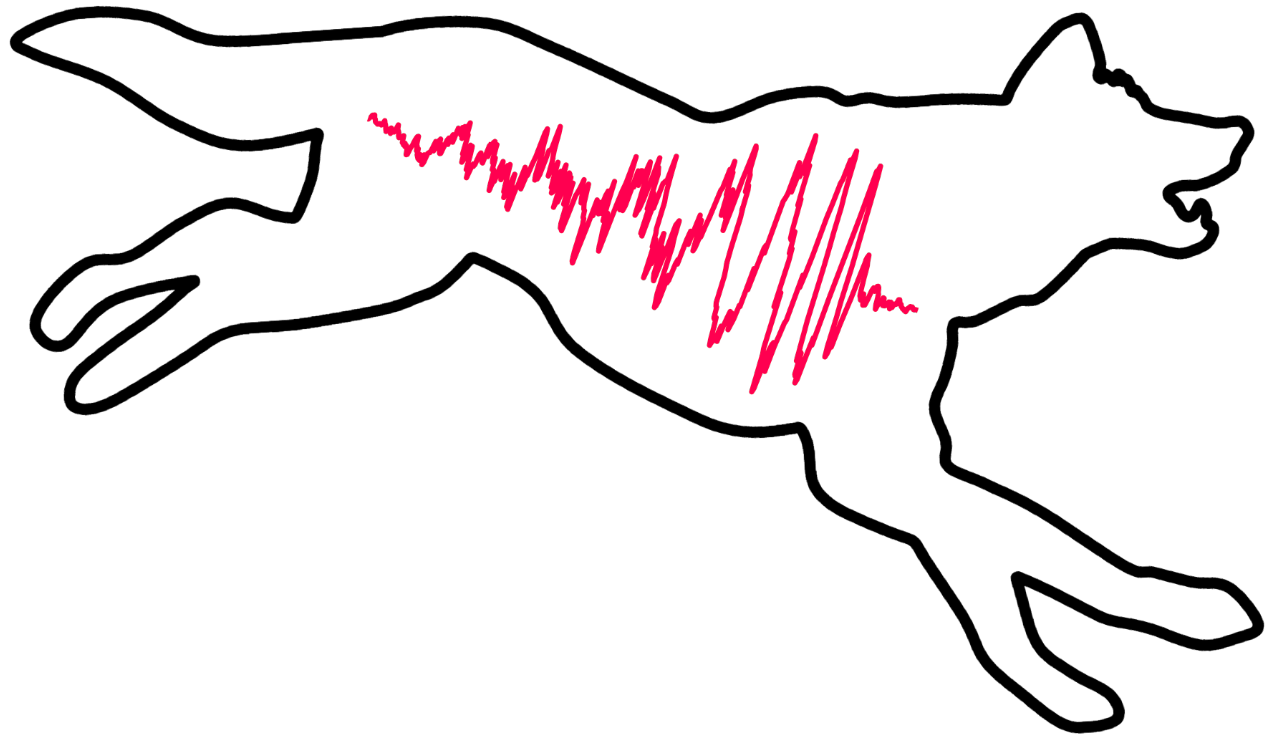
MAX PLANCK INSTITUTE
FOR INTELLIGENT SYSTEMS

Posterior Estimation with DINGO

An (incomplete) Summary



Annalena Kofler, 25.04.2024

AIslands 2024

The Dingo Pack


Max Dax


Stephen Green


Jonas Wildberger


Nihar Gupte


Michael Pürrer


Hector Estelles


Alex Roussopoulos


Samuel Clyne

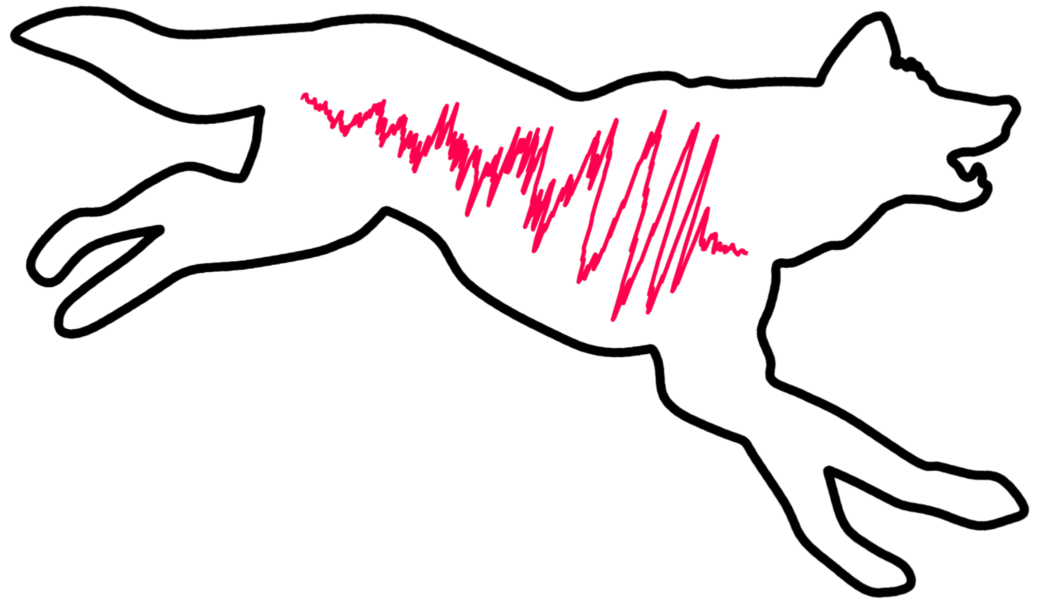

Annalena Kofler


Jonathan Gair


Jakob Macke


Bernhard Schölkopf

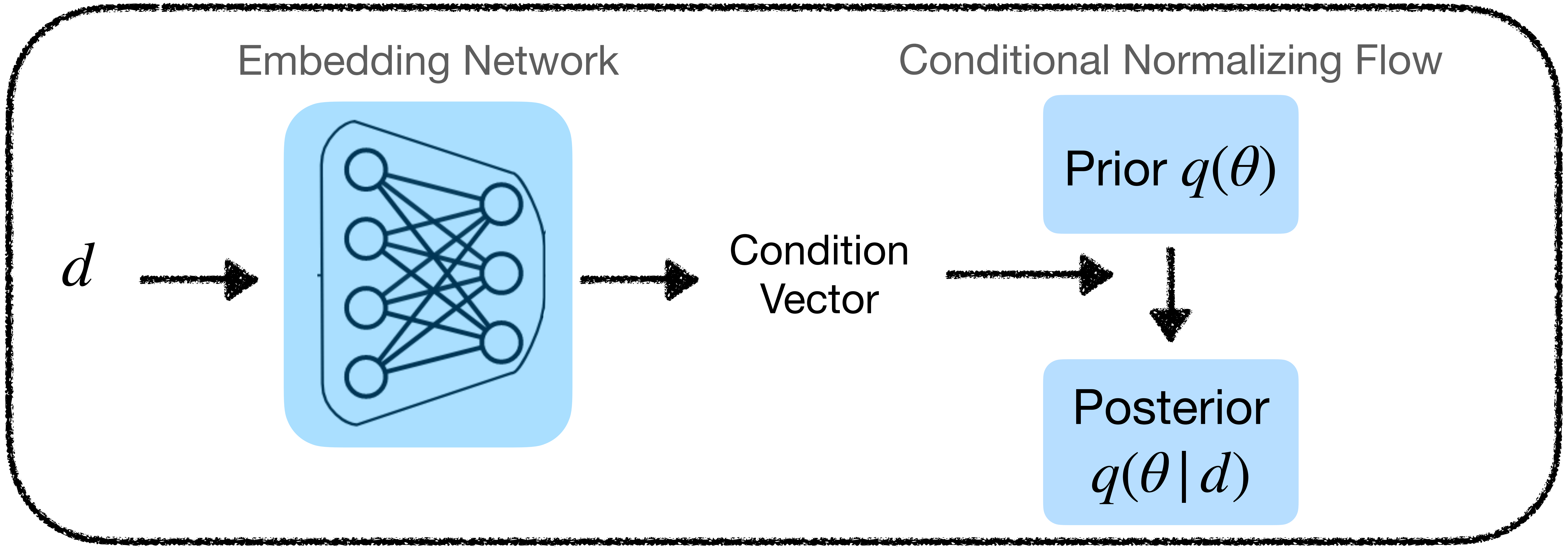

Alessandra Buonanno



What is DINGO?

- DINGO = **D**eep **IN**ference for **G**ravitational wave **O**bservations

DINGO Model



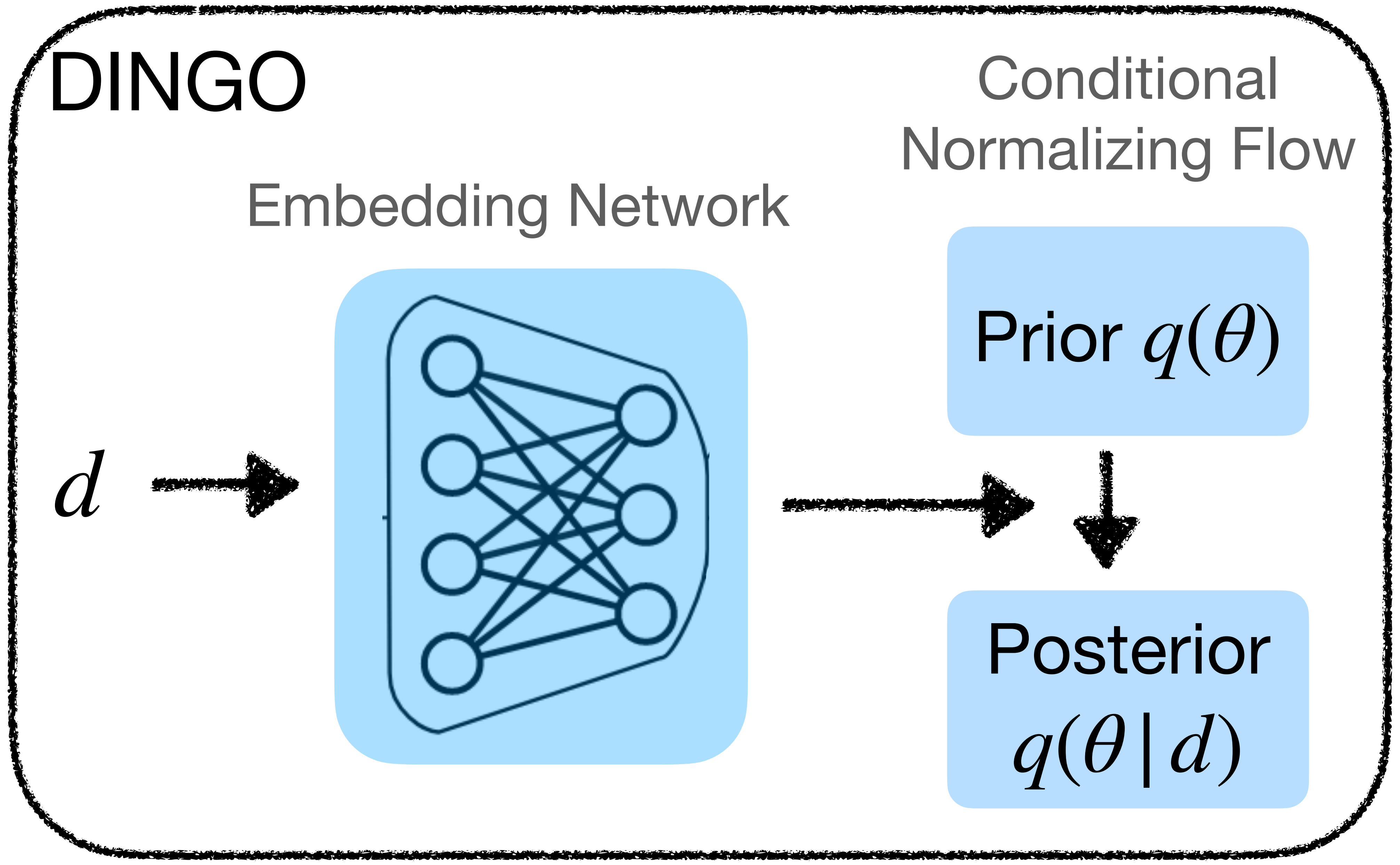
d: Data θ: Parameters

Dax+, PRL 2021

How to train DINGO?

- Sample from prior θ
- Simulate data: $d =$ signal + noise
- Evaluate DINGO Model
- Minimize negative log likelihood

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

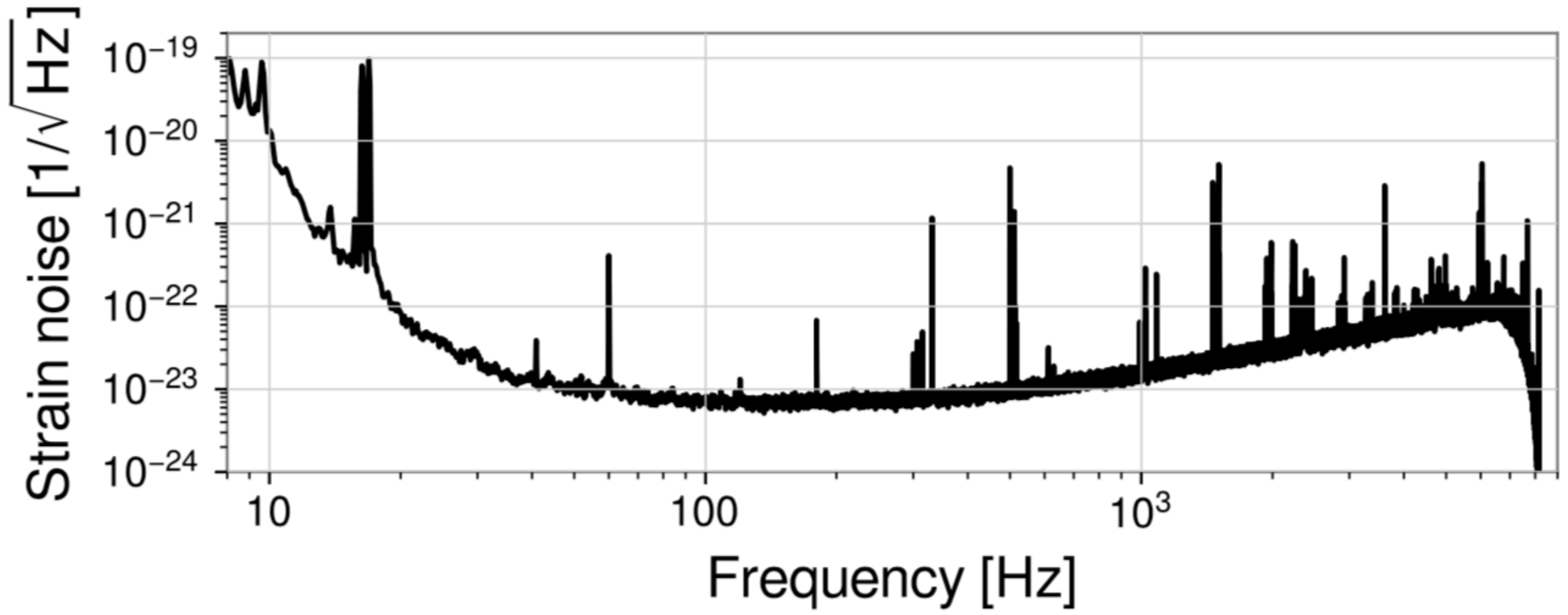


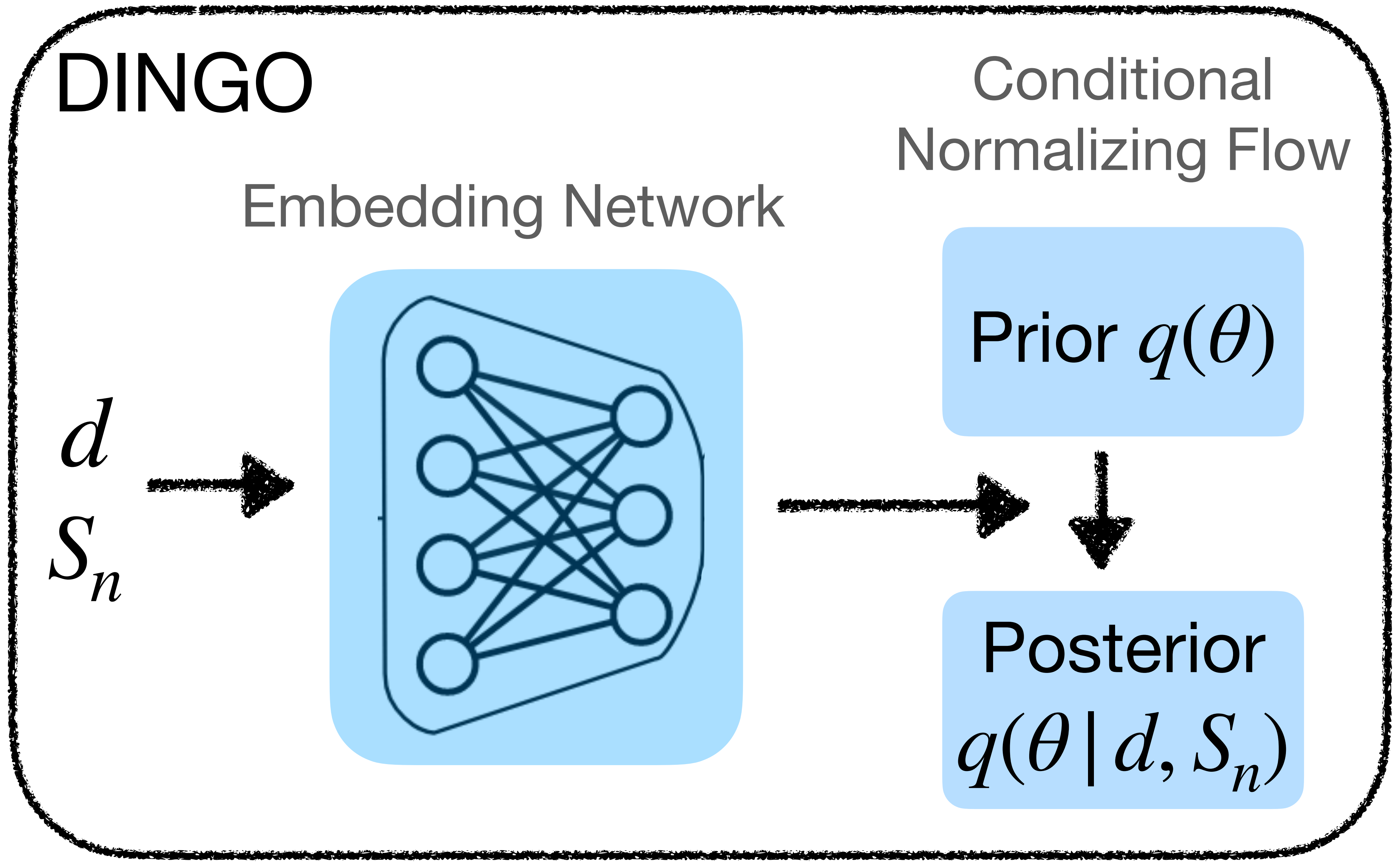
Dax+, PRL 2021

Changing PSDs

- Detector noise $S_n(f)$ varies from event to event
  → 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 $\quad d^{(i)} = h(\theta^{(i)}) + n^{(i)}$



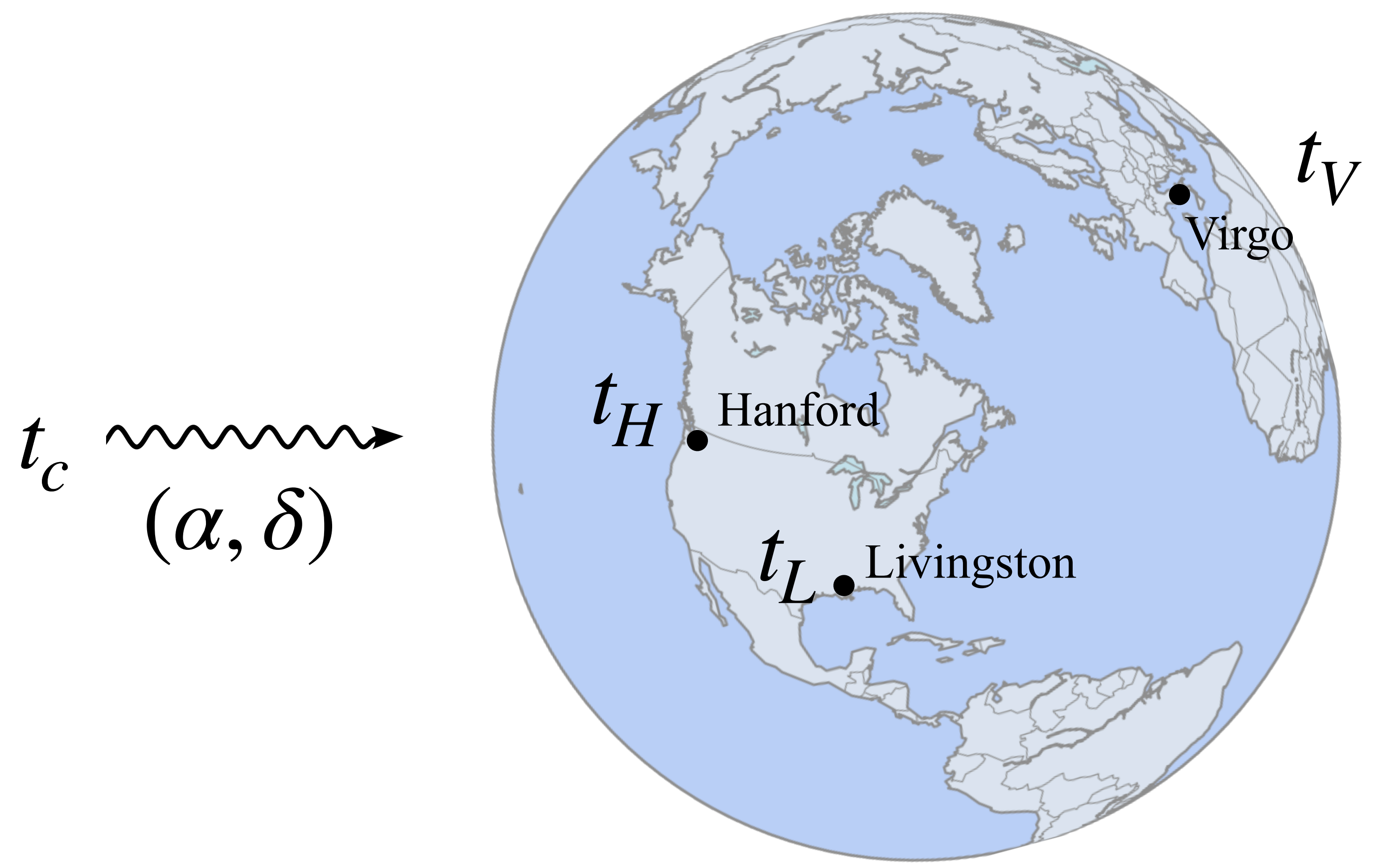


Dax+, PRL 2021

Simplifying Data using Symmetries

- Posterior should be …
  … equivariant wrt. overall coalescence time t_c

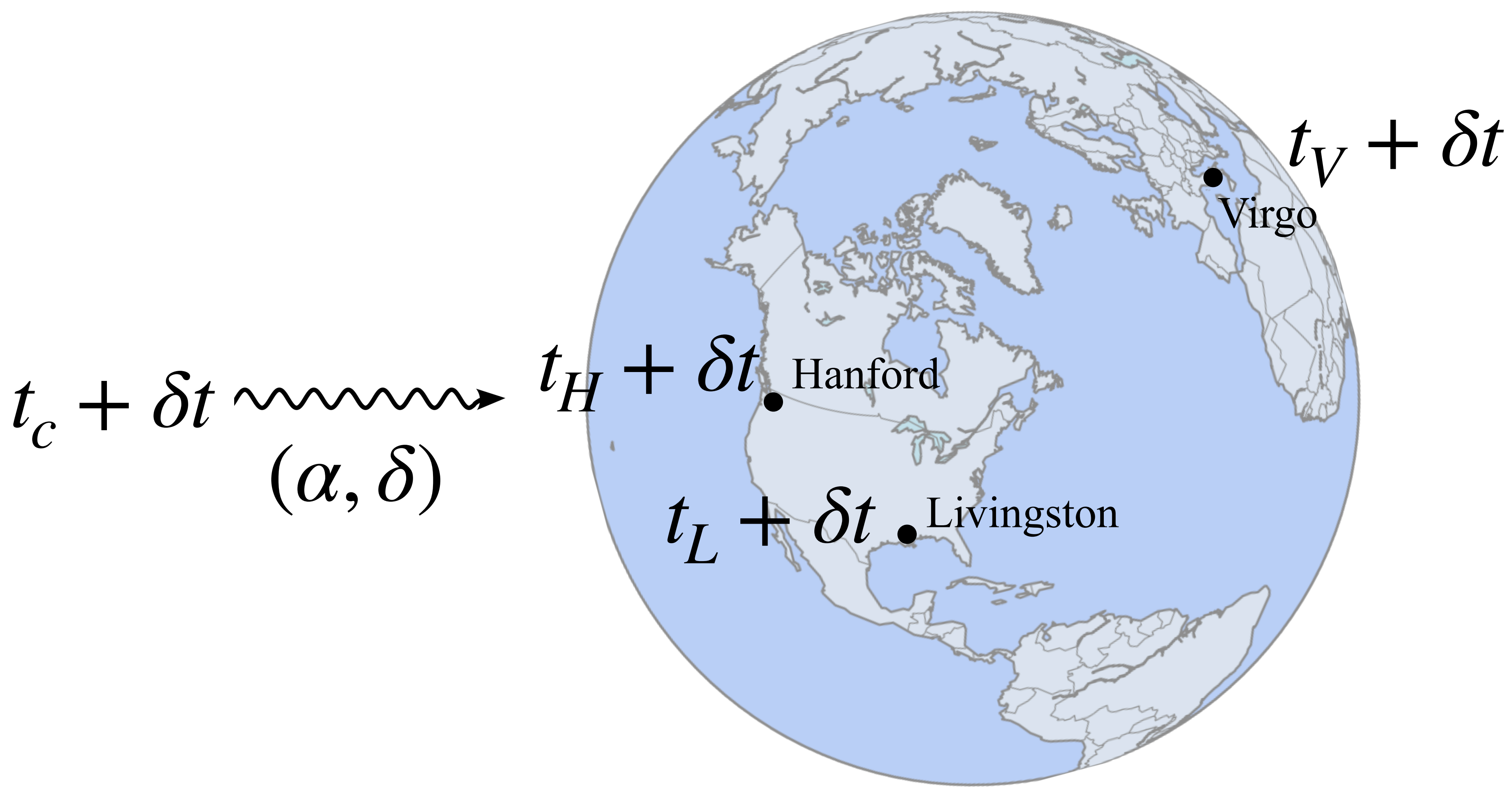
- Frequency-domain data
  → time shift = multiplication by $e^{2\pi i f \cdot \delta t}$



Simplifying Data using Symmetries

- Posterior should be …
  … equivariant wrt. overall coalescence time t_c

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



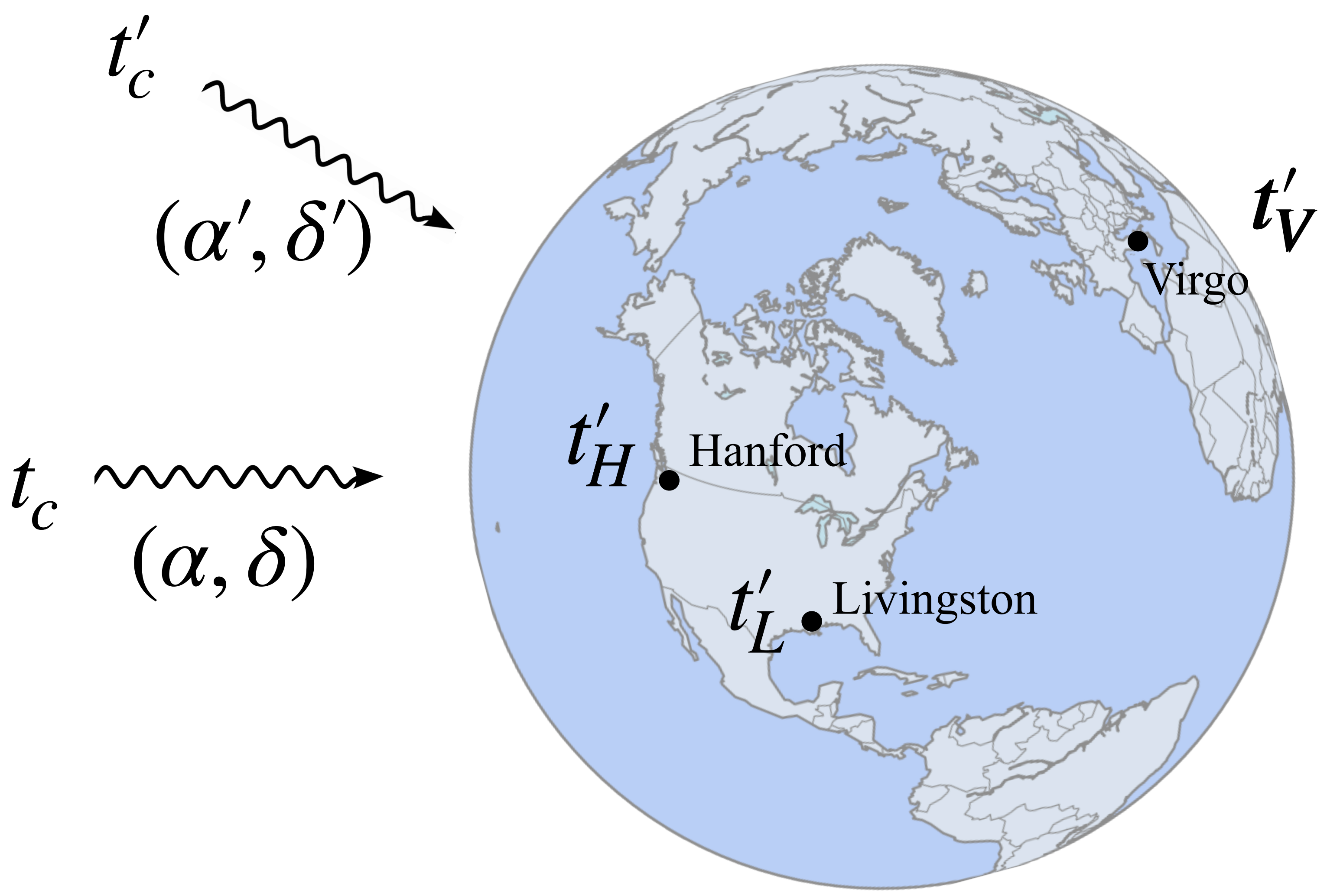
Dax+, ICLR 2022

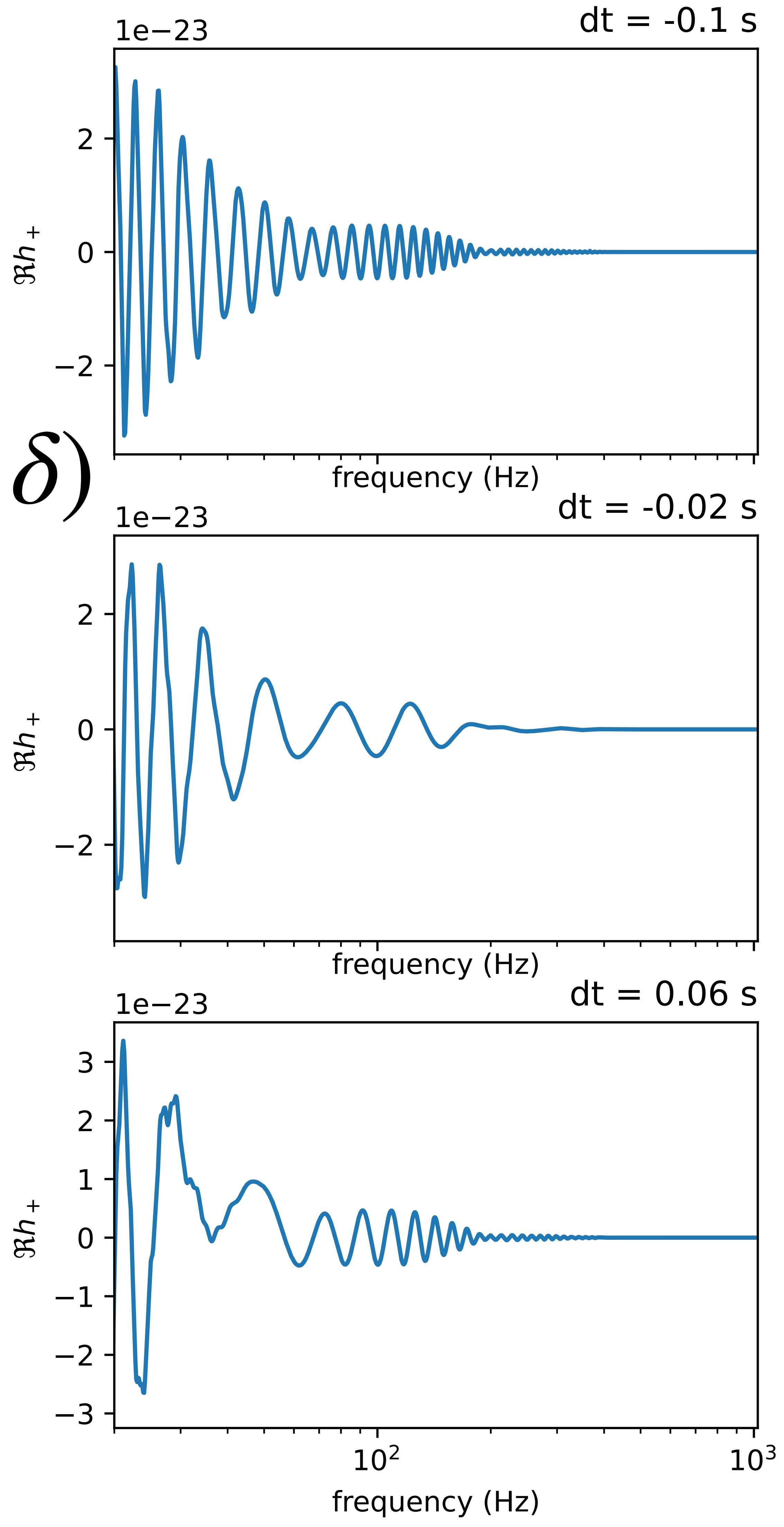
Simplifying Data using Symmetries

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

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

- Hard to learn just from data



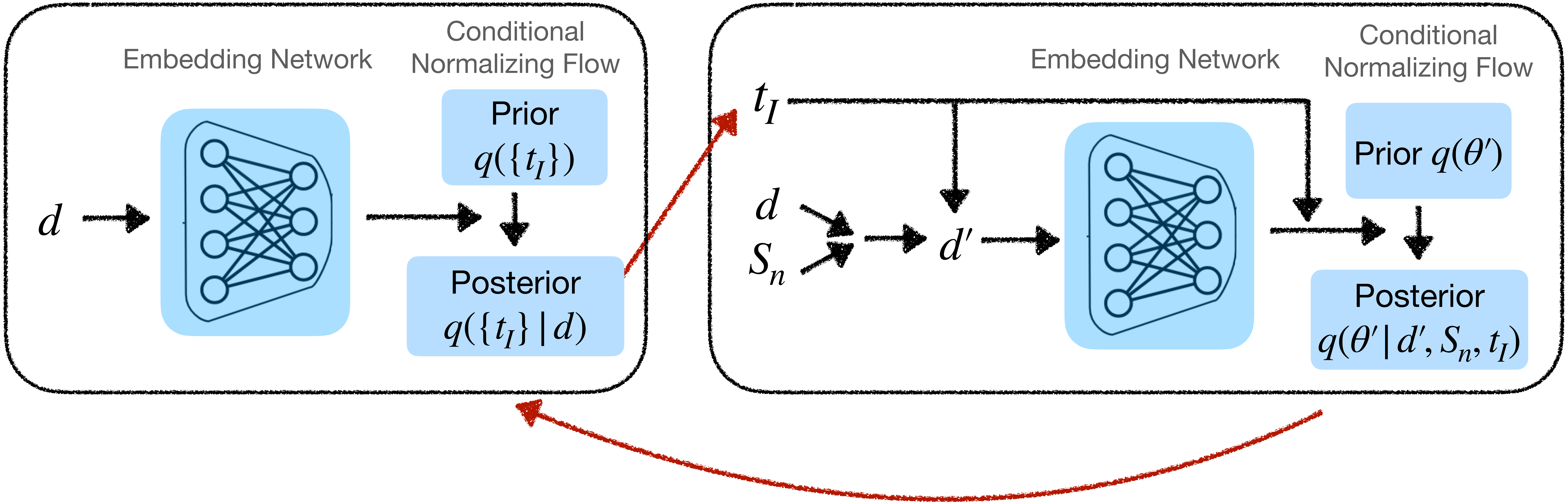


GNPE = Group-equivariant NPE

- Standardize data within band around $t_I = 0$
- Gibbs Sampling:
  - Start with blurred proxy $t_I + \epsilon_I$ with $\epsilon_I \sim \kappa(\epsilon_I)$
  - Learn approximate distribution $q_{\text{init}}(\{t_I\}_{I=H,L,V} \,|\, d)$ for time shift
  - Time-translate data $d' = d(t_I = 0)$
  - Train conditional density estimator $q(\theta' \,|\, d', t_I + \epsilon_I)$

GNPE = Group-equivariant NPE

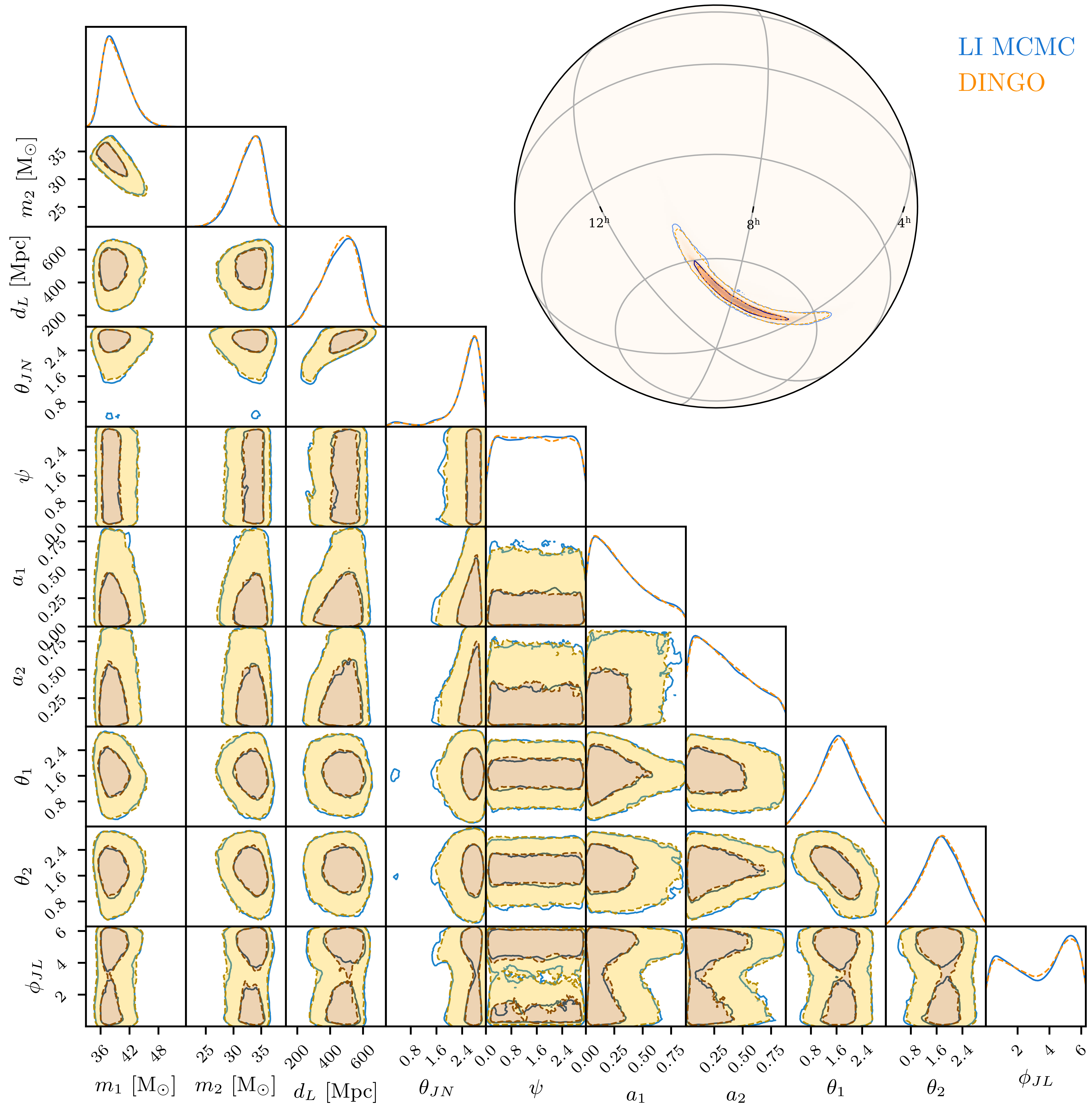
- Include condition on t_I in DINGO



Dax+, ICLR 2022

DINGO 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

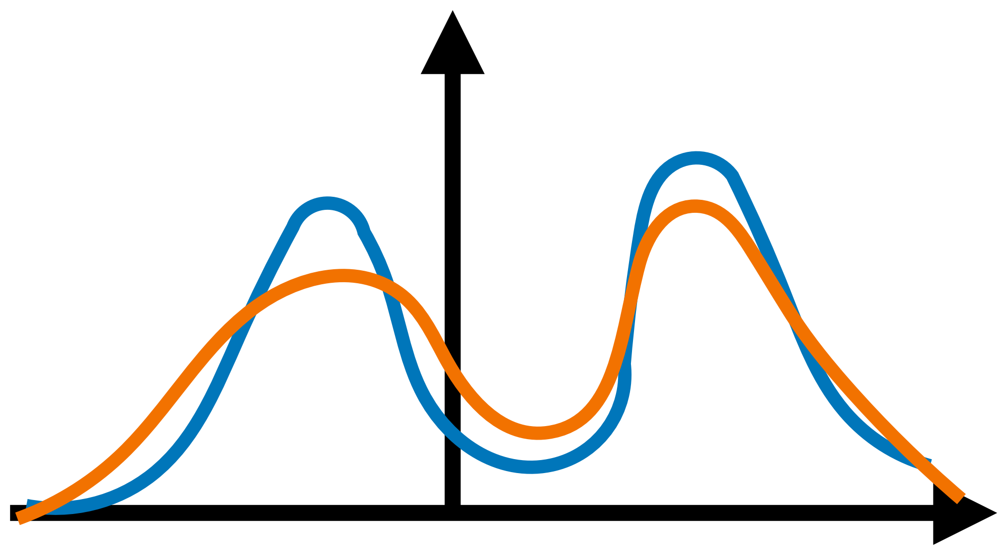
Verifying Results: DINGO-IS

- Neural Importance Sampling to compare learned NPE density and likelihood

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

Target = Likelihood x Prior

Proposal (NPE)



- Effective Sample Size:

$$n_{\text{eff}} = \frac{\left(\sum_i w_i\right)^2}{\sum_i w_i^2}$$

- Evidence: $$p(d) \approx \frac{1}{N}\sum_{i=1}^{N} w_i$$

Dax+, PRL 2023

Verifying Results: DINGO-IS

- Failure cases flagged via low sample efficiency ϵ
- Out-of-distribution data: = inconsistent with noise or signal model

→ Identification of events with known issues with data quality or modeling

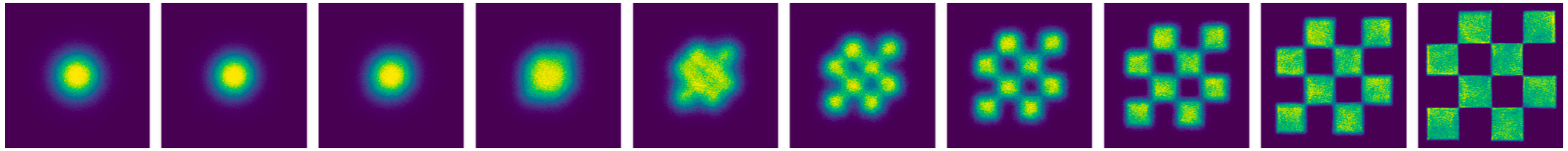
Event	$\log p(d)$	ϵ	Event	$\log p(d)$	ϵ	Event	$\log p(d)$	ϵ
GW190408	-16178.332 ± 0.012	6.9%	GW190727	-15992.017 ± 0.009	10.3%	GW191230	-15913.798 ± 0.009	12.2%
_181802	-16178.172 ± 0.010	9.3%	_060333	-15992.428 ± 0.005	30.8%	_180458	-15913.918 ± 0.010	8.8%
GW190413	-15571.413 ± 0.006	22.5%	GW190731	-16376.777 ± 0.005	32.6%	GW200128	-16305.128 ± 0.013	6.1%
_052954	-15571.391 ± 0.005	26.3%	_140936	-16376.763 ± 0.005	31.0%	_022011	-16304.510 ± 0.007	18.3%
GW190413	-16399.331 ± 0.009	12.4%	GW190803	-16132.409 ± 0.006	21.4%	‡GW200129	-16226.851 ± 0.109	0.1%
_134308	-16399.139 ± 0.014	4.7%	_022701	-16132.408 ± 0.005	27.8%	_065458	-16231.203 ± 0.051	0.4%
GW190421	-15983.248 ± 0.008	15.3%	GW190805	-16073.261 ± 0.006	20.0%	GW200208	-16136.381 ± 0.007	16.6%
_213856	-15983.131 ± 0.010	9.4%	_211137	-16073.656 ± 0.007	16.6%	_130117	-16136.531 ± 0.009	11.2%
GW190503	-16582.865 ± 0.022	2.0%	GW190828	-16137.220 ± 0.009	12.2%	GW200208	-16775.200 ± 0.011	7.4%
_185404	-16583.352 ± 0.027	1.4%	_063405	-16136.799 ± 0.010	9.1%	_222617	-16774.582 ± 0.021	2.2%
GW190513	-15946.462 ± 0.043	0.6%	GW190909	-16061.634 ± 0.011	7.4%	GW200209	-16383.847 ± 0.009	12.5%
_205428	-15946.581 ± 0.017	3.4%	_114149	-16061.275 ± 0.016	3.8%	_085452	-16384.157 ± 0.025	1.6%
GW190514	-16556.466 ± 0.009	11.6%	GW190915	-16083.960 ± 0.015	20.8%	GW200216	-16215.703 ± 0.017	3.4%
_065416	-16556.314 ± 0.017	3.5%	_235702	-16083.937 ± 0.027	4.8%	_220804	-16215.540 ± 0.018	3.1%
GW190517	-16271.048 ± 0.027	1.3%	GW190926	-16015.813 ± 0.019	2.8%	GW200219	-16133.457 ± 0.011	9.6%
_055101	-16272.428 ± 0.034	0.9%	_050336	-16015.861 ± 0.009	12.1%	_094415	-16133.157 ± 0.017	4.0%
GW190519	-15991.171 ± 0.008	15.2%	GW190929	-16146.666 ± 0.018	3.2%	GW200220	-16303.782 ± 0.007	17.3%
_153544	-15991.287 ± 0.068	0.2%	_012149	-16146.591 ± 0.021	2.4%	_061928	-16303.087 ± 0.026	1.5%
GW190521	-16008.876 ± 0.008	13.4%	GW191109	-17925.064 ± 0.025	1.7%	GW200220	-16136.600 ± 0.008	13.2%
_074359	-16008.037 ± 0.015	4.2%	_010717	-17922.762 ± 0.041	0.6%	_124850	-16136.519 ± 0.037	0.7%
GW190527	-16119.012 ± 0.008	13.8%	GW191127	-16759.328 ± 0.019	2.7%	GW200224	-16138.613 ± 0.006	22.5%
_092055	-16118.781 ± 0.013	6.1%	_050227	-16758.102 ± 0.029	1.2%	_222234	-16139.101 ± 0.006	21.4%
GW190602	-16036.993 ± 0.006	25.0%	‡GW191204	-15984.455 ± 0.015	4.2%	‡GW200308	-16173.938 ± 0.013	6.0%
_175927	-16037.529 ± 0.006	23.5%	_110529	-15983.618 ± 0.063	0.3%	_173609	-16173.692 ± 0.025	1.7%
GW190701	-16521.381 ± 0.040	0.6%	GW191215	-16001.286 ± 0.013	5.8%	GW200311	-16117.505 ± 0.011	7.4%
_203306	-16521.609 ± 0.010	10.1%	_223052	-16000.846 ± 0.052	0.4%	_115853	-16117.583 ± 0.009	11.9%
GW190719	-15850.492 ± 0.008	13.4%	GW191222	-15871.521 ± 0.007	16.5%	‡GW200322	-16313.568 ± 0.307	0.0%
_215514	-15850.339 ± 0.011	8.0%	_033537	-15871.450 ± 0.005	25.8%	_091133	-16313.110 ± 0.105	0.1%

Table II. 42 BBH events from GWTC-3 analyzed with DINGO-IS. We report the log evidence $\log p(d)$ and the sample efficiency ϵ for the two waveform models IMRPhenomXPHM (upper rows) and SEOBNRv4PHM (lower rows). Highlighting colors indicate the sample efficiency (green: high; yellow: medium; orange/red: low); DINGO-IS results can be trusted for medium and high ϵ (see Supplemental Material). Events in gray suffer from data quality issues [1, 21]. ‡See remarks on these events in text.

How to improve DINGO?

- Larger networks?
  $\rightarrow$ (discrete) normalizing flows do not scale well beyond $\sim 10^8$ trainable parameters and ~ 15 inference parameters

- Alternative: Flow Matching Lipman et al., ICLR 2023



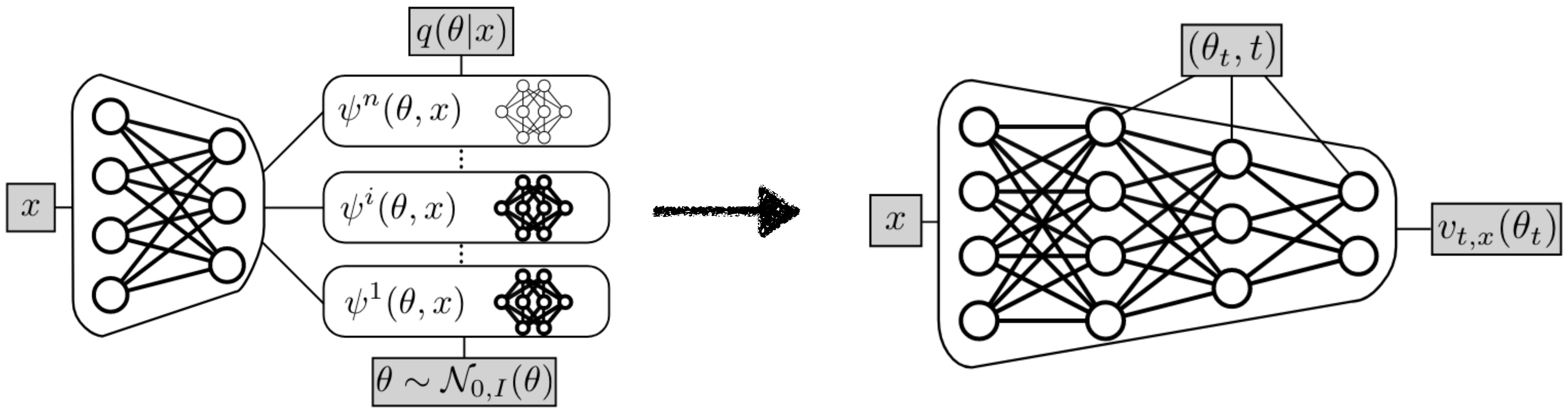
Flow Matching w/ OT

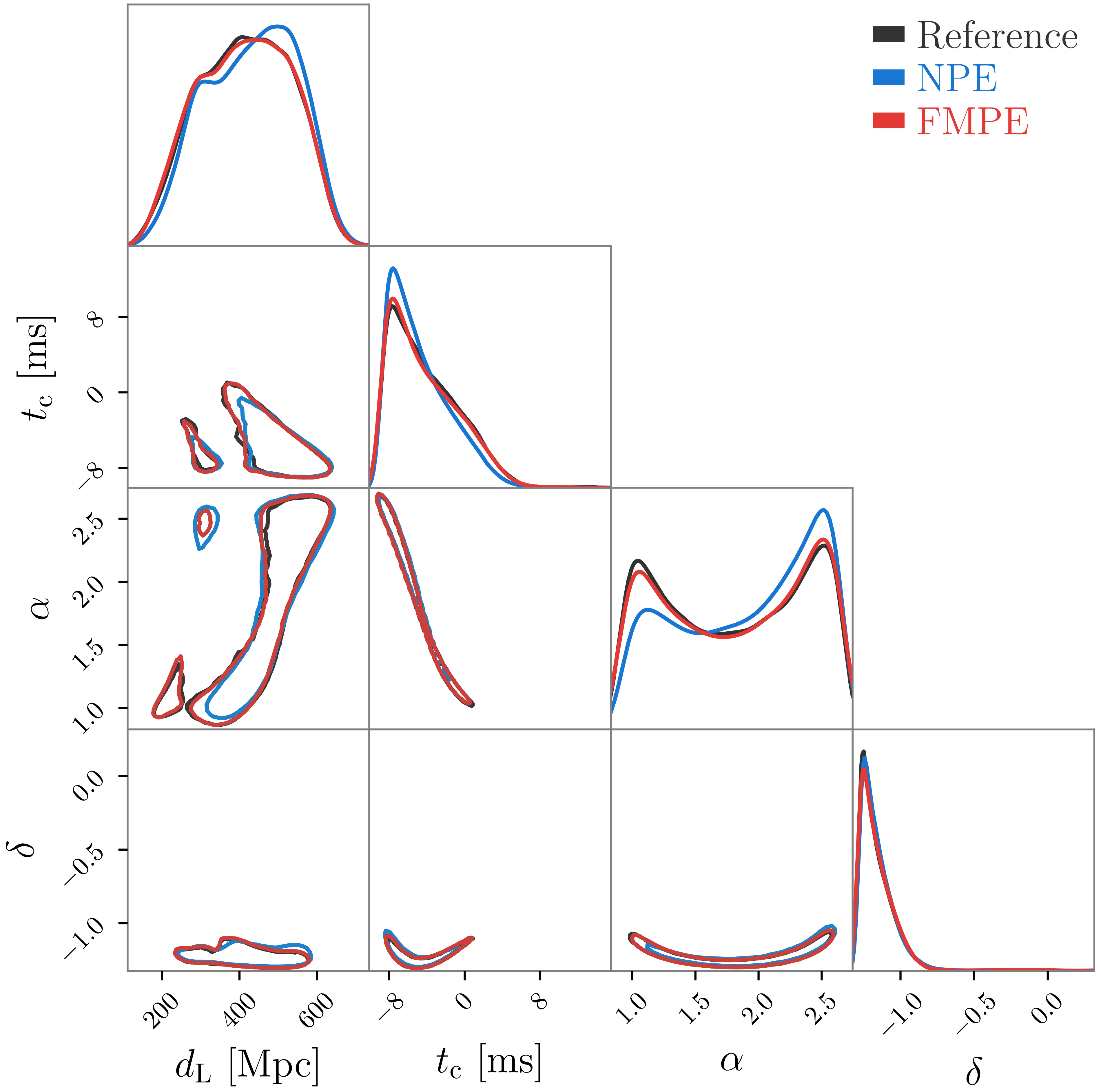
$t = 0$ $\longrightarrow t$ $t = 1$

- Network learns vector field $$\frac{\mathrm{d}\theta_t}{\mathrm{d}t} = v_{t,d}(\theta_t)$$

Flow Matching Posterior Estimation

- Estimate posterior with flow matching





- Outperforms NPE with discrete flows
  - Faster training
  - Better scaling to larger networks

Dax+, NeurIPS 2023

Where/How is DINGO used?

- Analyze large collection of observations
  - Learn population posterior directly   Leyde et al., 2023
  - Evidence for eccentricity → Nihar's talk on Monday   Gupte et al., 2024
- DINGO approved in internal LIGO review 🥳
  → Evaluation of O4 events!

Max Dax

Stephen Green

Jonas Wildberger

Nihar Gupte

Hector Estelles

Wildberger+, PRD 2023

Take-Aways for DINGO

- Accurate inference for BBHs in seconds - minutes
  → Rapid inference of large number of events
- Validation of results with Importance Sampling
- Ready to be used:
  Code @ https://github.com/dingo-gw/dingo, Documentation, Tutorials

Outlook

- Extension to BNS → Max talk on Friday
- New architectures to deliver ever-improving performance

The Dingo Pack


Max Dax


Stephen Green


Jonas Wildberger


Nihar Gupte


Michael Pürrer


Hector Estelles


Alex Roussopoulos


Samuel Clyne

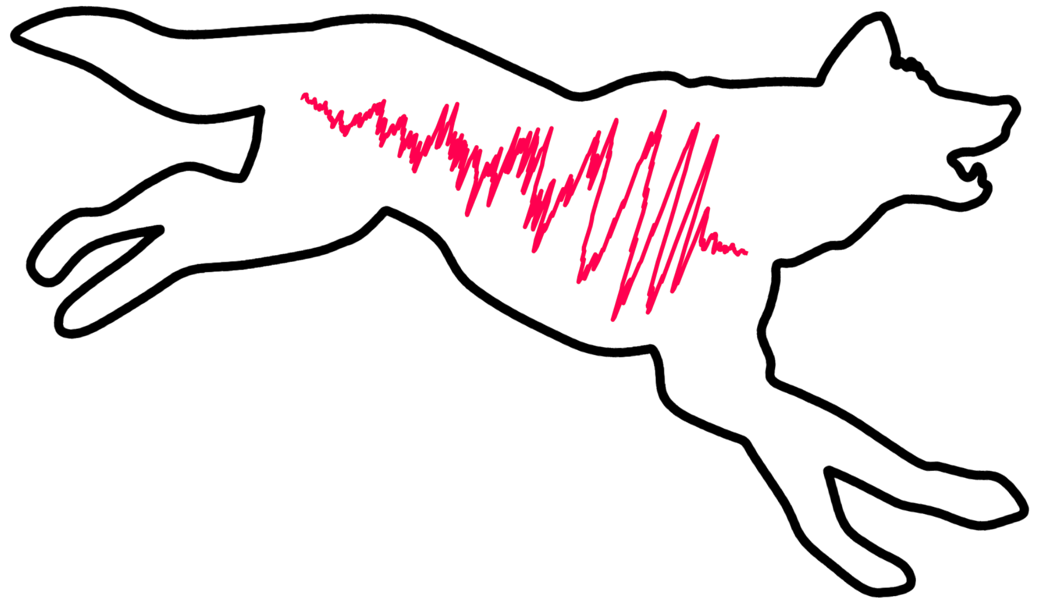

Annalena Kofler


Jonathan Gair


Jakob Macke


Bernhard Schölkopf


Alessandra Buonanno

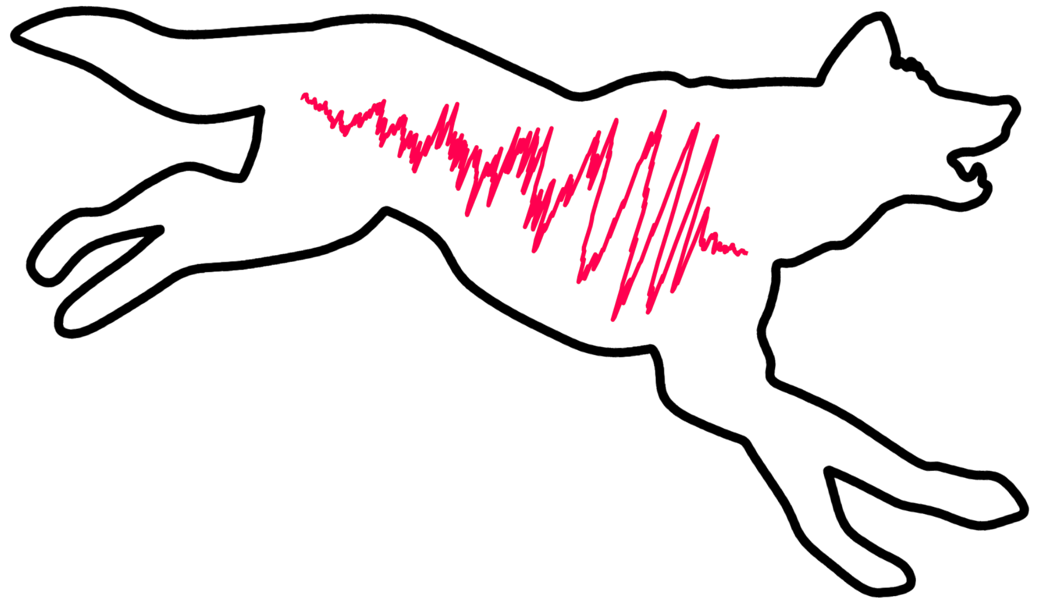


Discussion Questions

- Do you know of other scenarios where DINGO can be applied?
- Do you want to use DINGO or know people that are looking into using DINGO?

Thank you!

Do you have any questions?



References

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