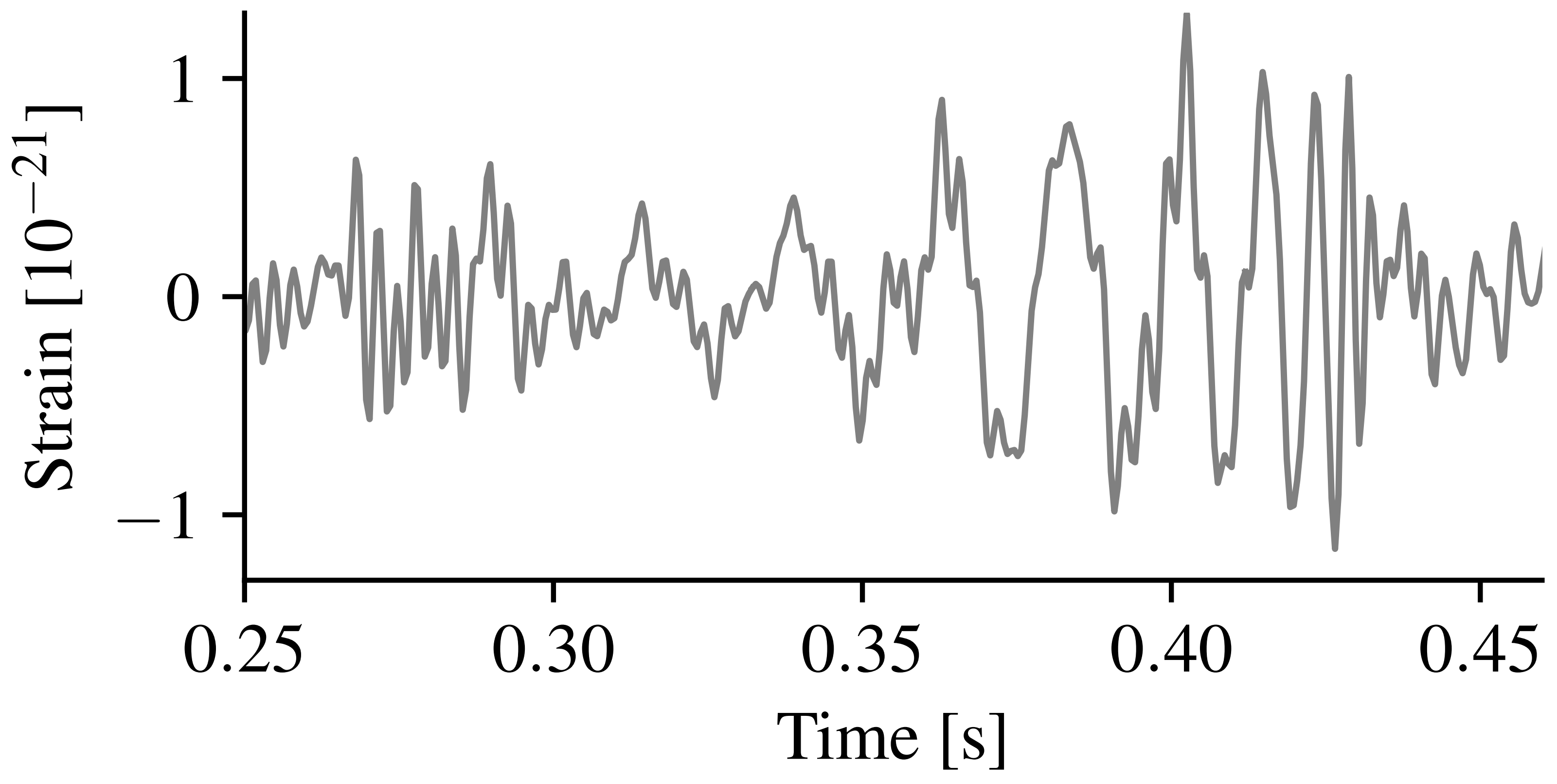
How can we guarantee accurate inference?

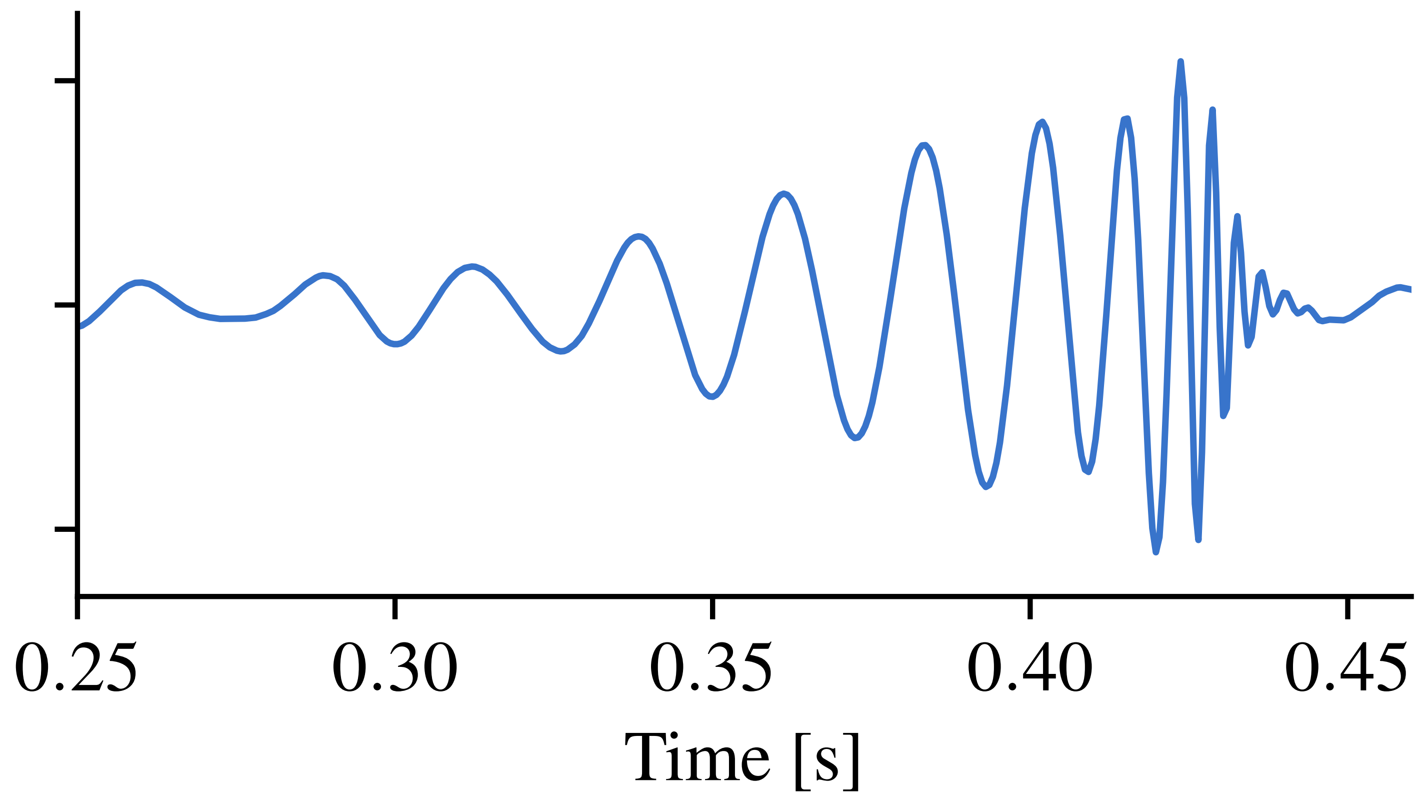
Forward model

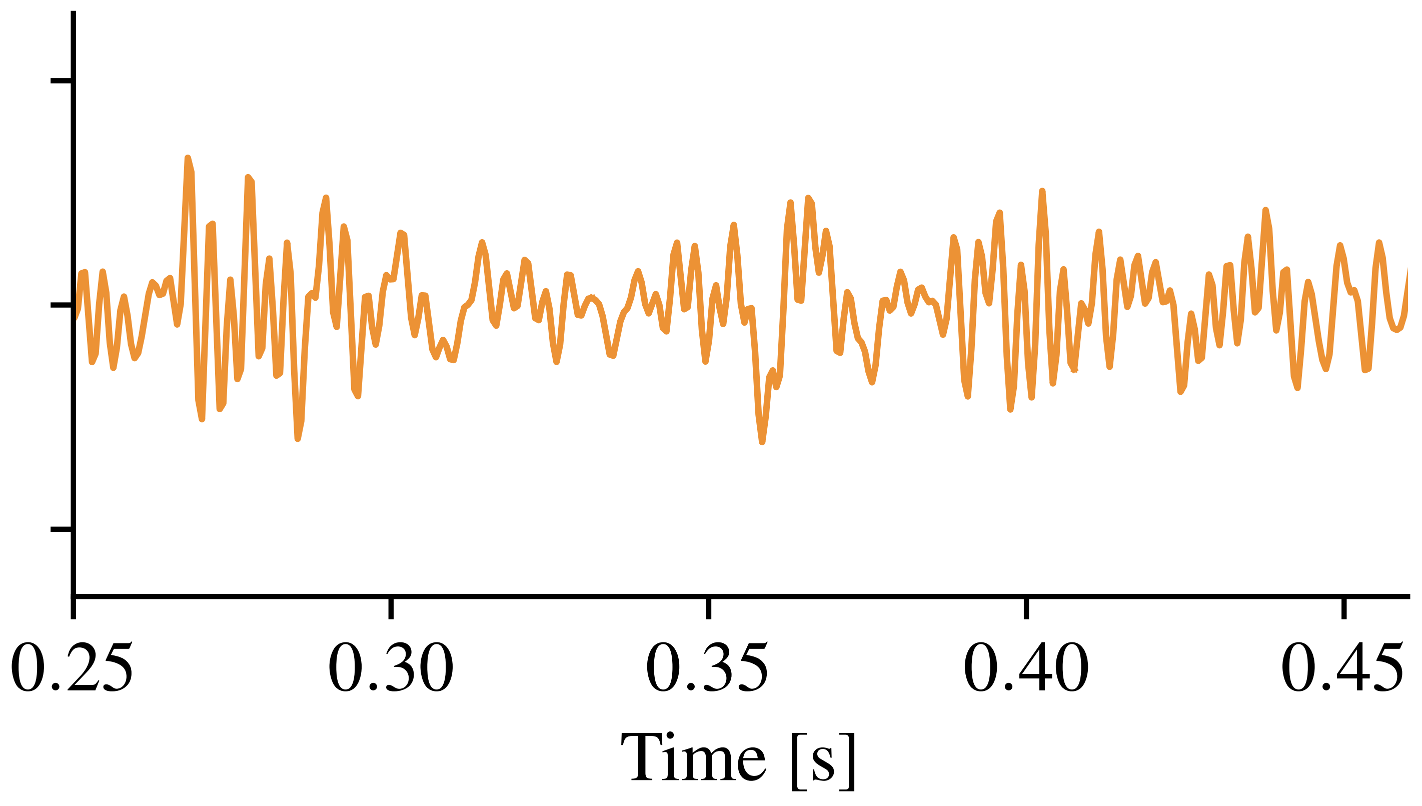


GW measurement: **signal** + **noise**

$d \sim p(d|\theta)$ = $h(\theta)$ + $n \sim \mathcal{N}(0, S_n)$

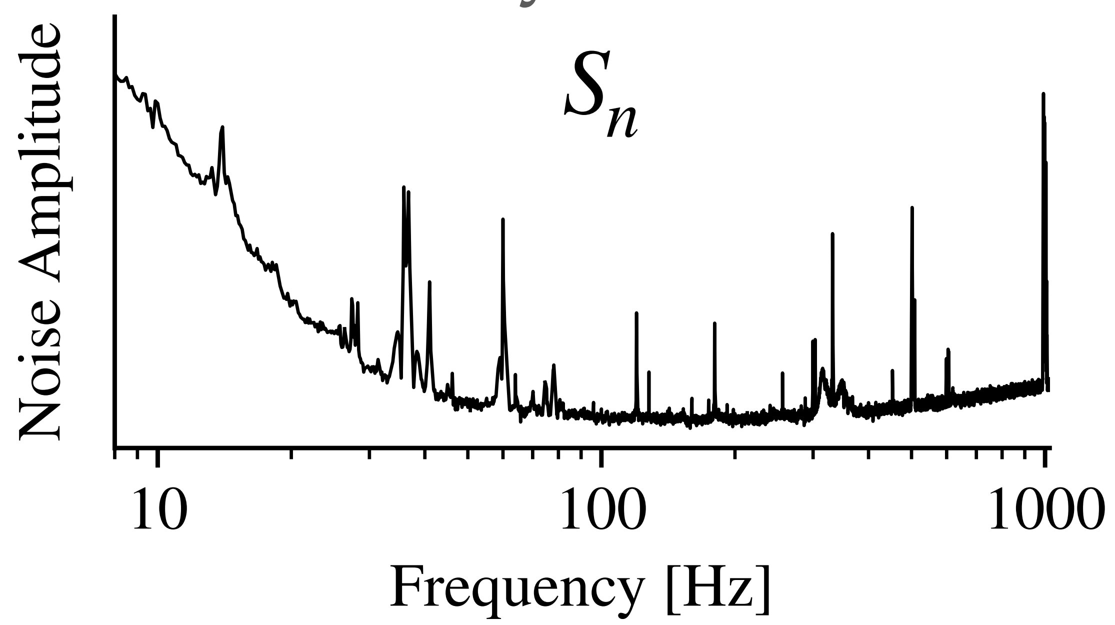






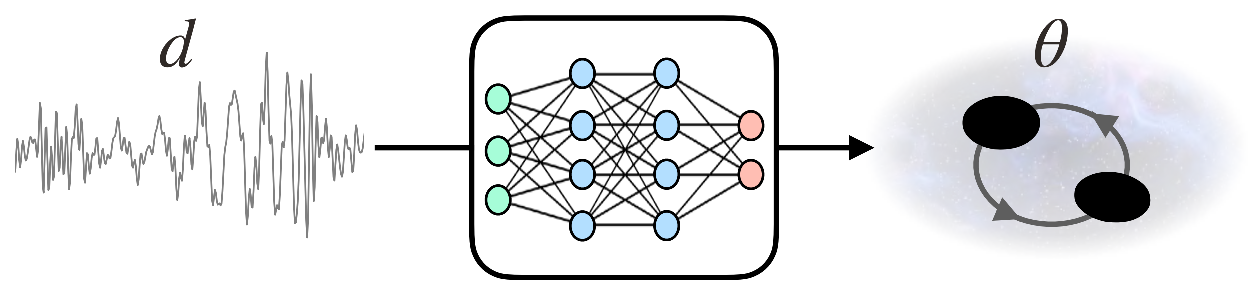
GW signal model
(general relativity)

Detector noise model
(Stationary, Gaussian)

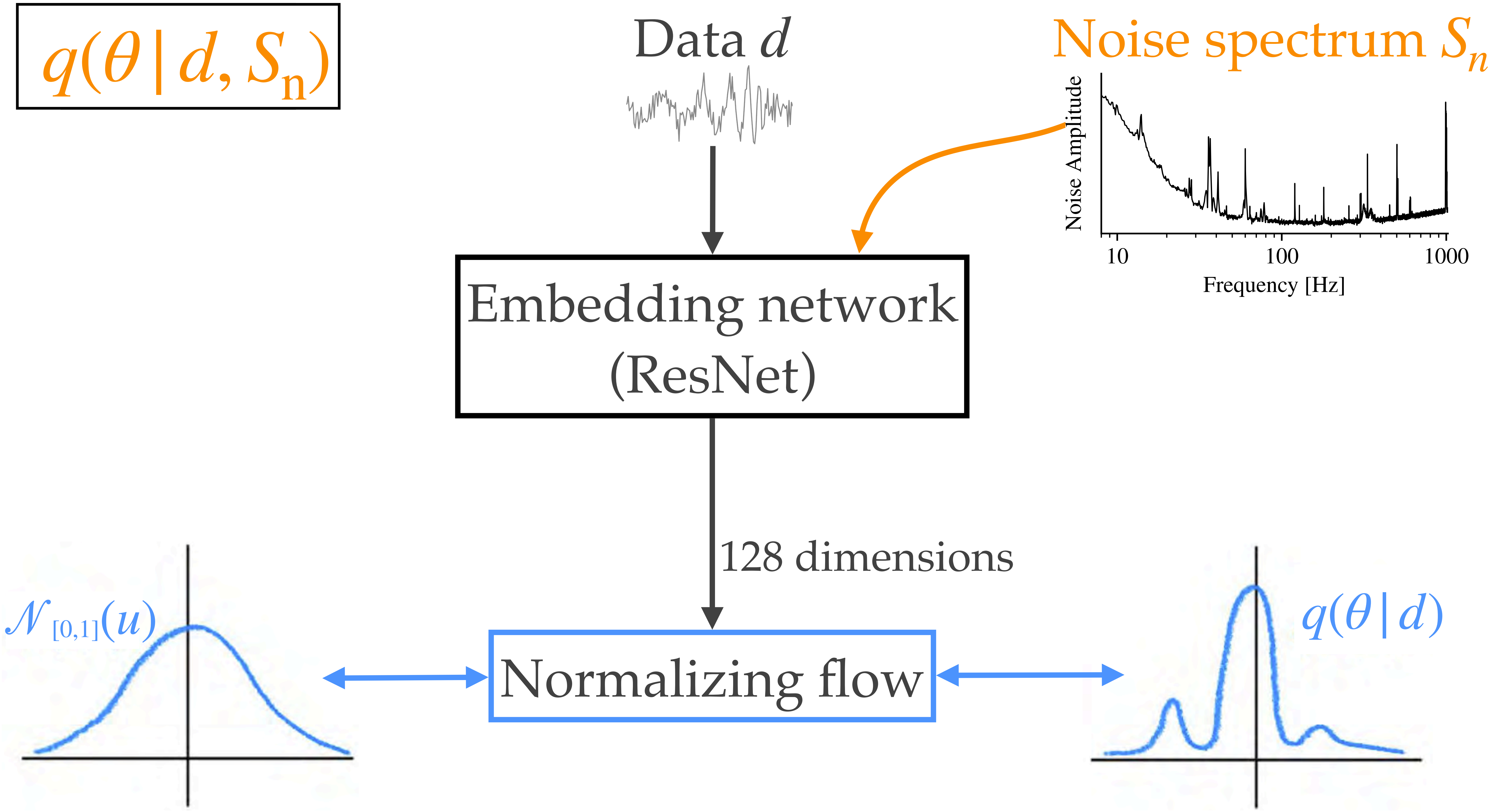


Dax, Green, Gair, Macke, Buonanno, Schölkopf. *Real-time gravitational wave science with neural posterior estimation.* **Physical Review Letters**, 2021.

Inverse model



$q(\theta \,|\, d, S_n)$



Simulation-based training

$\theta \sim p(\theta),$

$S_n \sim p(S_n),$

$d \sim p(d \,|\, \theta, S_n)$

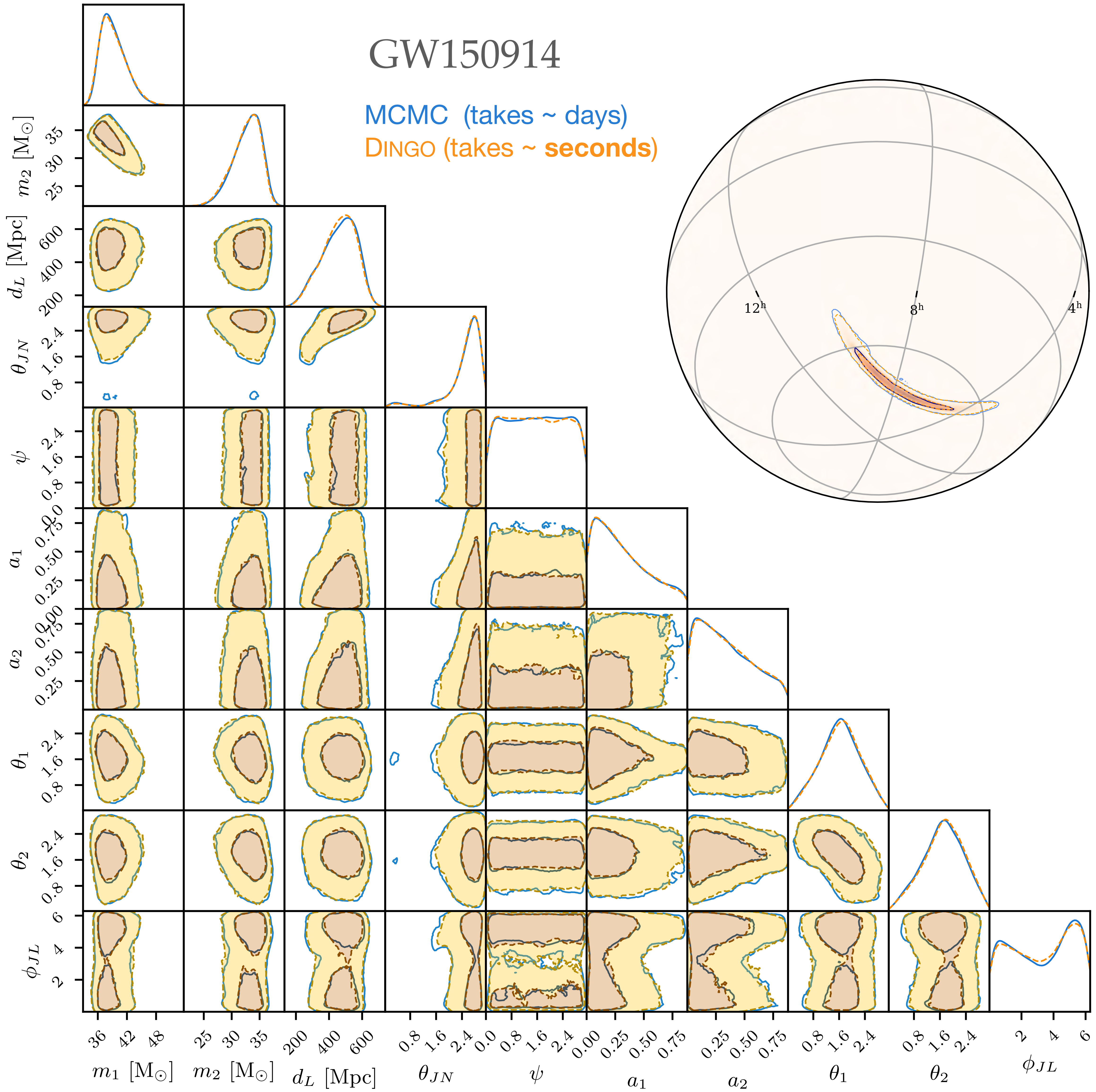
$\Rightarrow$ NPE model **tunable** to varying noise levels

NPE results

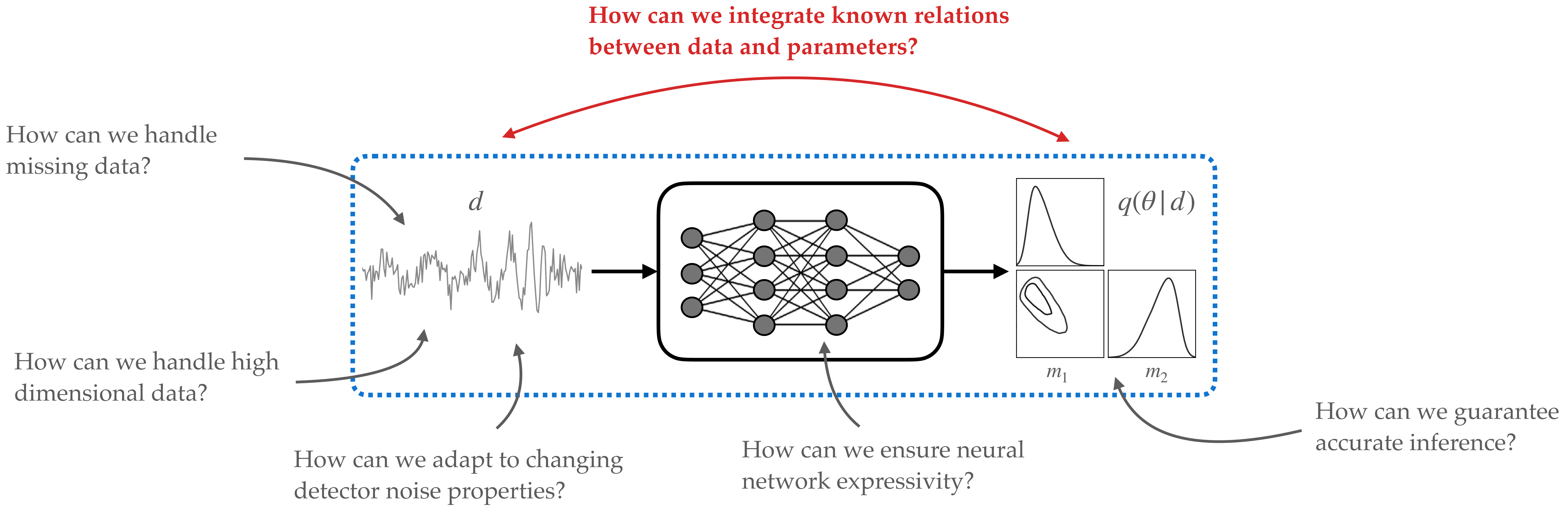
- Noise conditioning: a **single NPE model** is applicable to **entire observing run**
- Agreement with standard samplers on eight real GW events
- Inference in seconds using pre-trained networks (**1000x speed up**)

DINGO

Deep INference for Gravitational-wave Observations



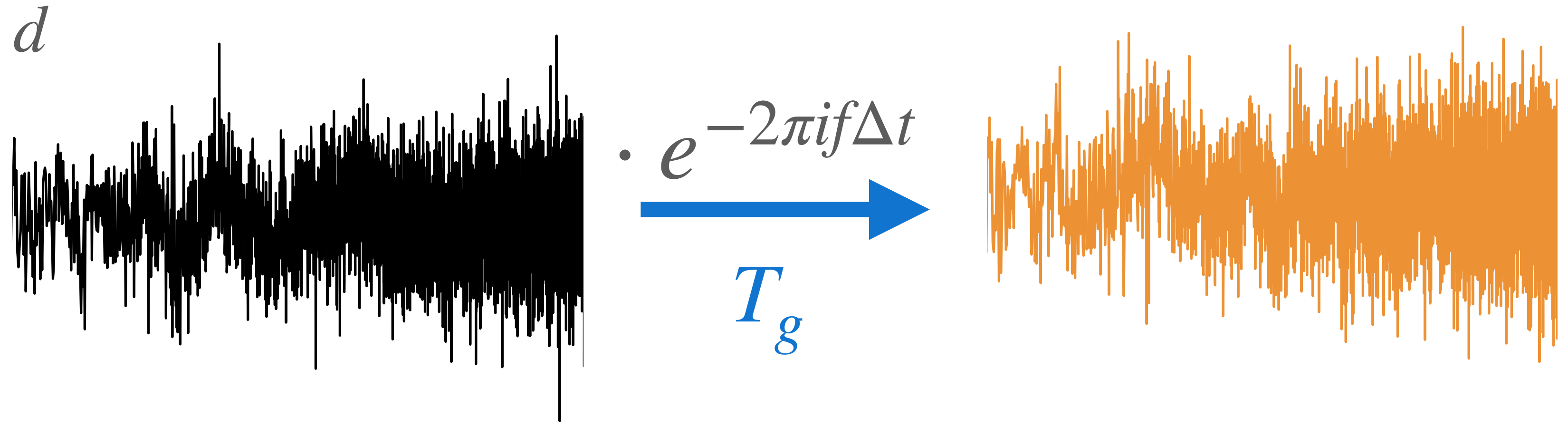
Integrating Physical Symmetries



NPE with symmetries: Group-equivariant NPE

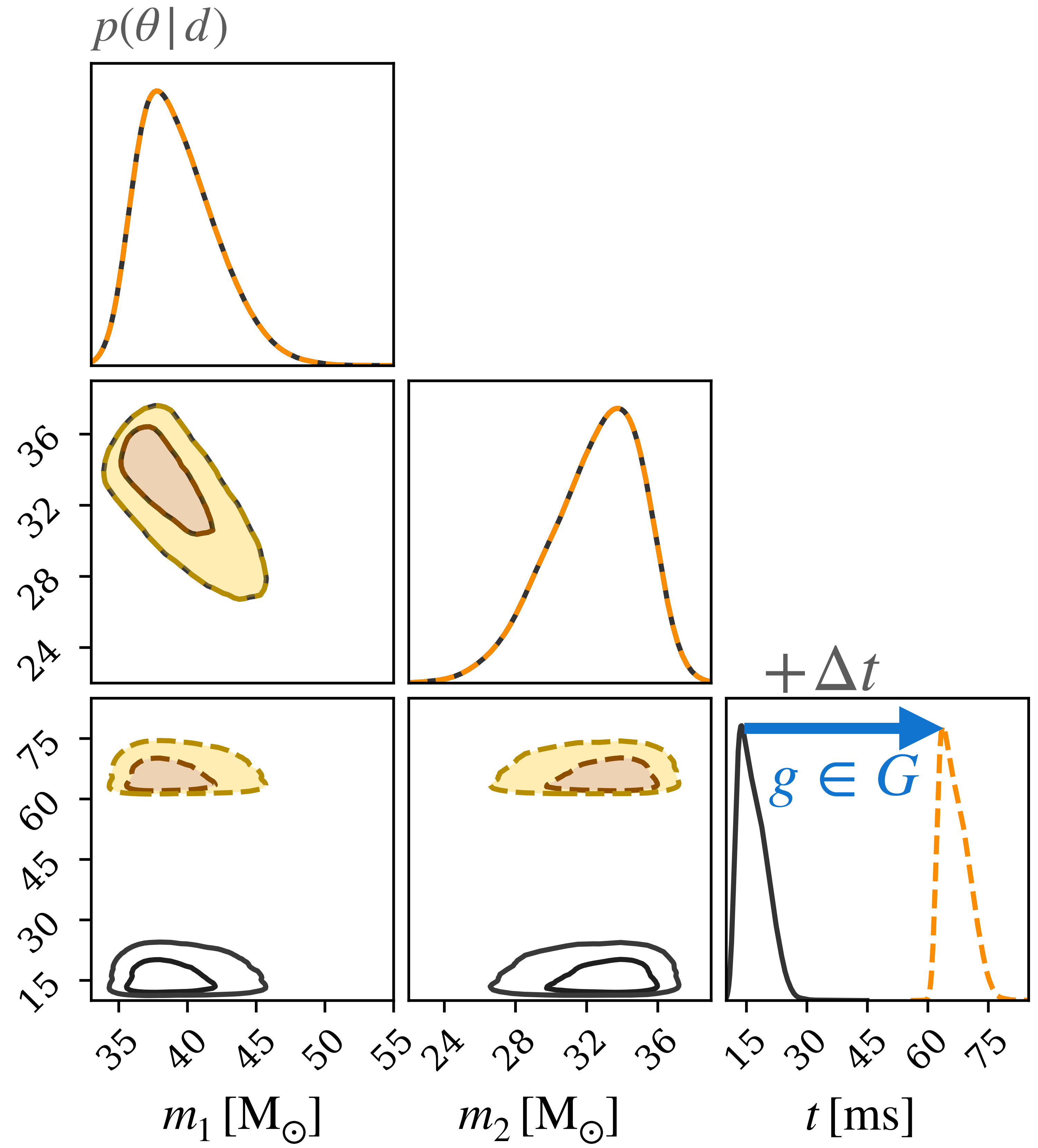
- Equivariance (covariance) under time shift

$$p(\theta|d) = p(g\theta|T_g d)\,|\det J_g| \qquad \forall g \in G$$



- NPE learns such symmetries from simulation data
  ⇒ requires network and training capacity
  ⇒ can we instead **enforce such symmetries**?

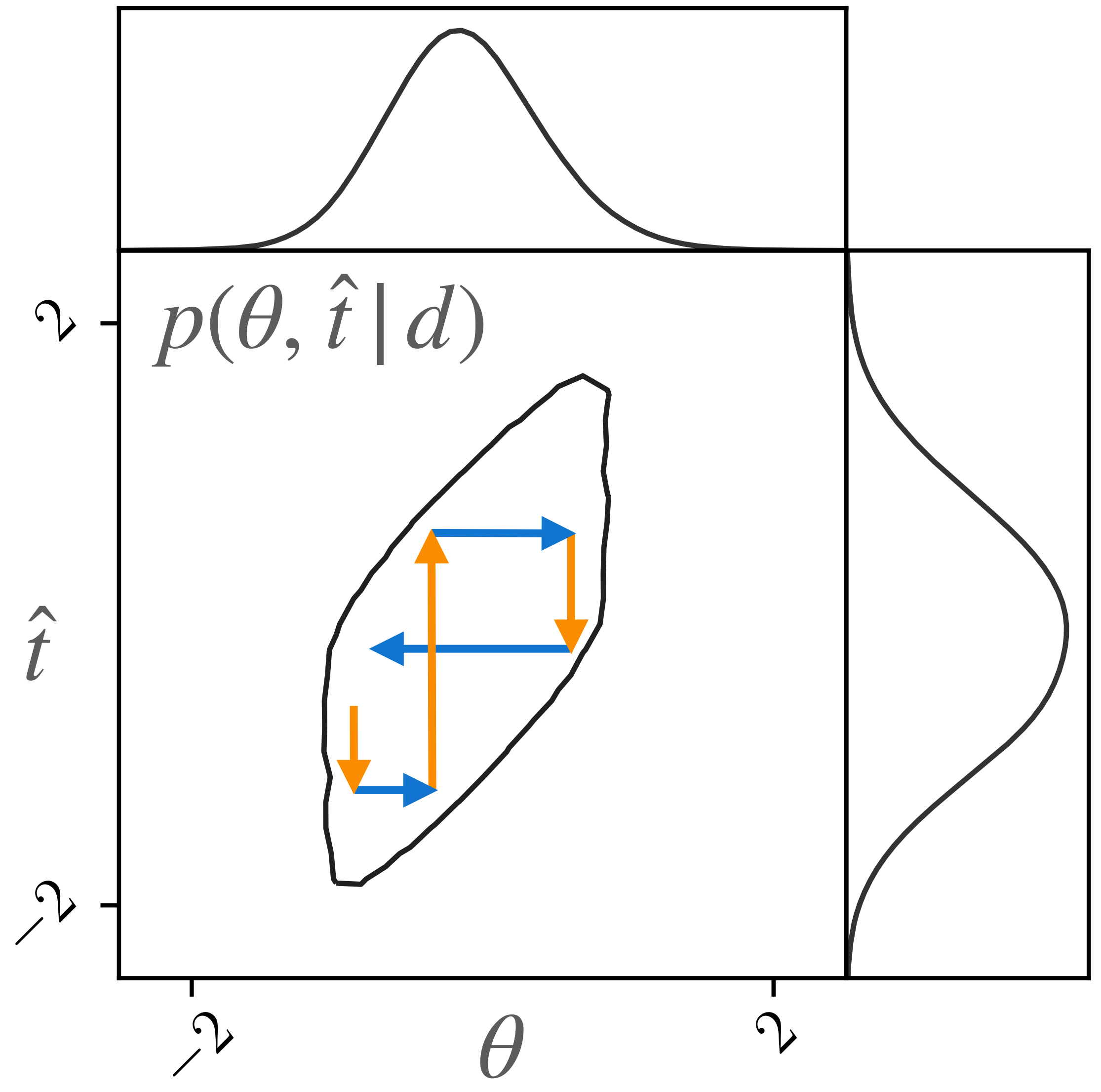
- Group-equivariant NPE (GNPE)
  - Integrate **symmetries via data-standardisation** (works with exact and approximate symmetries)
  - Requires iterative inference over nuisance parameters (⇒ slower)
  - For GWs: great **accuracy improvements**



Group equivariant neural posterior estimation (GNPE)

- Idea: **Standardize data d to $t = 0$**
  Problem: t unknown at inference time

- **Define** $\hat{t}$ via $p(\hat{t}\,|\,t) = U_{[-\delta,\delta]}\left(\hat{t} - t\right)$
- **Sample** $\left(\theta, \hat{t}\right) \sim p\left(\theta, \hat{t}\,|\,d\right)$
- **Marginalize** over $\hat{t}$



[$\theta = t$ for simplicity]

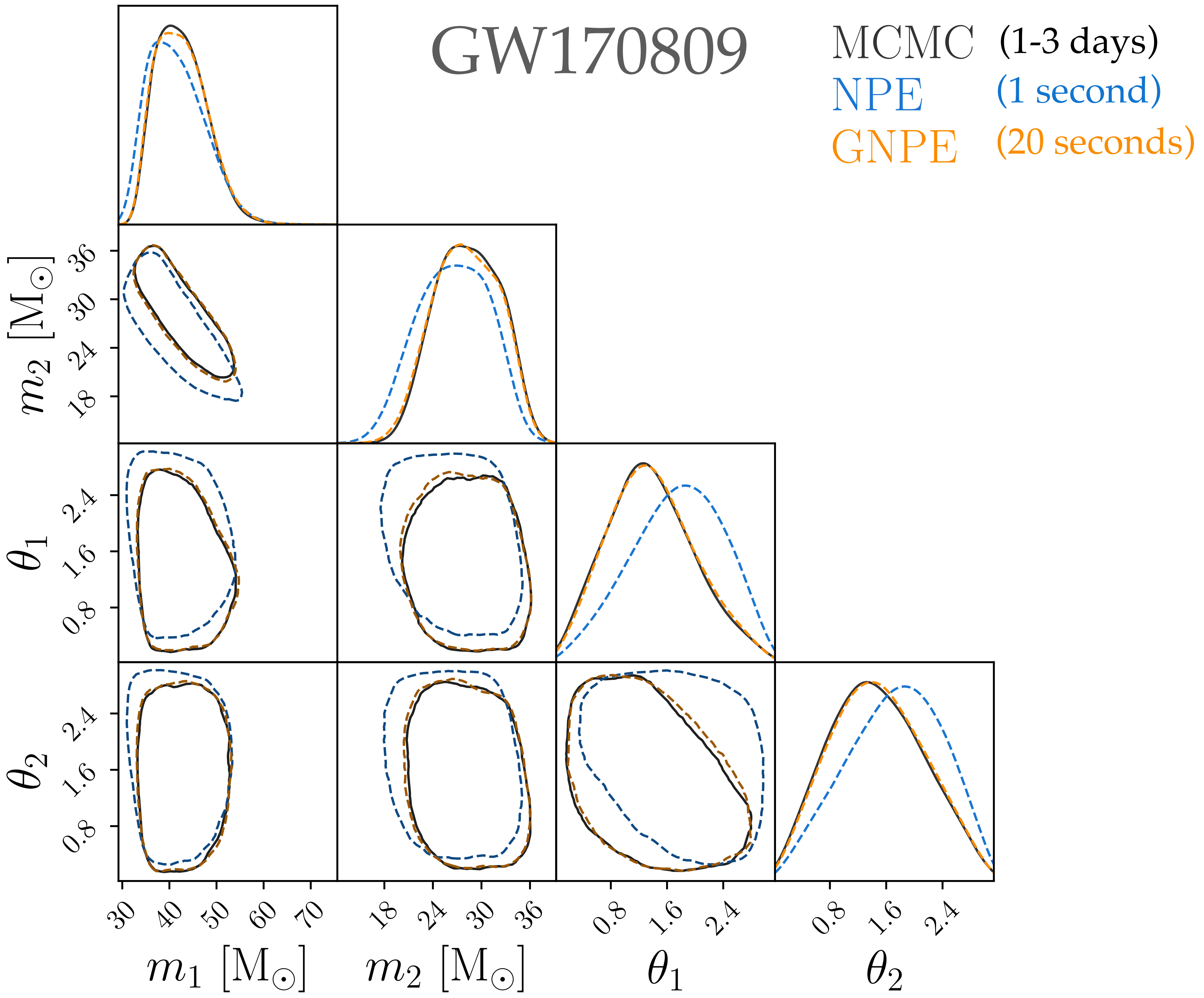
Gibbs sampling

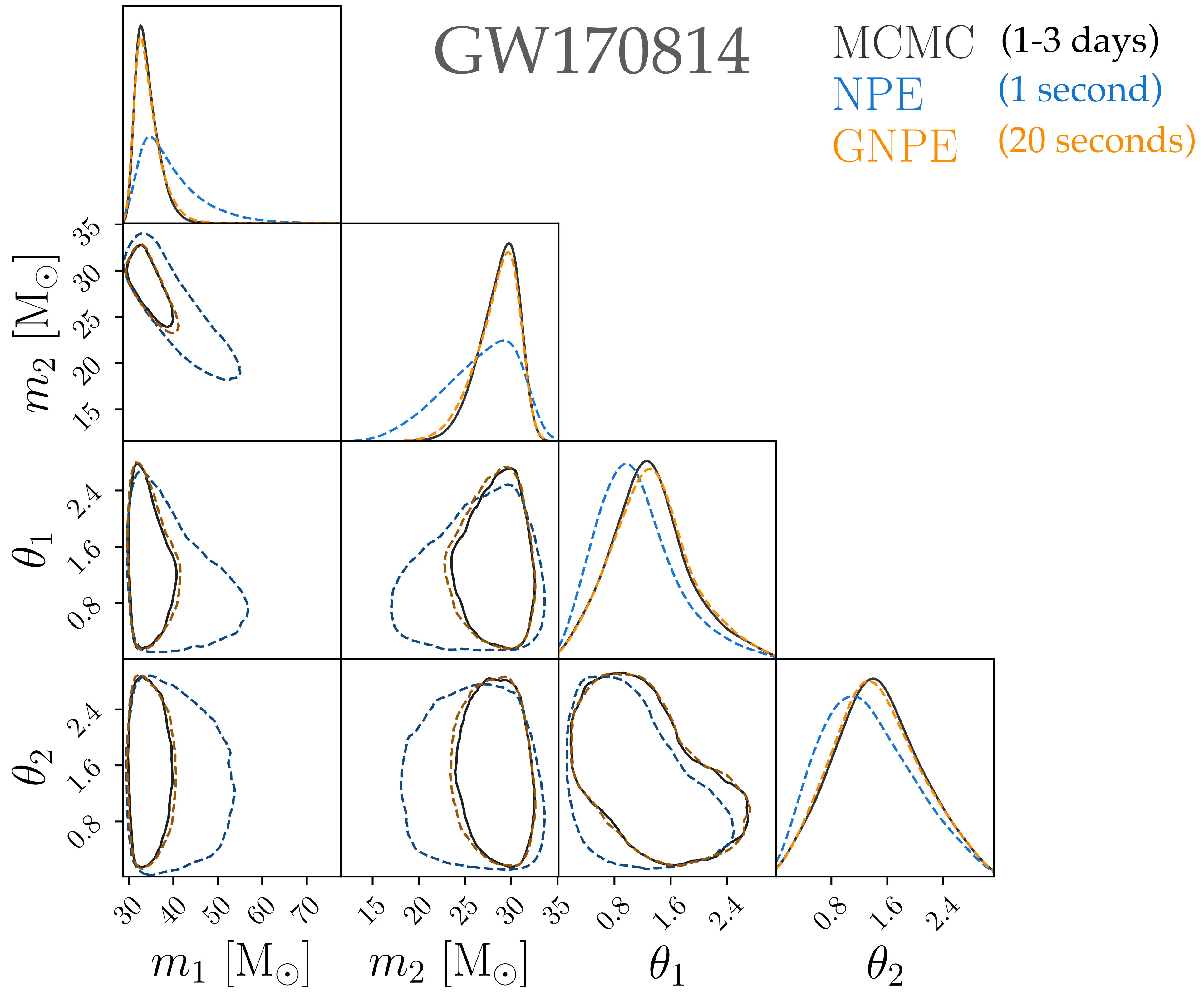
1) $\hat{t} \sim p(\hat{t}\,|\,d, \theta)$ via $\hat{t} = t + \epsilon, \quad \epsilon \sim U_{[-\delta,\delta]}(\epsilon)$ here: $\delta = 1$ ms
2) $\theta \sim p(\theta\,|\,d, \hat{t})$ via $q(\theta\,|\,d_{-\hat{t}}, \hat{t})$ with NPE

Data $d_{-\hat{t}}$ has time shifts of only $|t - \hat{t}| \leq \delta = 1$ ms!
(prior: $[-100, 100]$ ms)

Dax, Green, Gair, Macke, Buonanno, Schölkopf. *Real-time gravitational wave science with neural posterior estimation.* **Physical Review Letters**, 2021.
Dax, Green, Gair, Deistler, Schölkopf, Macke. *Group equivariant neural posterior estimation.* **ICLR**, 2022.

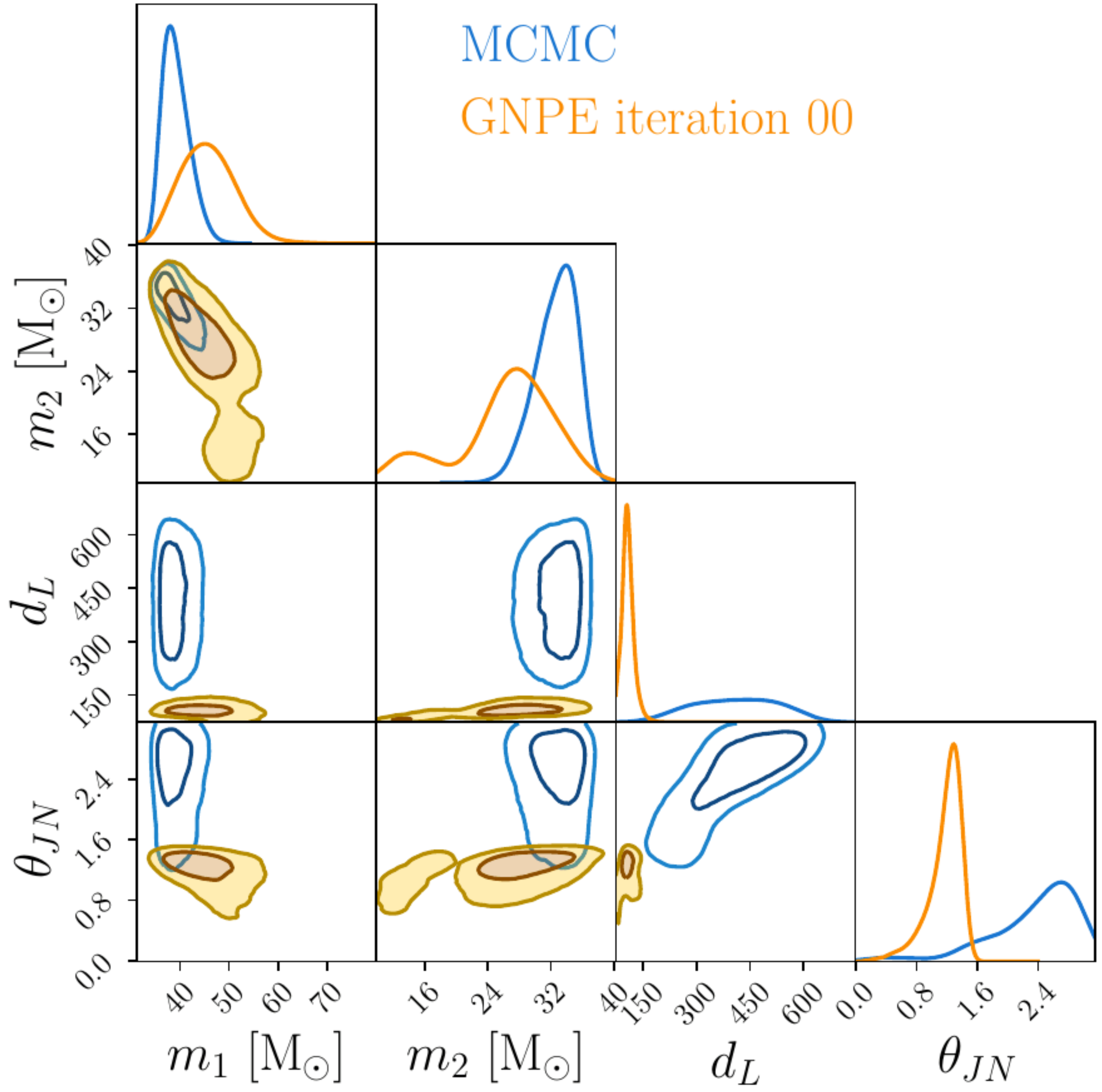
NPE vs. GNPE





Dax, Green, Gair, Macke, Buonanno, Schölkopf. *Real-time gravitational wave science with neural posterior estimation.* **Physical Review Letters**, 2021.
Dax, Green, Gair, Deistler, Schölkopf, Macke. *Group equivariant neural posterior estimation.* **ICLR**, 2022.

GNPE iterations



Real time

50k samples

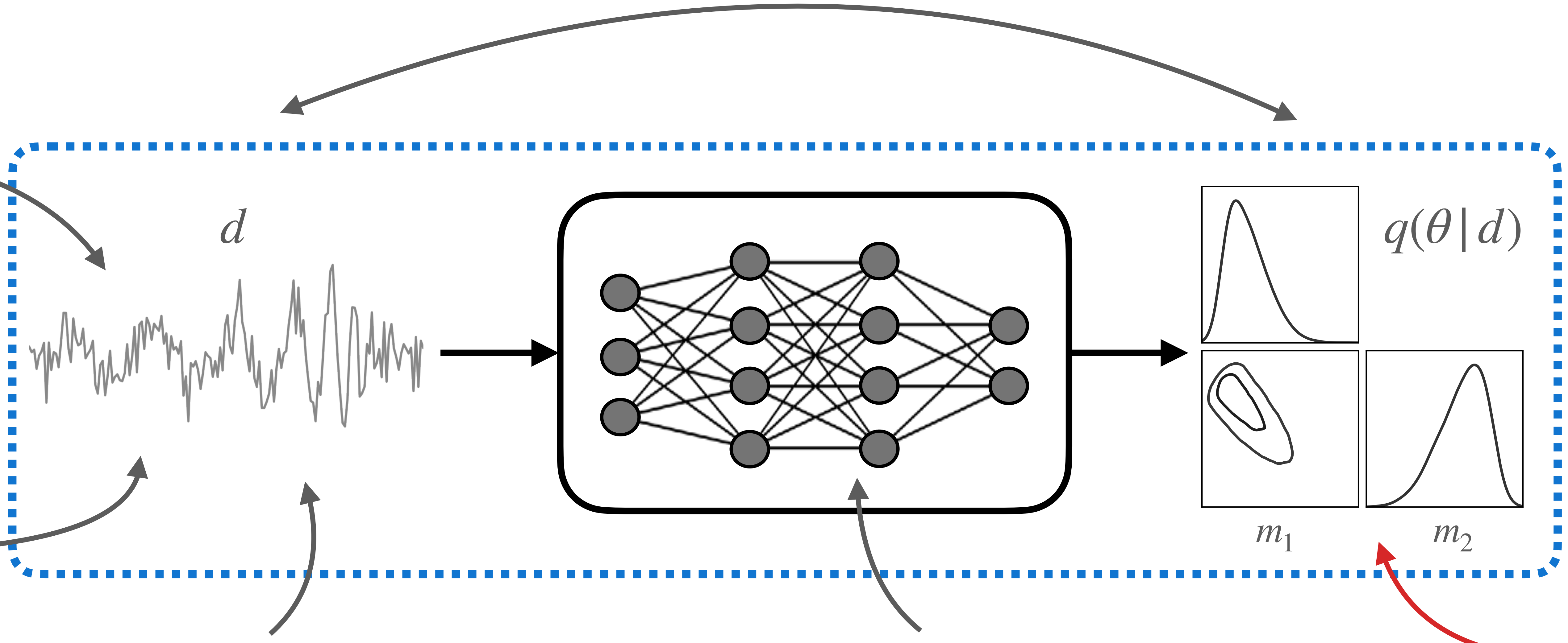
Sampling done in parallel,
$\sim N_{it.}$ times slower than NPE

Accuracy Guarantees

How can we integrate known relations between data and parameters?

How can we handle missing data?

How can we handle high dimensional data?



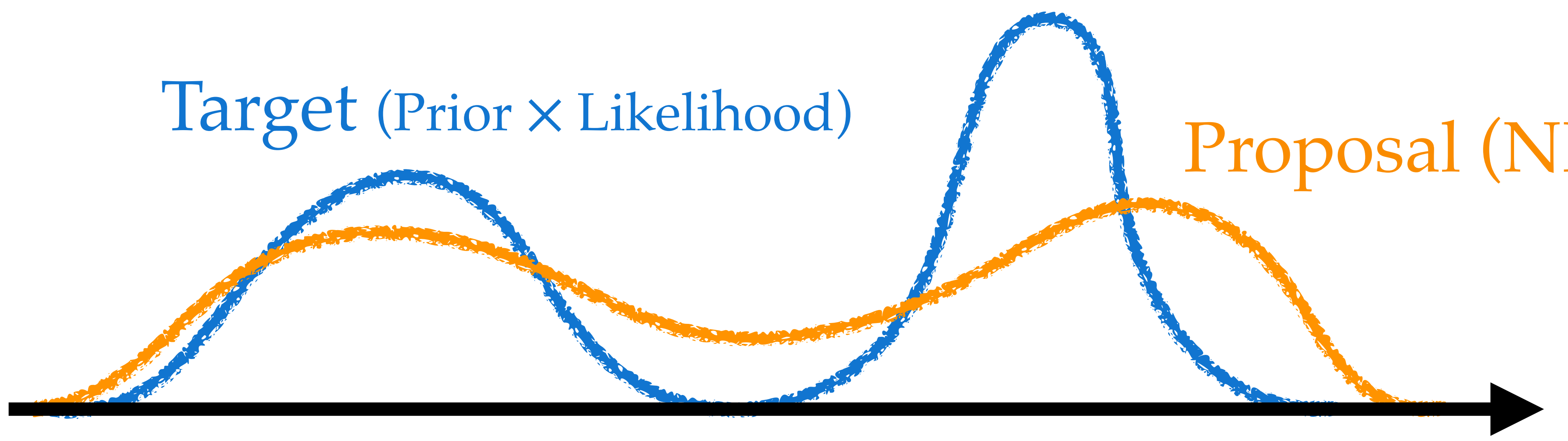
How can we adapt to changing detector noise properties?

How can we ensure neural network expressivity?

How can we guarantee accurate inference?

Verification and correction with **importance-sampled NPE (NPE-IS)**

- If likelihood is tractable, can reweight NPE results



$$\theta_i \sim q(\theta | d)$$

$$w_i = \frac{p(\theta_i)p(d|\theta_i)}{q(\theta_i|d)}$$

⇒ **asymptotically exact**

- **Effective number of samples** as performance metric

$$n_{\text{eff}} = \left(\Sigma_i w_i\right)^2 / \Sigma_i \left(w_i^2\right) \qquad \epsilon = n_{\text{eff}}/n \in (0,1]$$

⇒ **verification** without for ground truth posterior

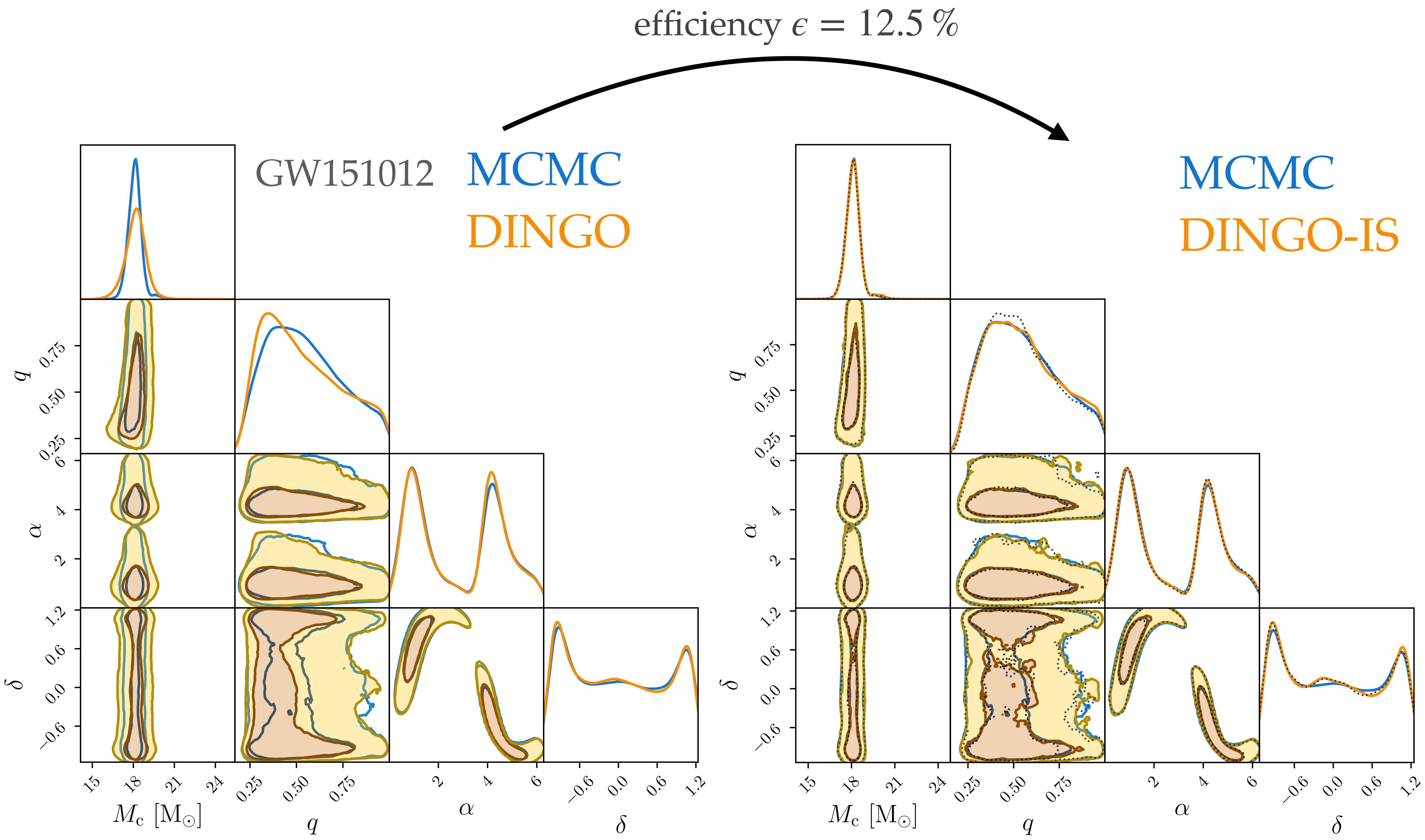
- Estimate of **Bayesian evidence**

$$p(d) = \frac{1}{n}\sum_i w_i \qquad \sigma_{\log p(d)} = \sqrt{(1-\epsilon)/(n\cdot\epsilon)}$$

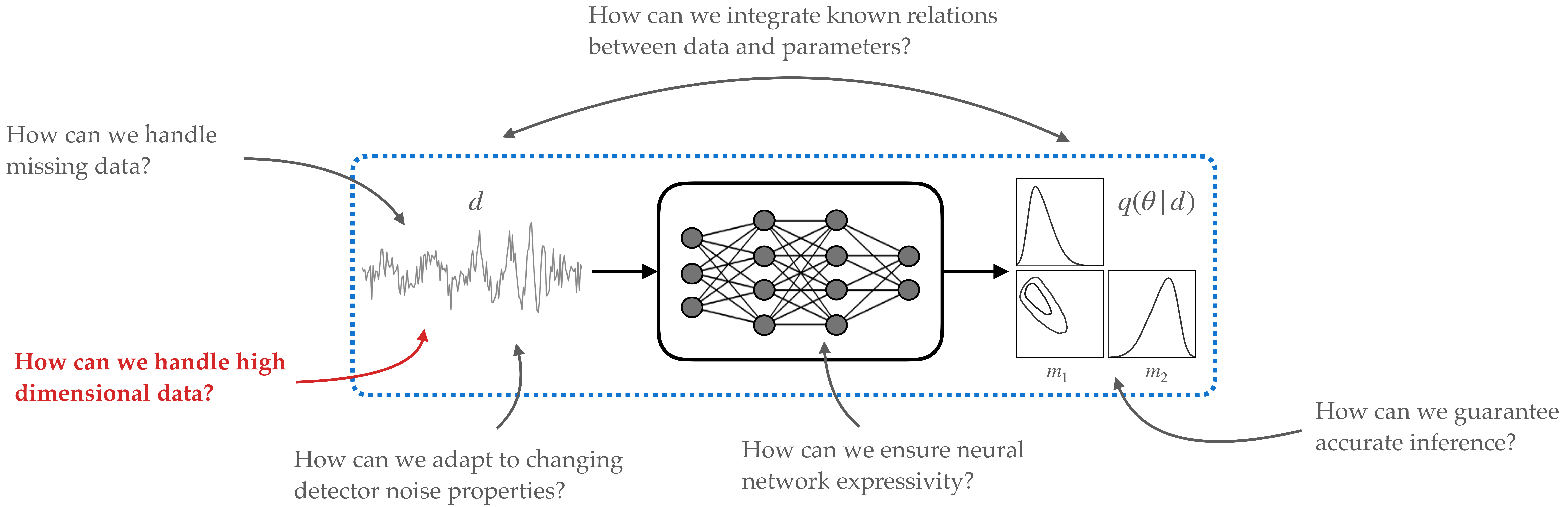
⇒ unbiased & precise estimate of **evidence**

NPE-IS results

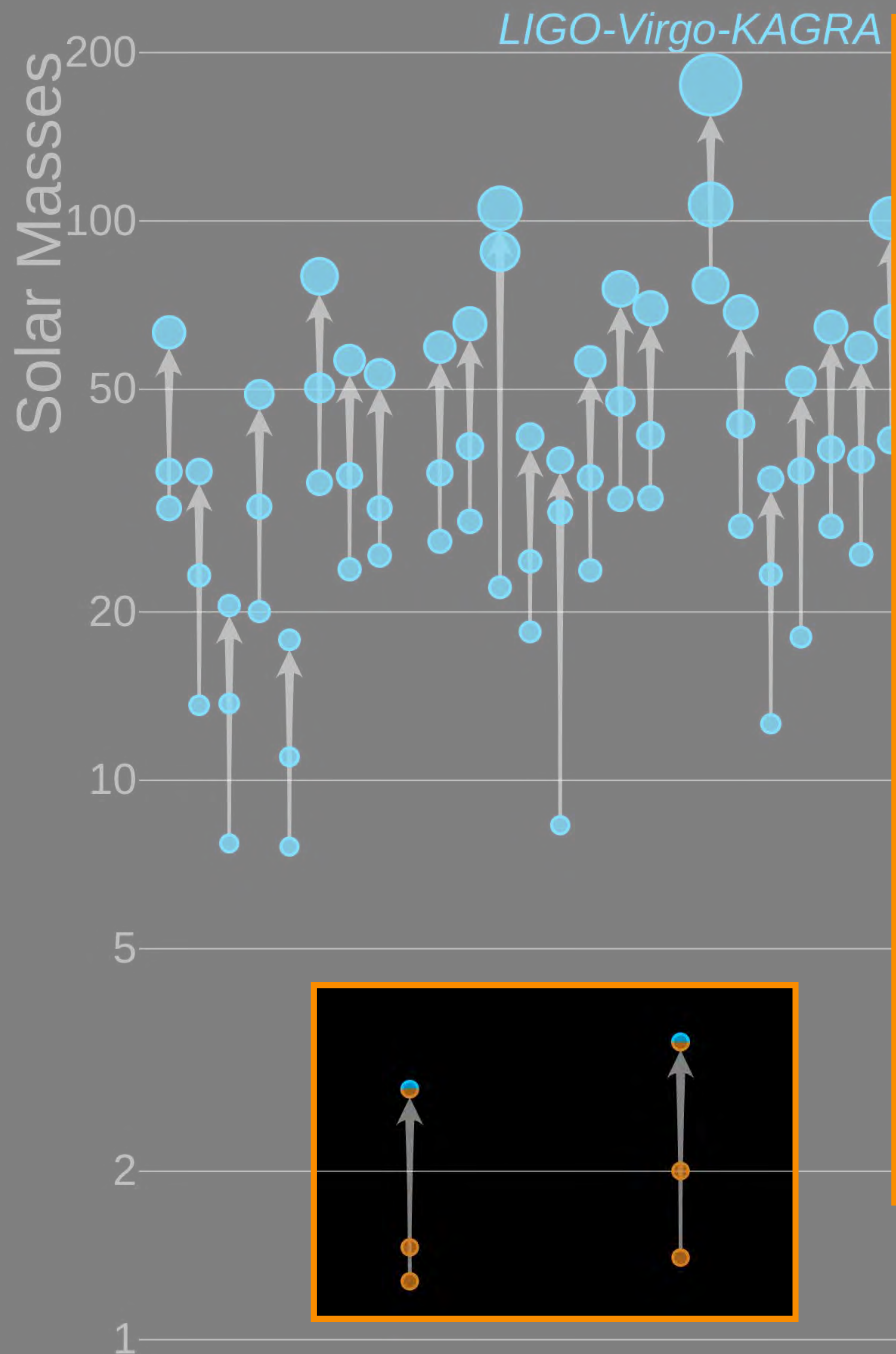
- Evaluation on 42 real GW events, **efficiencies of ≈ 10%**
- Can use GW models for which MCMC is too costly
- Low efficiencies **flag OOD data** and adversarial attacks
- **Evidences** consistent with nested sampling, but **10x more precise**



Fast Inference for Binary Neutron Stars



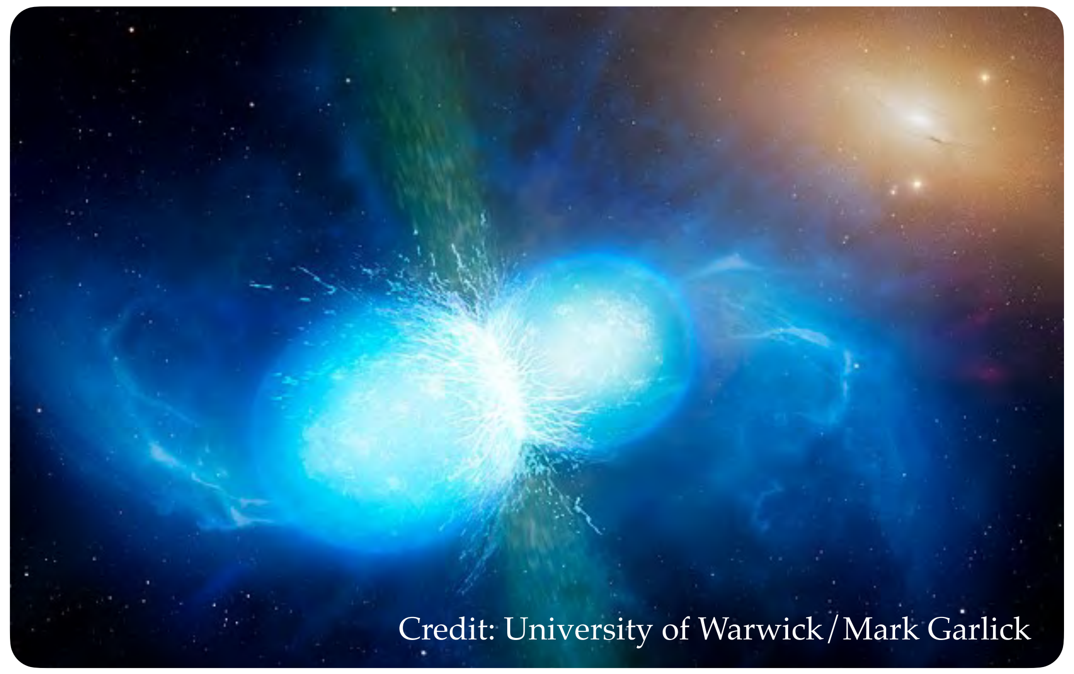
Masses in the Stellar Graveyard



Binary neutron stars (BNS)

- BNS kilonova emits **electromagnetic (EM) signals**
  ⇒ enables **Multi-Messenger Astronomy**

- Only **one successful EM observation** to date (2017), first observed **11 hours after the BNS merger**

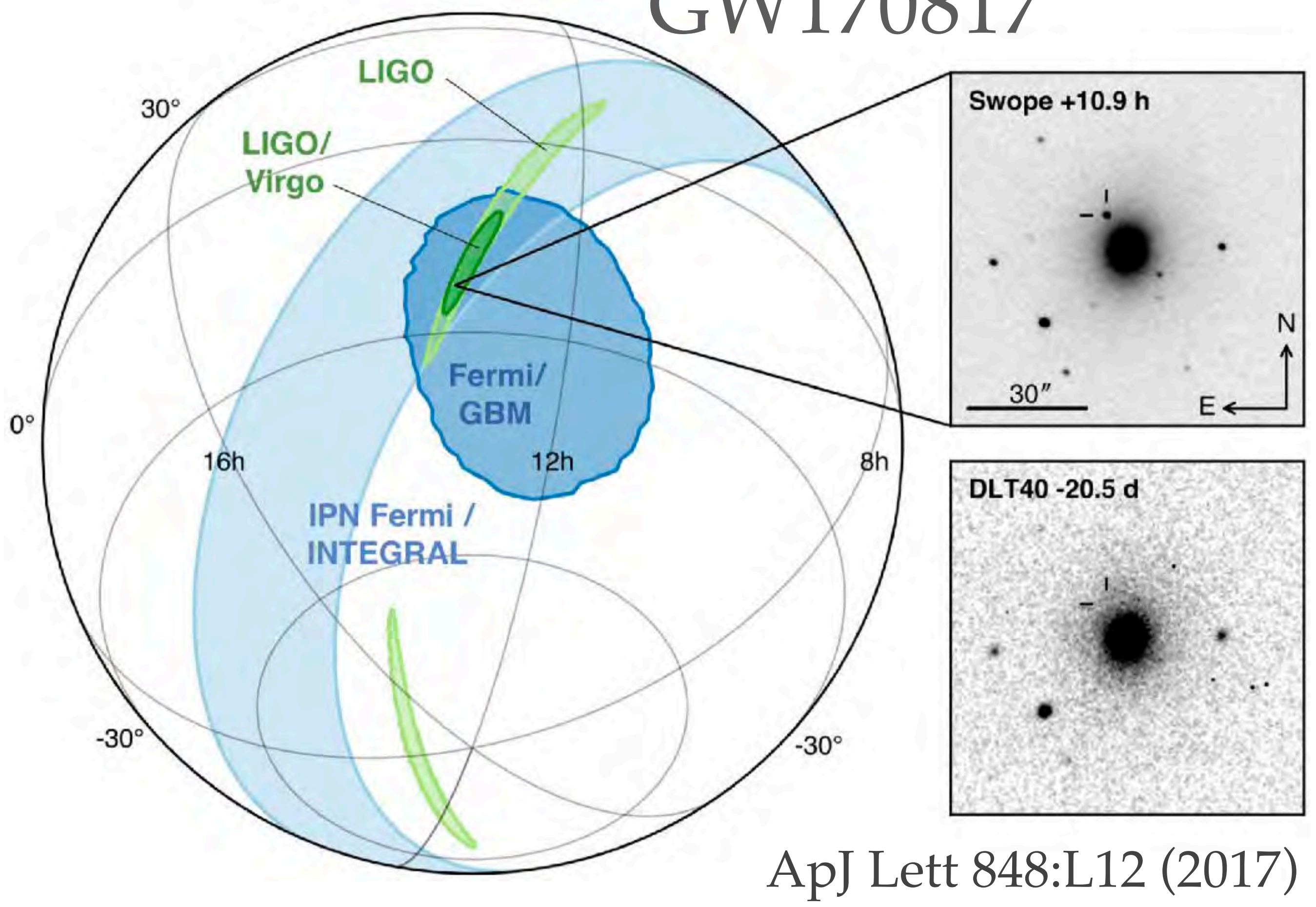


- EM observation relies on **BNS localization**
  ⇒ need fast & accurate GW inference

- Currently: accuracy sacrificed for speed

- Fast **ML inference** could help enable **more and earlier kilonova detections**

GW170817



ApJ Lett 848:L12 (2017)

Detection of BNS kilonovae is exciting;

this requires rapid and accurate GW inference;

which is really hard due to long signals.

Credit: A. Posada, MPI for intelligent systems

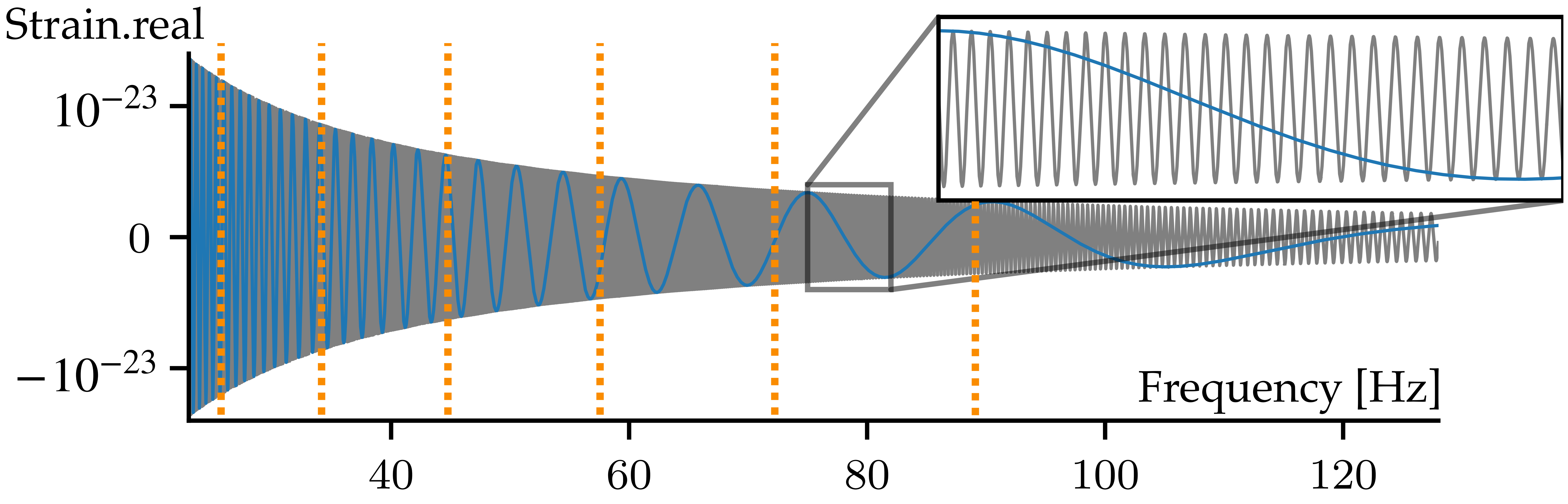
Main challenge: high dimensional data

BNS signals are extremely long, and thus high-dimensional (10M+ bins)

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

If we know the chirp mass M_c, could apply M_c-based compression

1. **Heterodyning** (Cornish 2010) — factor out overall phase $\propto (M_c^{\text{est}} f)^{-5/3}$
2. **Multibanding** (Vinciguerra+, 2017) — use reduced resolution at higher f

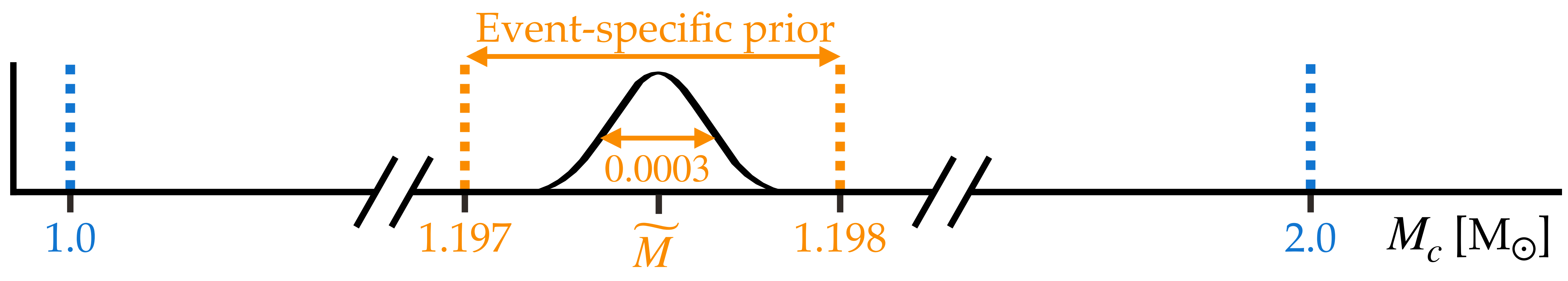


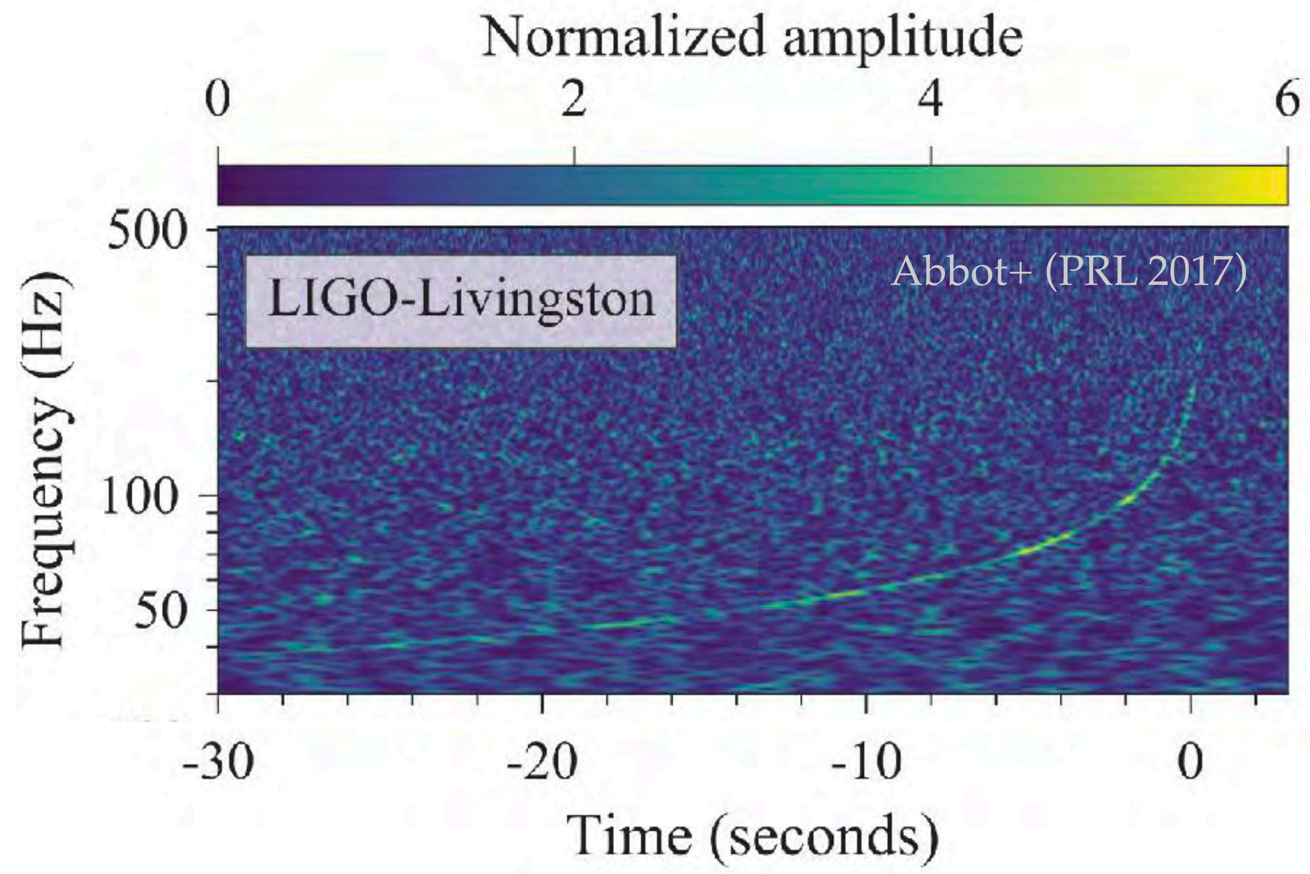
⇒ Loss-free compression by 100x

But: can't use M_c for compression, since this is an (unknown) inference parameter

Prior-conditioning

- BNS data constrains the chirp mass M_c extremely well
  - **Between events**: large chirp mass range, e.g. [1.0, 2.0]
  - **Specific event**: tightly constrained posterior





- **Prior-conditioning** enables prior-tunable NPE networks

1) Sample the training prior hierarchically $\widetilde{M}_i \sim \hat{p}(\widetilde{M}),\ \theta_i \sim p_{\widetilde{M}_i}(\theta)$

2) Condition NPE network on choice of prior $q(\theta \,|\, d, \widetilde{M})$

3) Compress based on prior center $d \rightarrow d_{\widetilde{M}}$

For BNS:

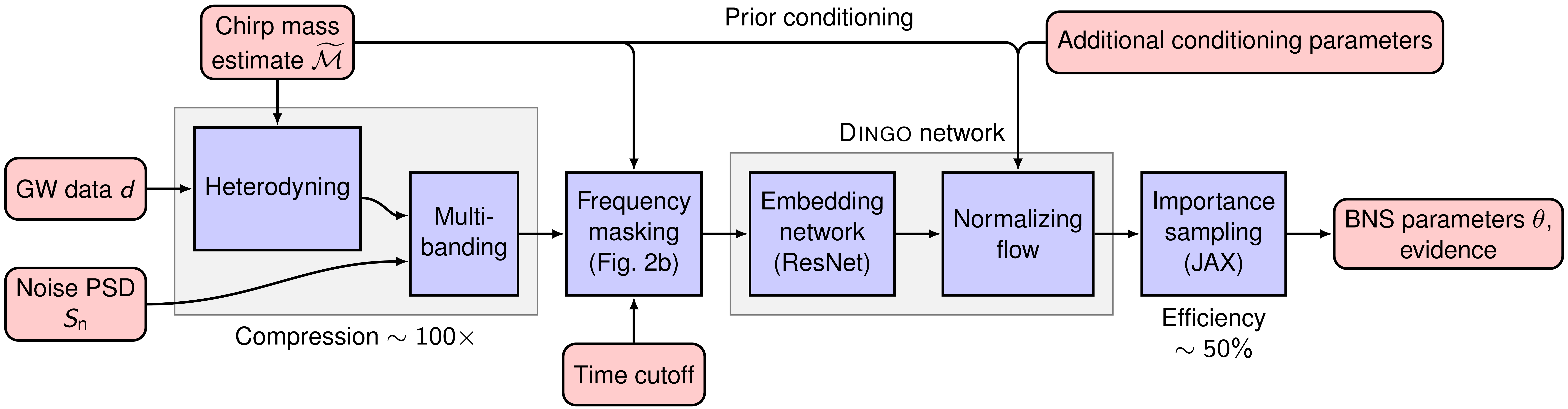
$\hat{p}(\widetilde{M}) = U[1.0, 2.0]\ \mathrm{M}_\odot$

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

⇒ The lightweight version of GNPE.

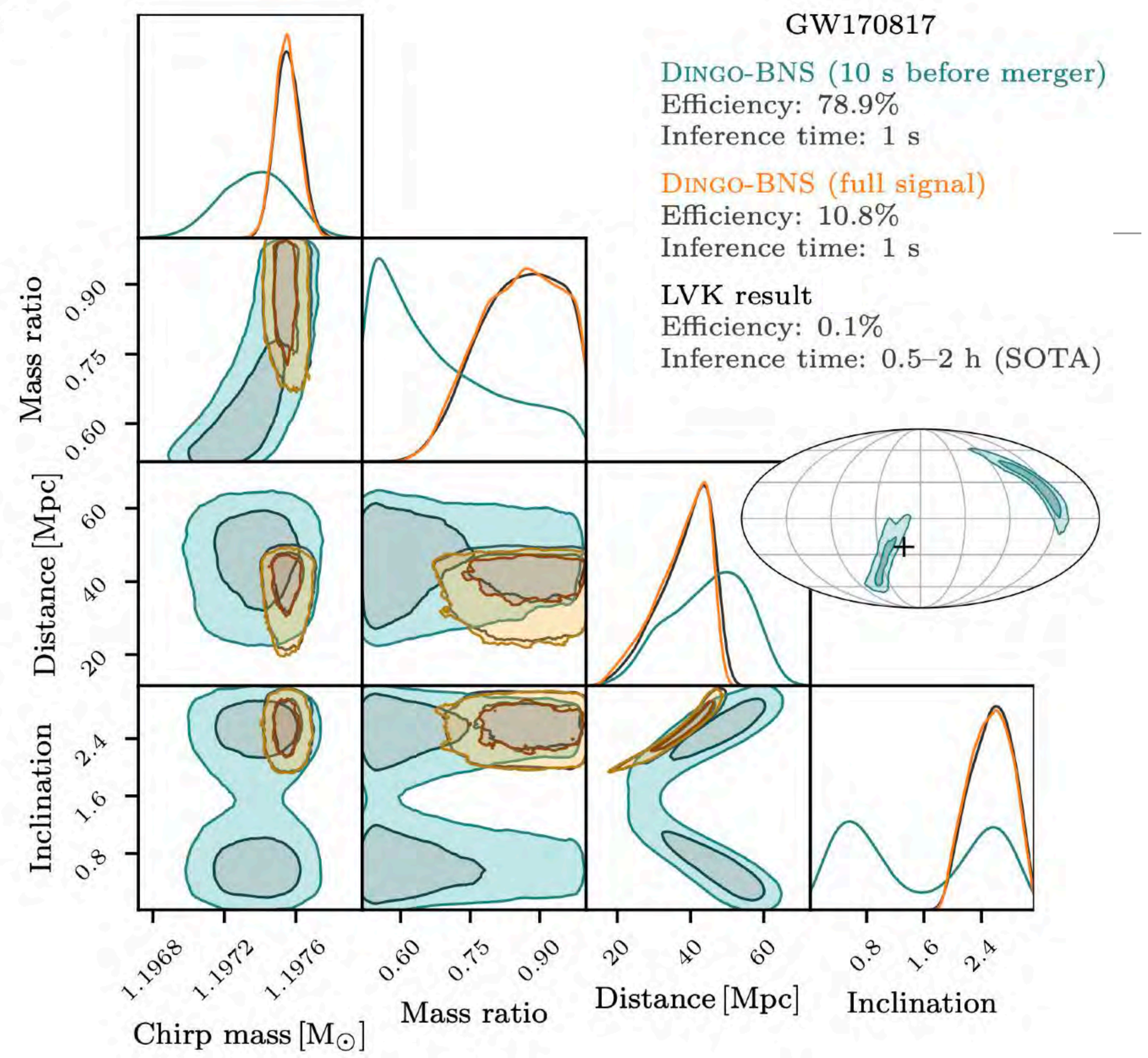
Technical Extensions

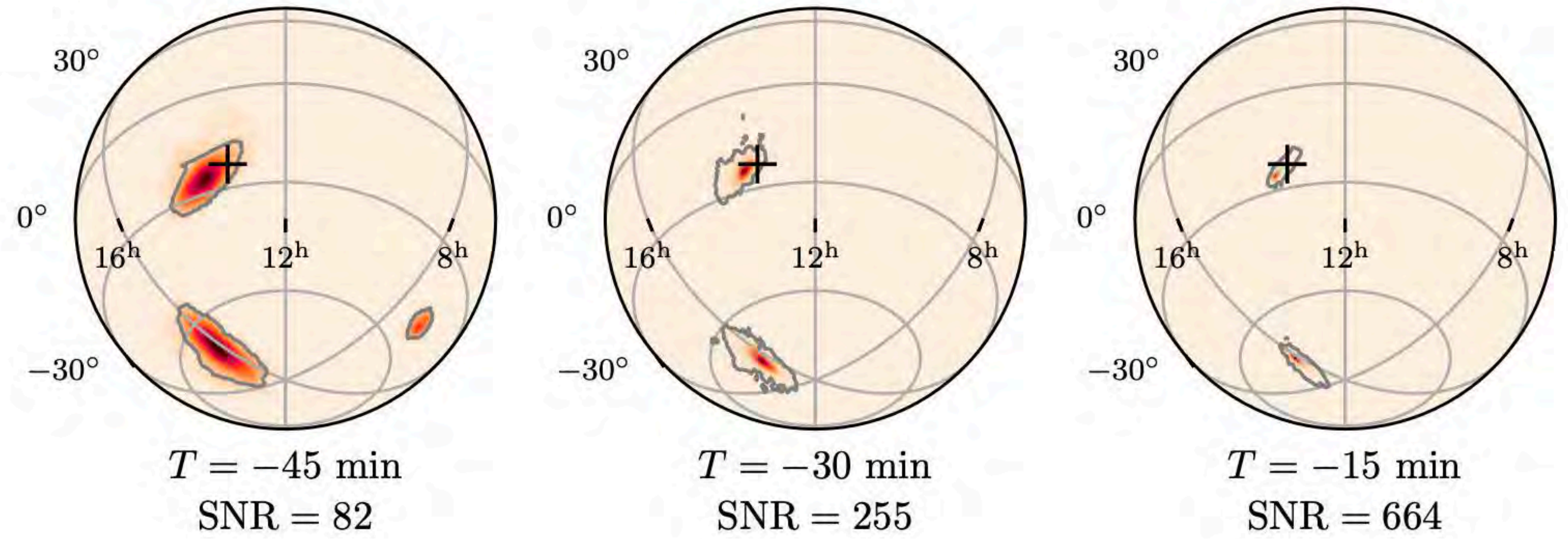
- Prior-conditioning for event **adaptive data compression**
- Data masking during training to enable **inference with partial data**
- Conditioning on parameter subset to enable instant **integration of external results**
- JAX implementation of likelihood for **importance sampling in 0.5 s**

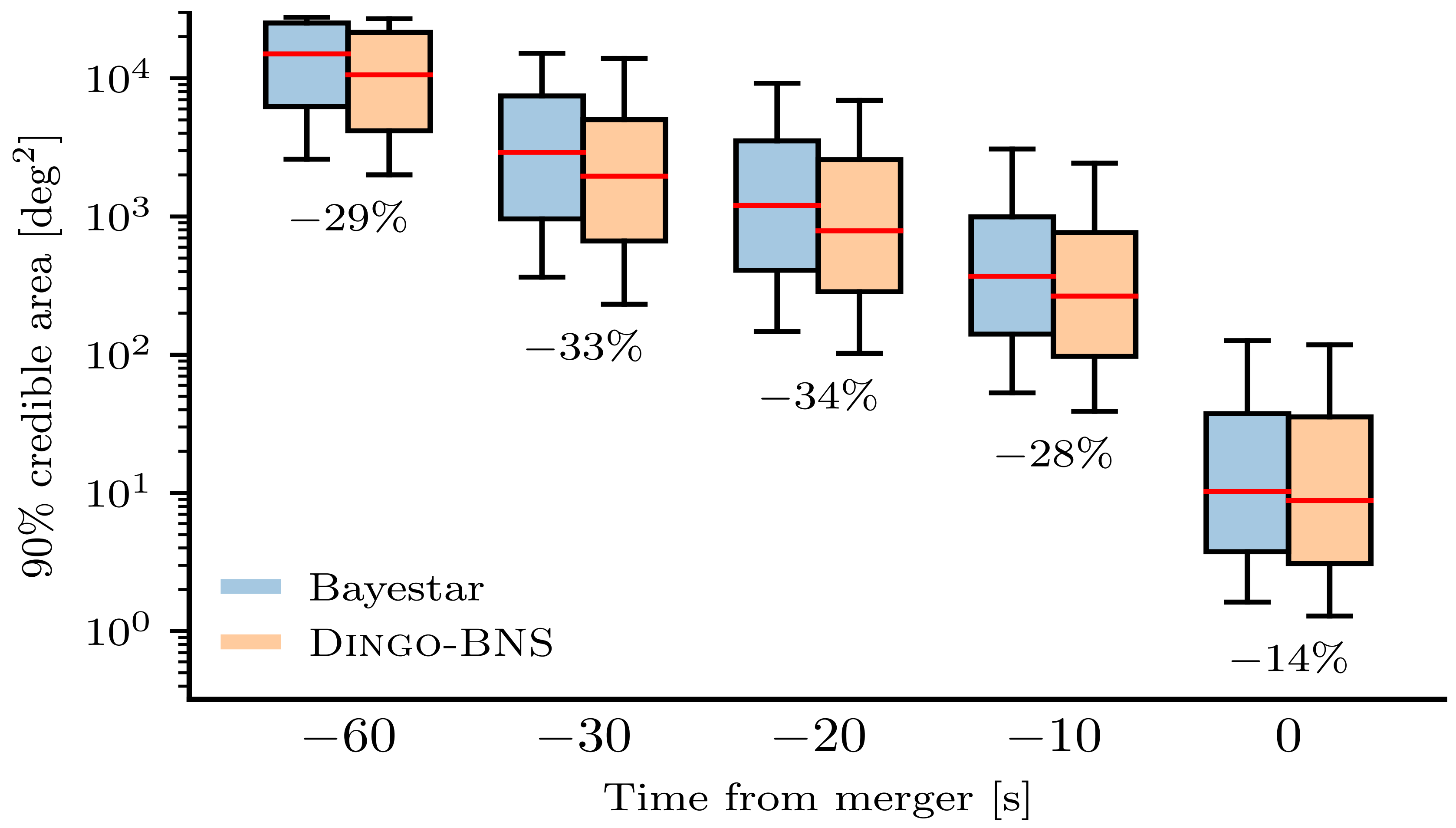


BNS: Results

- DINGO-BNS reproduces public LVK results with only **1 second inference time**
- Inference at arbitrary times **before to the merger**
- Complete inference without approximations
  ⇒ **30% improvement in low-latency localization**
- Scales to hour-long signals of **next-gen detectors**



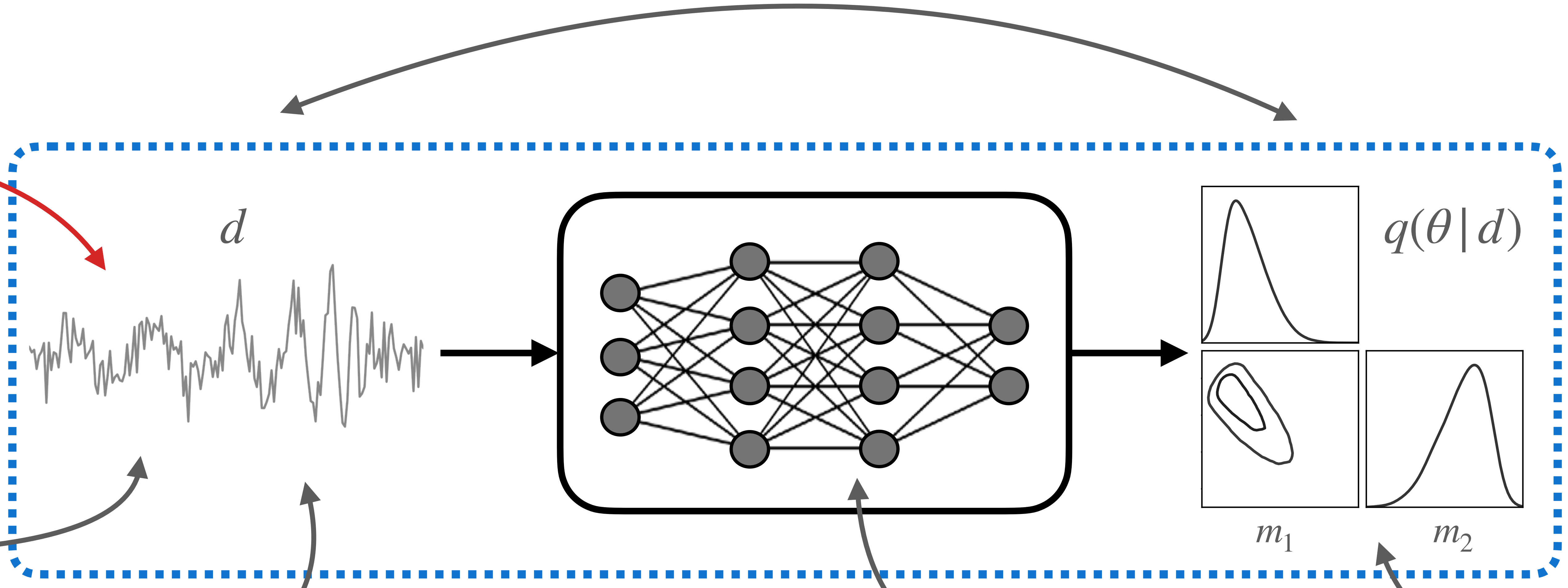




Flexible Inference with Transformers

How can we integrate known relations between data and parameters?

How can we handle missing data?



How can we handle high dimensional data?

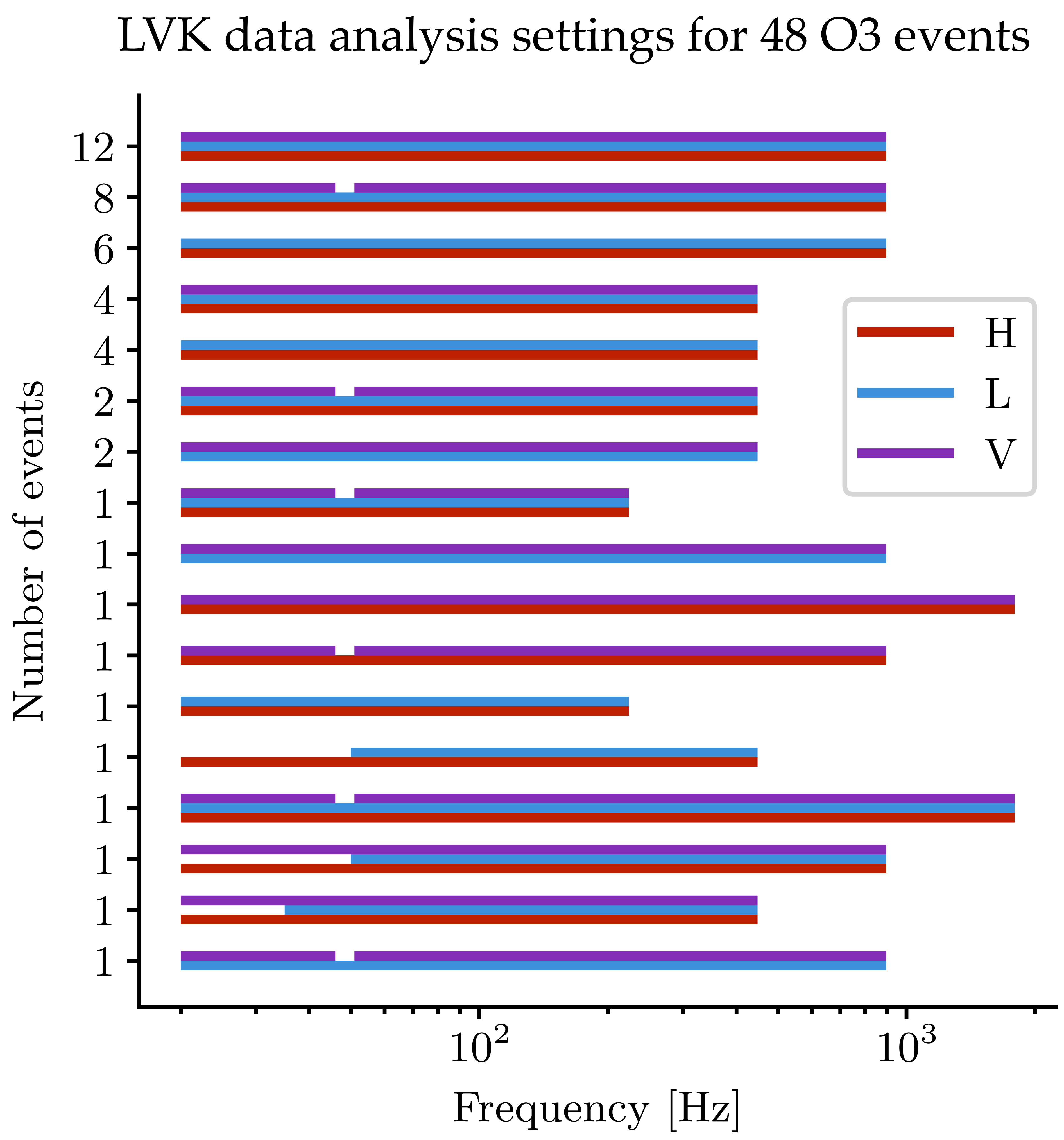
How can we adapt to changing detector noise properties?

How can we ensure neural network expressivity?

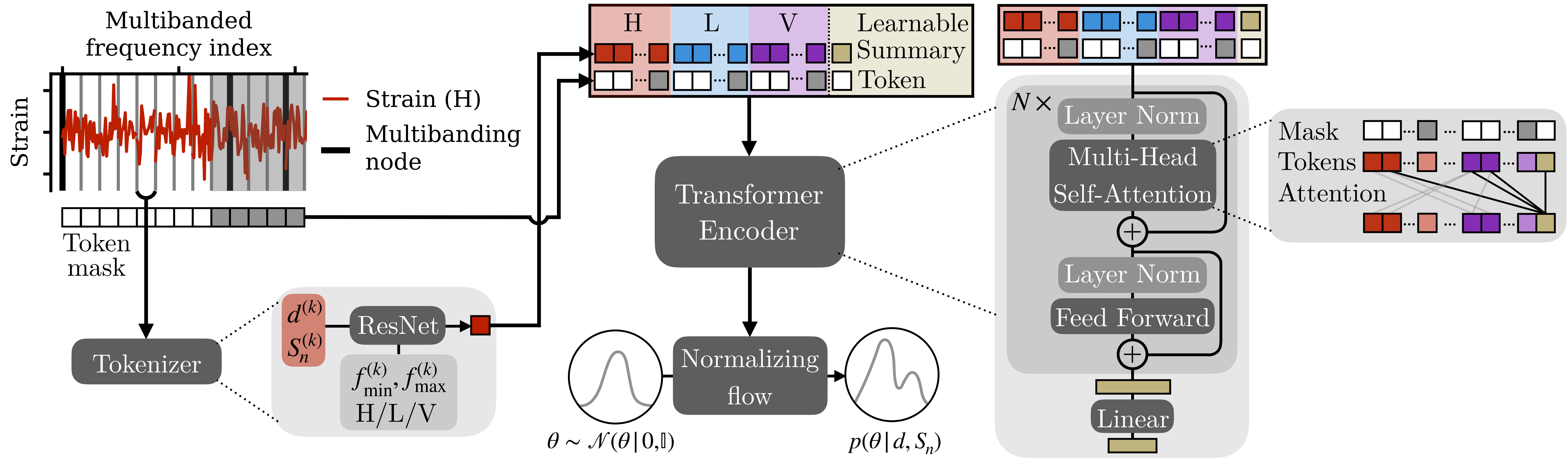
How can we guarantee accurate inference?

Heterogeneous GW data

- GW data is represented as a set of frequency series in multiple detectors (here: H, L, V)
- In many cases, not all data is available (e.g., a detector may be offline)
- Most neural networks have fixed input dimension
  → trained network accommodates only a single configuration
- In practice, **would need to train one DINGO network per analysis configuration**



Inference with Heterogeneous GW data: DINGO-T1

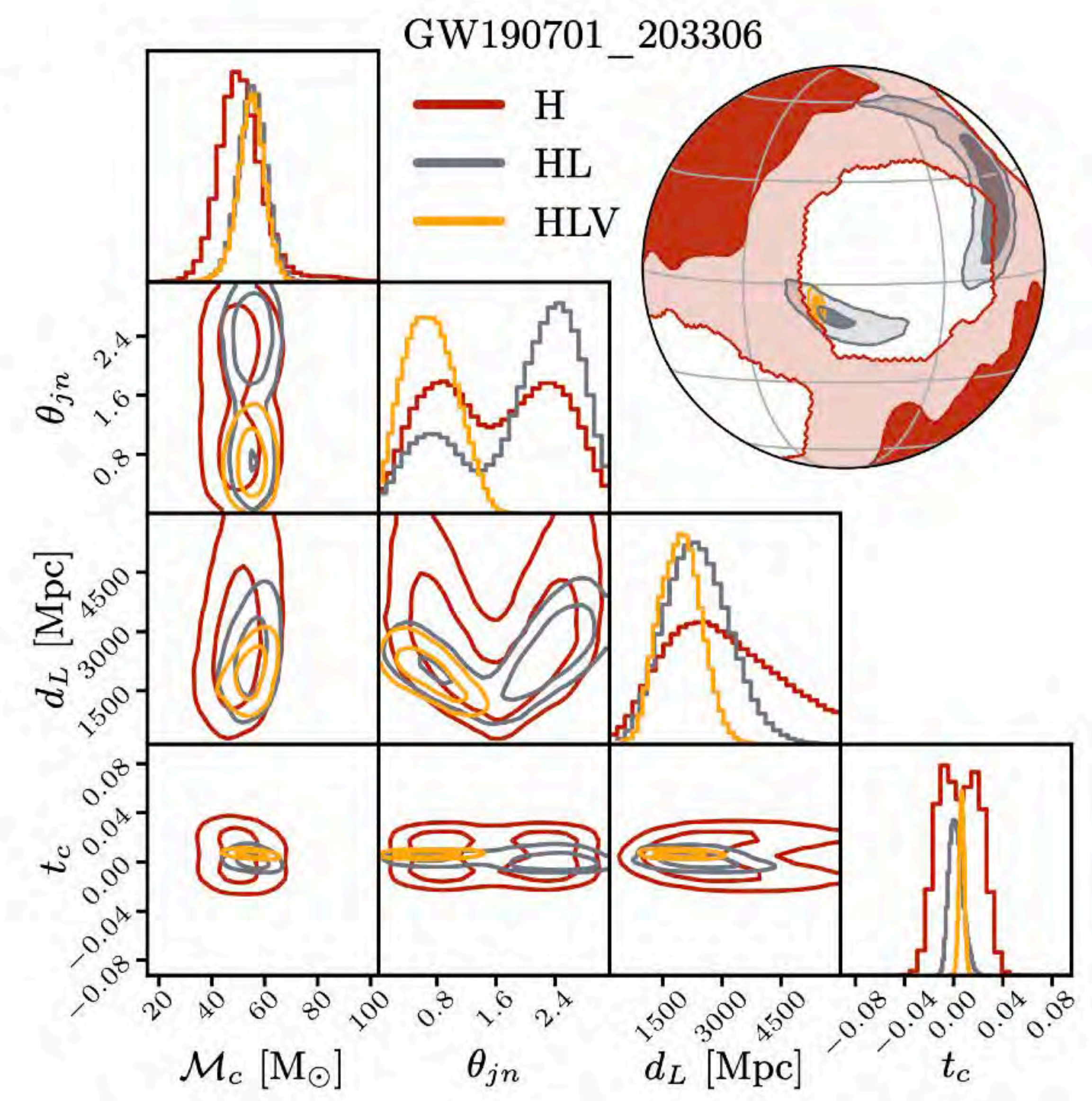


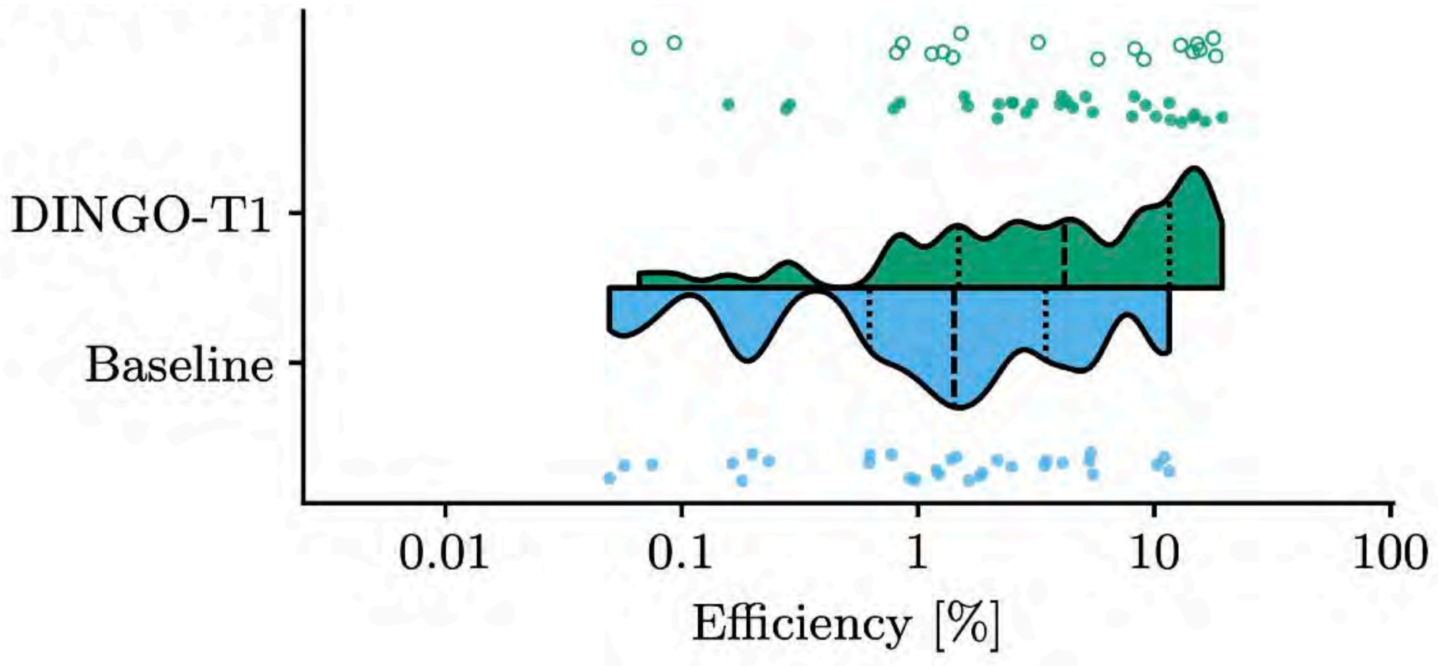
- Partition **GW data** into a **set of** smaller segments (**tokens**; ~ corresponding to a single word in an LLM)
- Process token sequence with a **transformer** (can handle sequences of arbitrary length)
- During training, **randomly mask out tokens**
  → Inference network **learns to deal with partial data**

DINGO-T1 results

- **DINGO-T1** is a single model that accommodates **flexible data configurations** (frequency ranges and detector combinations) **without retraining**

- It **outperforms non transformer-based models** in terms of accuracy (partially due to scaling)





48 O3 events analysed in all possible detector configurations, using only a single model

Event	Detectors	HLV	HL	HV	LV	H	L	V
GW190408_181802	HLV	8.07 %	8.28 %	0.64 %	1.35 %	2.79 %	6.71 %	62.23 %
GW190413_052954	HLV	11.77 %	9.52 %	2.41 %	0.74 %	9.5 %	18.61 %	37.48 %
GW190413_134308$^{*\dagger L}$	HLV	4.52 %	7.84 %	8.29 %	1.36 %	21.16 %	11.78 %	32.2 %
GW190426_190642	HLV	0.79 %	2.27 %	7.41 %	0.0 %	29.53 %	0.04 %	49.52 %
GW190503_185404^{*L}	HLV	1.57 %	1.31 %	1.0 %	0.41 %	17.09 %	2.14 %	60.85 %
GW190513_205428^{*L}	HLV	0.16 %	0.52 %	4.19 %	2.04 %	11.74 %	4.17 %	22.9 %
GW190517_055101	HLV	2.17 %	4.28 %	1.79 %	0.27 %	1.36 %	0.36 %	13.06 %
GW190519_153544	HLV	4.26 %	0.85 %	3.71 %	4.88 %	3.38 %	9.09 %	59.66 %
GW190602_175927	HLV	13.09 %	13.18 %	28.09 %	5.4 %	27.24 %	15.95 %	26.58 %
GW190701_203306^{*L}	HLV	14.83 %	14.05 %	9.13 %	5.45 %	31.43 %	5.93 %	38.78 %
GW190706_222641	HLV	2.49 %	5.65 %	13.73 %	6.21 %	0.1 %	10.14 %	27.66 %
GW190727_060333$^{\dagger L}$	HLV	0.84 %	17.6 %	2.95 %	3.29 %	5.09 %	0.45 %	23.71 %
GW190803_022701	HLV	16.39 %	28.23 %	19.48 %	9.04 %	28.74 %	7.89 %	42.92 %
GW190828_063405	HLV	3.99 %	17.35 %	8.7 %	7.11 %	24.58 %	9.26 %	64.41 %
GW190915_235702	HLV	8.23 %	13.56 %	0.89 %	5.24 %	1.9 %	14.02 %	42.52 %
GW190916_200658	HLV	19.41 %	20.19 %	5.16 %	18.13 %	11.2 %	27.25 %	42.29 %
GW190926_050336	HLV	1.63 %	4.25 %	11.17 %	12.38 %	16.39 %	19.97 %	53.75 %
GW190929_012149	HLV	3.04 %	4.74 %	10.61 %	0.0 %	24.68 %	0.01 %	50.24 %
GW191127_050227$^{\diamond H}$	HLV	0.28 %	0.97 %	8.21 %	5.15 %	16.13 %	28.83 %	6.04 %
GW191215_223052	HLV	4.05 %	5.79 %	5.13 %	1.25 %	6.55 %	2.02 %	56.09 %
GW191230_180458	HLV	14.61 %	8.4 %	2.68 %	16.51 %	7.79 %	16.88 %	58.75 %
GW200129_065458$^{\ddagger L}$	HLV	0.29 %	0.29 %	2.16 %	0.01 %	14.15 %	0.07 %	2.66 %
GW200208_130117	HLV	11.58 %	12.15 %	5.17 %	5.99 %	32.27 %	8.9 %	17.75 %
GW200208_222617	HLV	2.86 %	2.47 %	4.73 %	0.05 %	5.55 %	18.82 %	18.48 %
GW200209_085452	HLV	2.52 %	11.06 %	3.65 %	16.62 %	9.66 %	34.08 %	41.28 %
GW200216_220804	HLV	2.2 %	12.56 %	12.41 %	7.44 %	21.27 %	10.96 %	52.1 %
GW200219_094415	HLV	5.12 %	7.36 %	3.27 %	2.77 %	11.76 %	11.09 %	1.9 %
GW200220_061928	HLV	10.17 %	18.1 %	6.07 %	7.68 %	8.67 %	11.95 %	61.93 %
GW200224_222234	HLV	5.49 %	7.18 %	1.65 %	6.85 %	10.2 %	20.89 %	4.01 %
GW200311_115853	HLV	9.2 %	5.56 %	11.57 %	4.81 %	7.71 %	4.23 %	3.64 %
GW190421_213856	HL	-	17.8 %	-	-	10.08 %	17.4 %	-
GW190514_065416$^{*\dagger L}$	HL	-	15.22 %	-	-	19.48 %	4.83 %	-
GW190521_074359	HL	-	1.28 %	-	-	7.24 %	1.26 %	-
GW190527_092055	HL	-	9.04 %	-	-	2.95 %	5.56 %	-
GW190719_215514	HL	-	14.56 %	-	-	17.56 %	0.05 %	-
GW190731_140936	HL	-	18.22 %	-	-	38.6 %	4.69 %	-
GW191109_010717$^{\diamond HL}$	HL	-	1.52 %	-	-	0.11 %	5.15 %	-
GW191204_110529	HL	-	0.07 %	-	-	19.49 %	13.28 %	-
GW191222_033537	HL	-	15.61 %	-	-	6.84 %	16.18 %	-
GW200128_022011	HL	-	8.28 %	-	-	38.67 %	18.79 %	-
GW200220_124850	HL	-	12.94 %	-	-	23.86 %	15.56 %	-
GW200306_093714	HL	-	0.09 %	-	-	6.0 %	0.08 %	-
GW190925_232845	HV	-	-	0.86 %	-	8.16 %	-	16.97 %
GW200302_015811	HV	-	-	1.4 %	-	2.25 %	-	31.8 %
GW190620_030421	LV	-	-	-	0.81 %	-	1.5 %	55.53 %
GW190630_185205	LV	-	-	-	1.15 %	-	7.98 %	16.08 %
GW190910_112807	LV	-	-	-	3.22 %	-	4.2 %	19.1 %
GW200112_155838	LV	-	-	-	5.79 %	-	16.7 %	20.93 %

Open questions

- **LIGO-Virgo-KAGRA**: many problems already solved, but some challenges remain
  → **close final gaps in parameter space** (e.g., neutron star-black hole mergers)
  → **improve accuracy**, in particular for OOD data and complex waveform models
  → **enhance flexibility** of pretrained inference networks

- **Next generation** detectors (Einstein Telescope, Cosmic Explorer, LISA)
  → **Extremely long signals**
  → **Non-stationary detector noise** (and lack of a good noise model)
  → **Overlapping signals**
  → Accounting for **Earth rotation**
  → **Data gaps**
  → **Needle-in-the-haystack** problems (e.g., extreme mass ratio inspirals)

mdax@tue.mpg.de

Thanks for your attention!

Publications featured in this talk

Dax, Green, Gair, Macke, Buonanno, Schölkopf. *Real-time gravitational wave science with neural posterior estimation.* **Physical Review Letters**, 2021. 🔗
Dax, Green, Gair, Deistler, Schölkopf, Macke. *Group equivariant neural posterior estimation.* **ICLR**, 2022. 🔗
Dax, Green, Gair, Pürrer, Wildberger, Macke, Buonanno, Schölkopf. *Neural Importance Sampling for Rapid and Reliable Gravitational-Wave Inference.* **Physical Review Letters**, 2023. 🔗
Dax, Green, Gair, Gupte, Pürrer, Raymond, Wildberger, Macke, Buonanno, Schölkopf. *Real-time inference for binary neutron star mergers using machine learning*. **Nature**, 2025. 🔗
Kofler, **Dax**, Green, Wildberger, Gupte, Macke, Gair, Buonanno, Schölkopf. Flexible Gravitational-Wave Parameter Estimation with Transformers. **Preprint**, 2025. 🔗

DINGO Pack


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