FROM ALGORITHMS TO ASTROPHYSICS: MACHINE LEARNING IN GRAVITATIONAL WAVE ASTRONOMY

NIKOLAOS STERGIOULAS

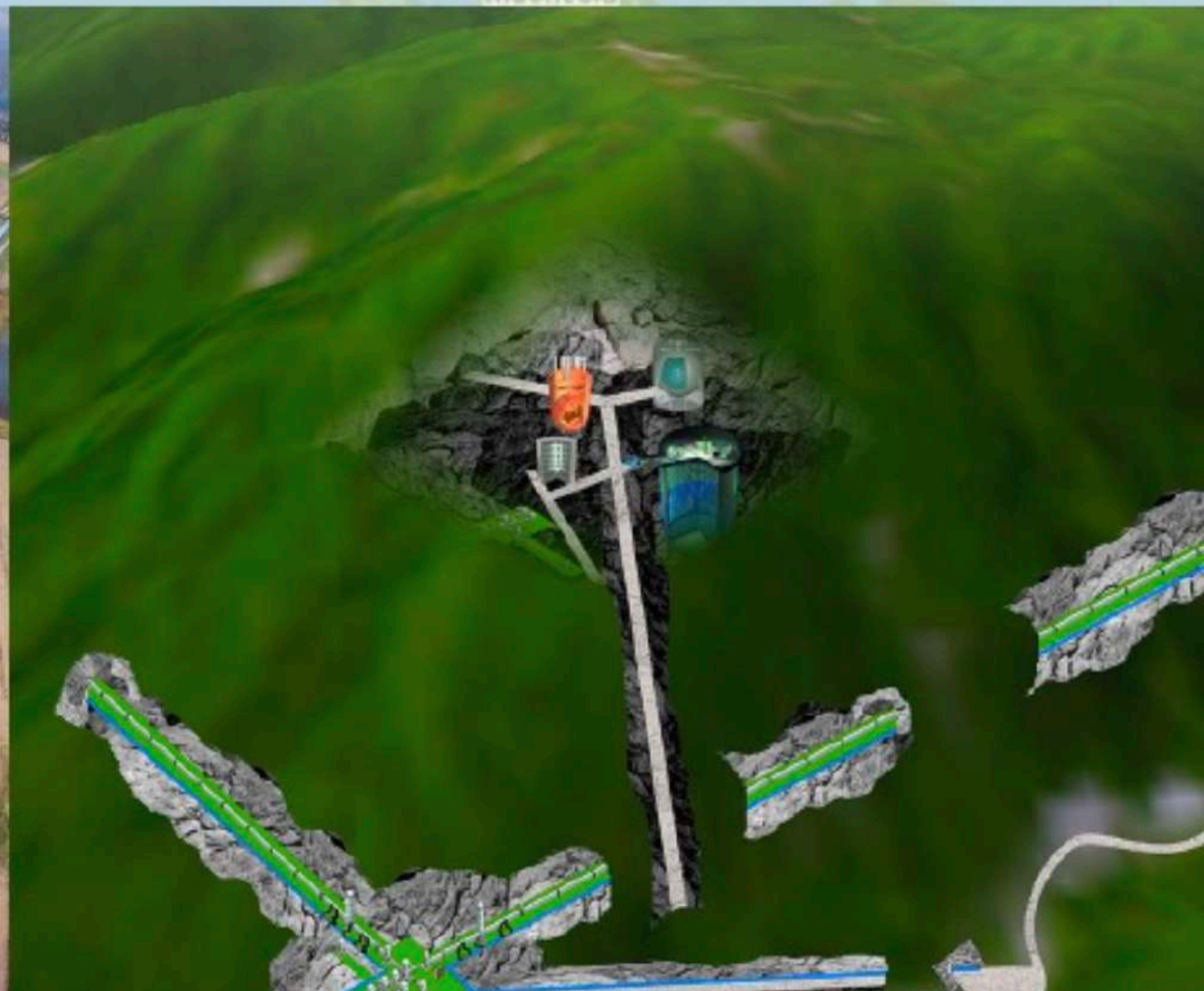
DEPARTMENT OF PHYSICS

ARISTOTLE UNIVERSITY OF THESSALONIKI

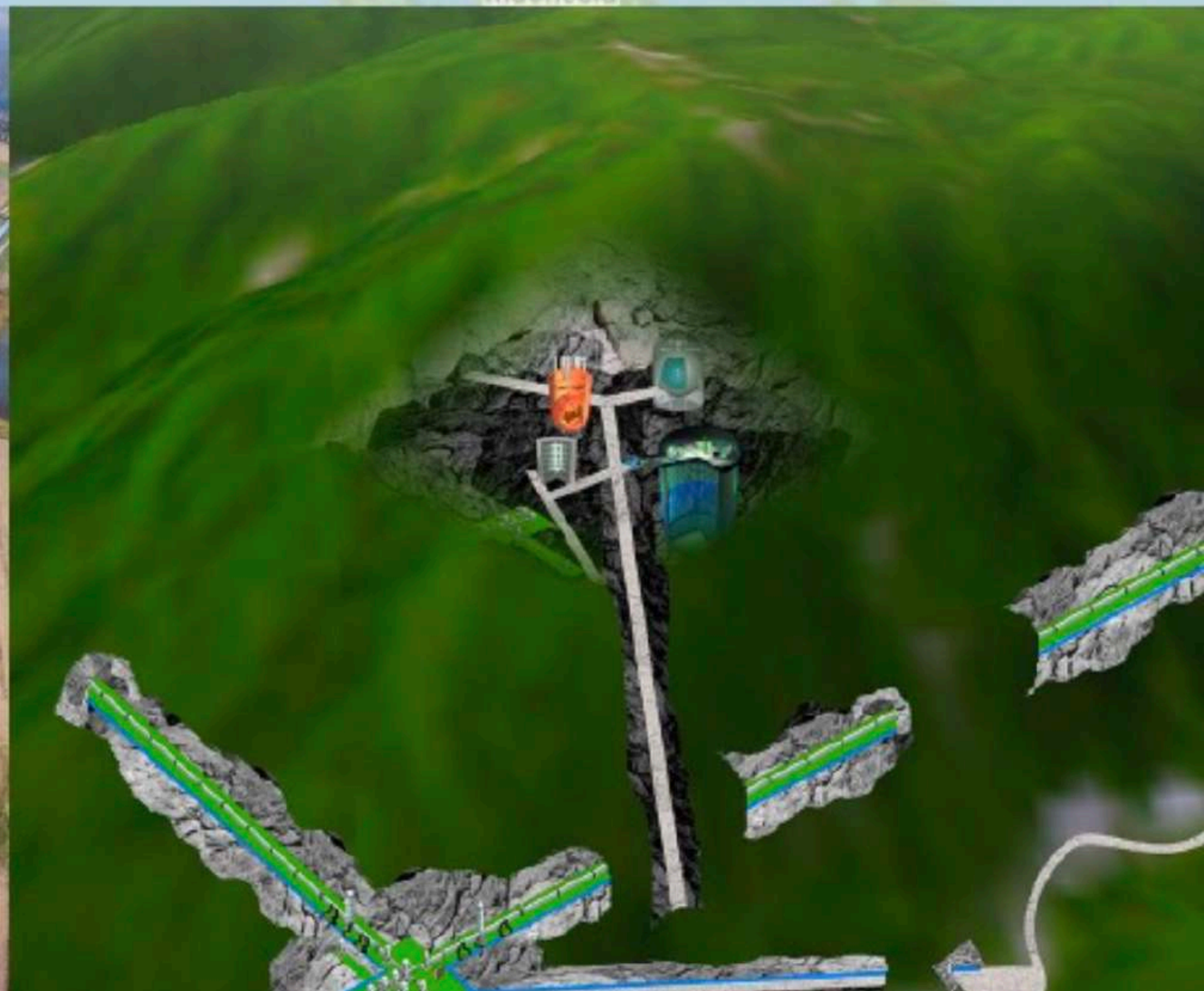
21st Conference of HSRGC, Corfu, September 4, 2025

PART A. BINARY BLACK HOLE MERGERS

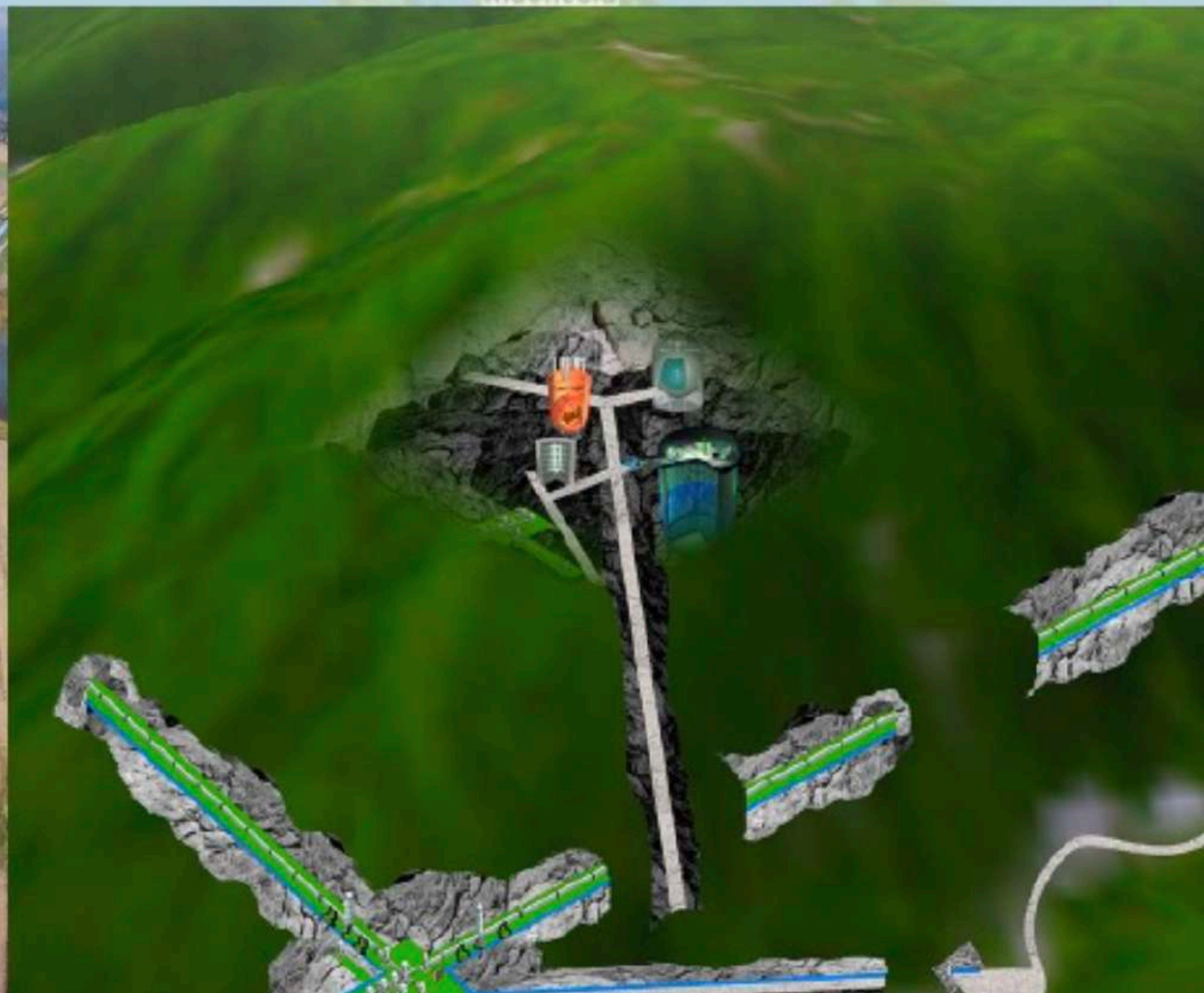
INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



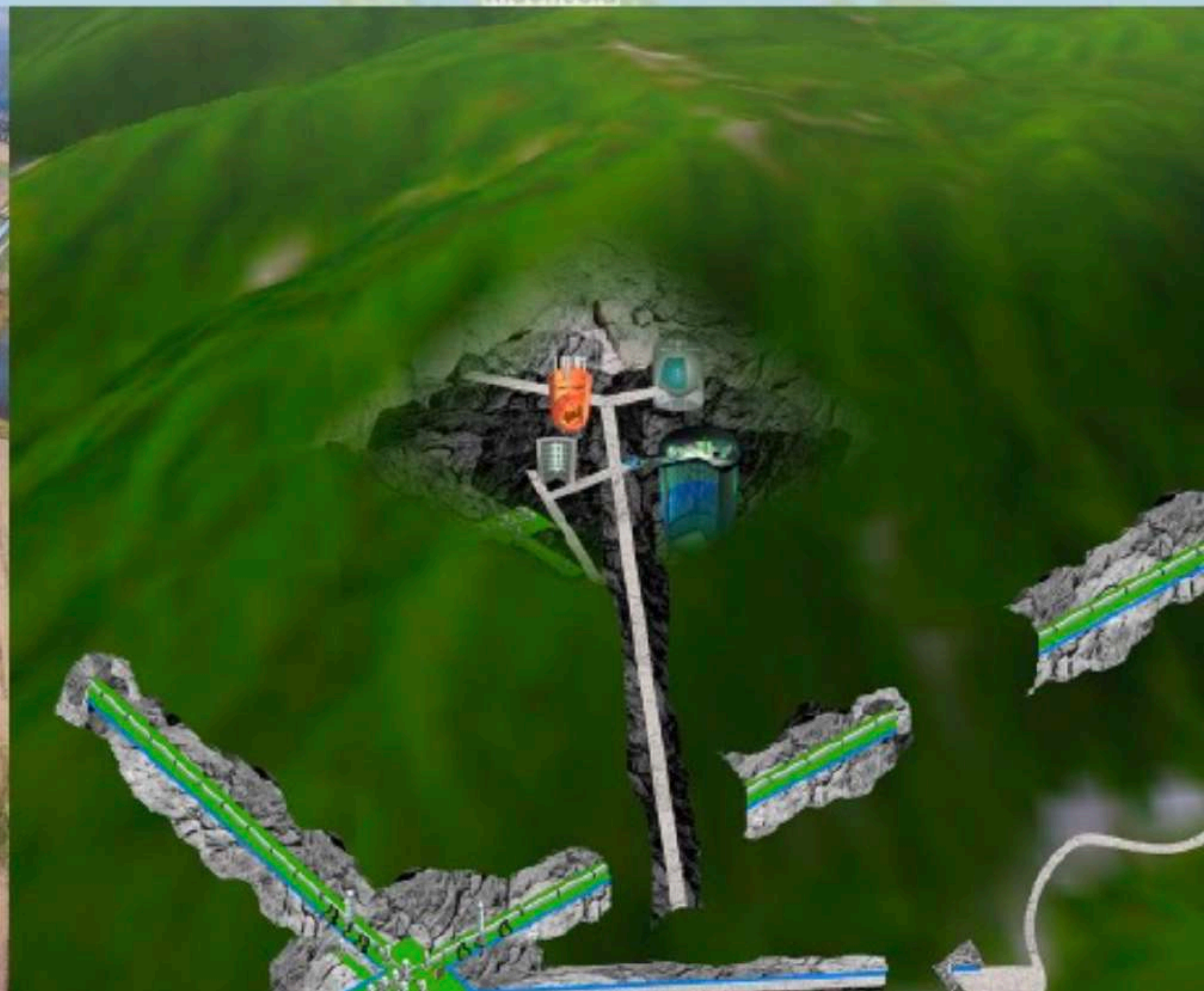
INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



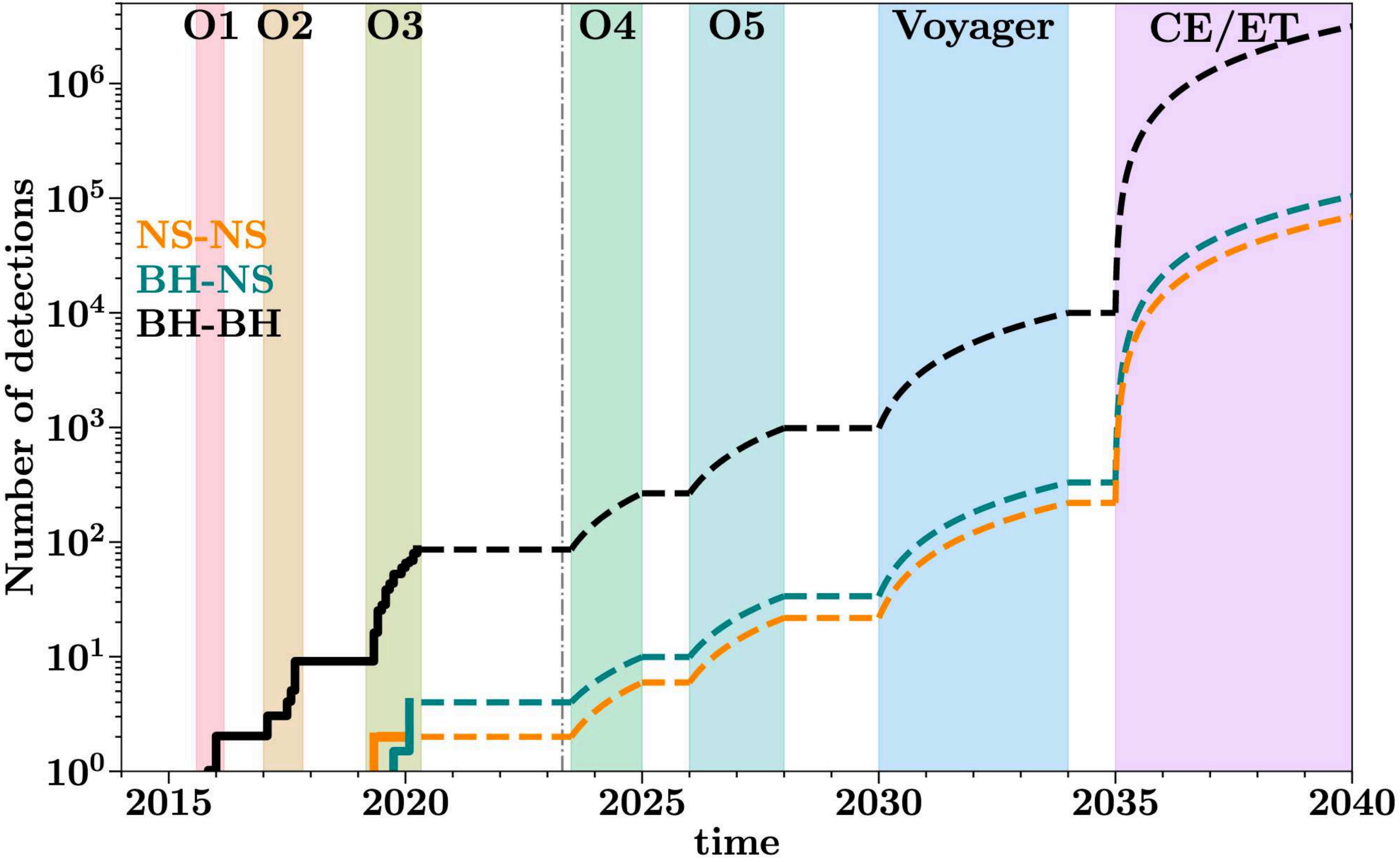
INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)

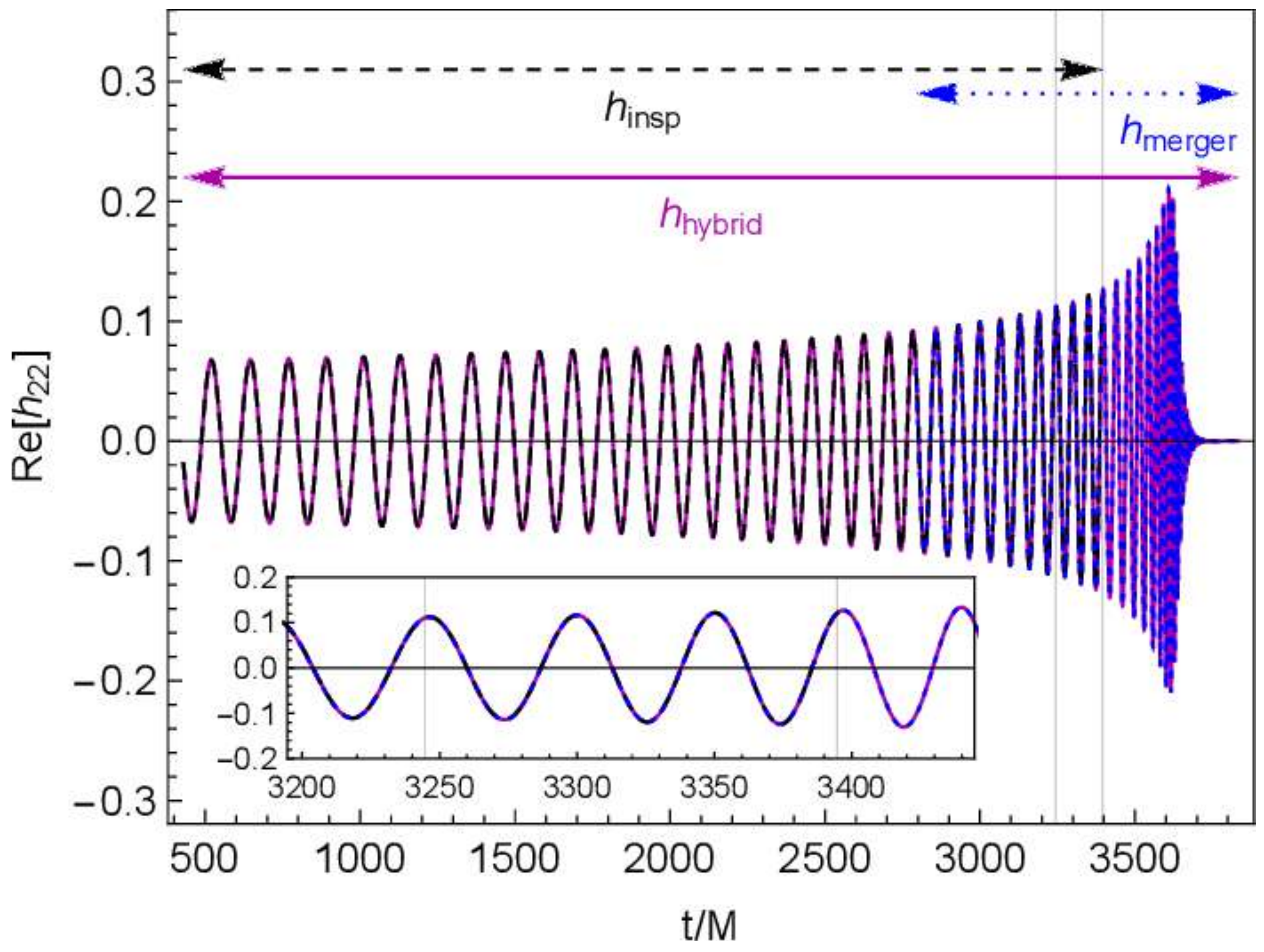


EXPECTED NUMBER OF DETECTIONS

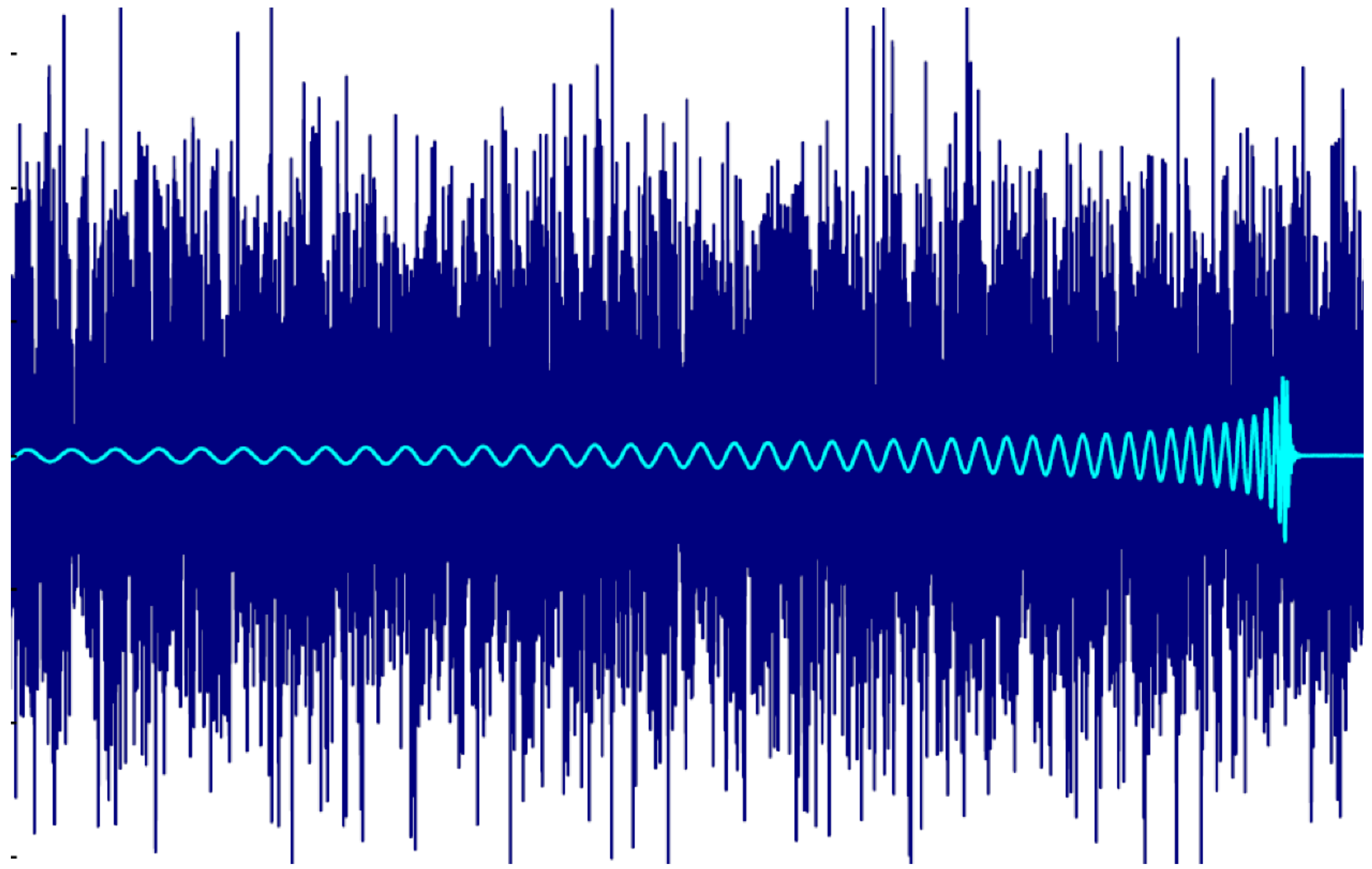


MACHINE LEARNING IN GRAVITATIONAL WAVE ASTRONOMY

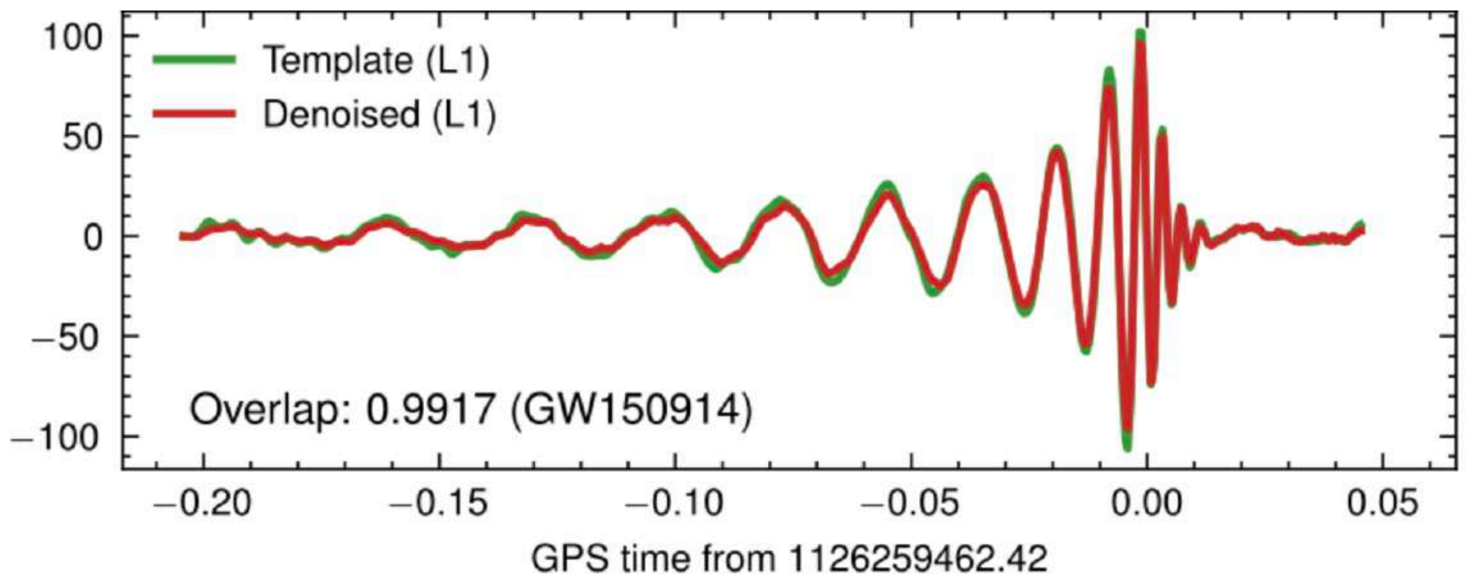
Waveform Modeling



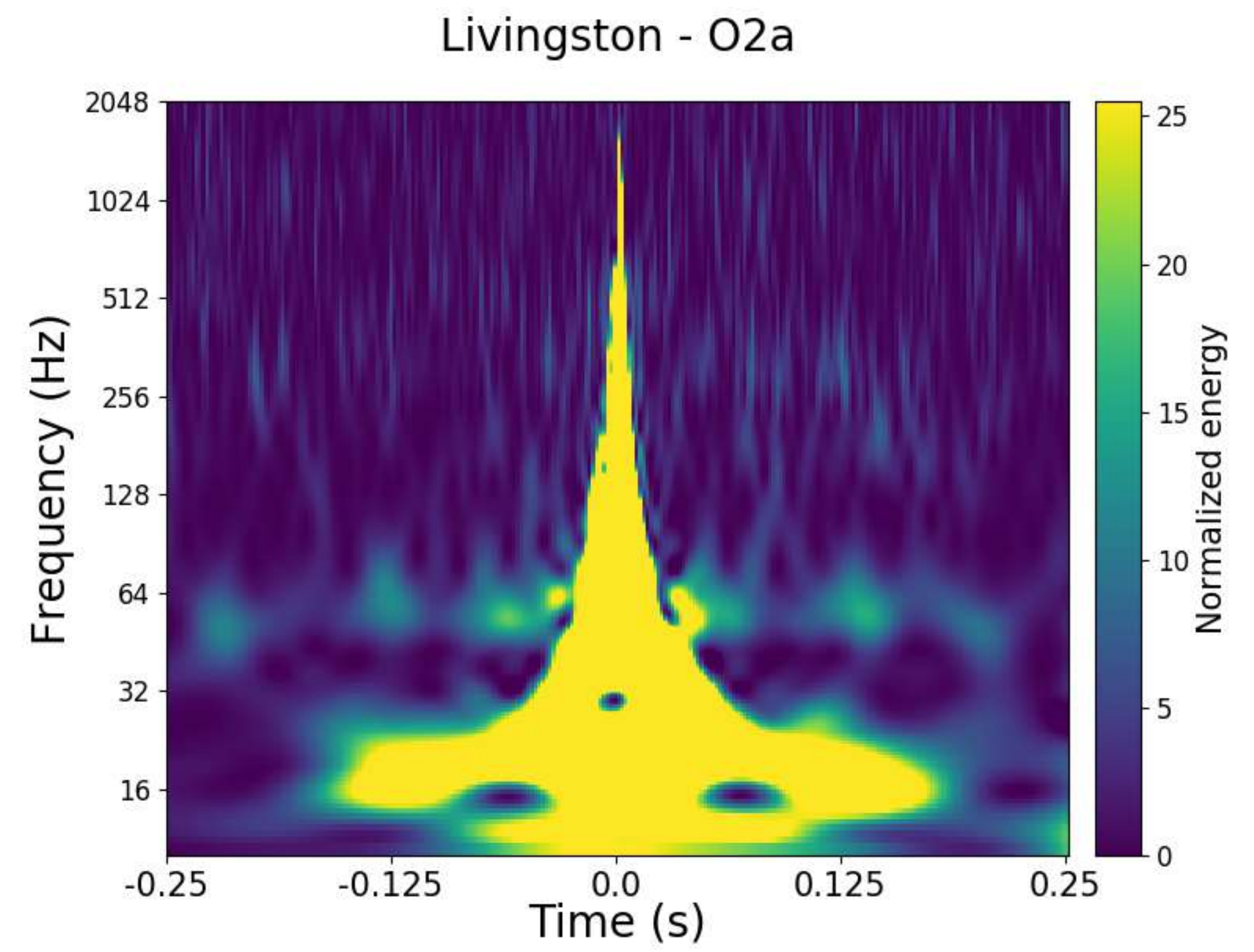
Detection



Denoising

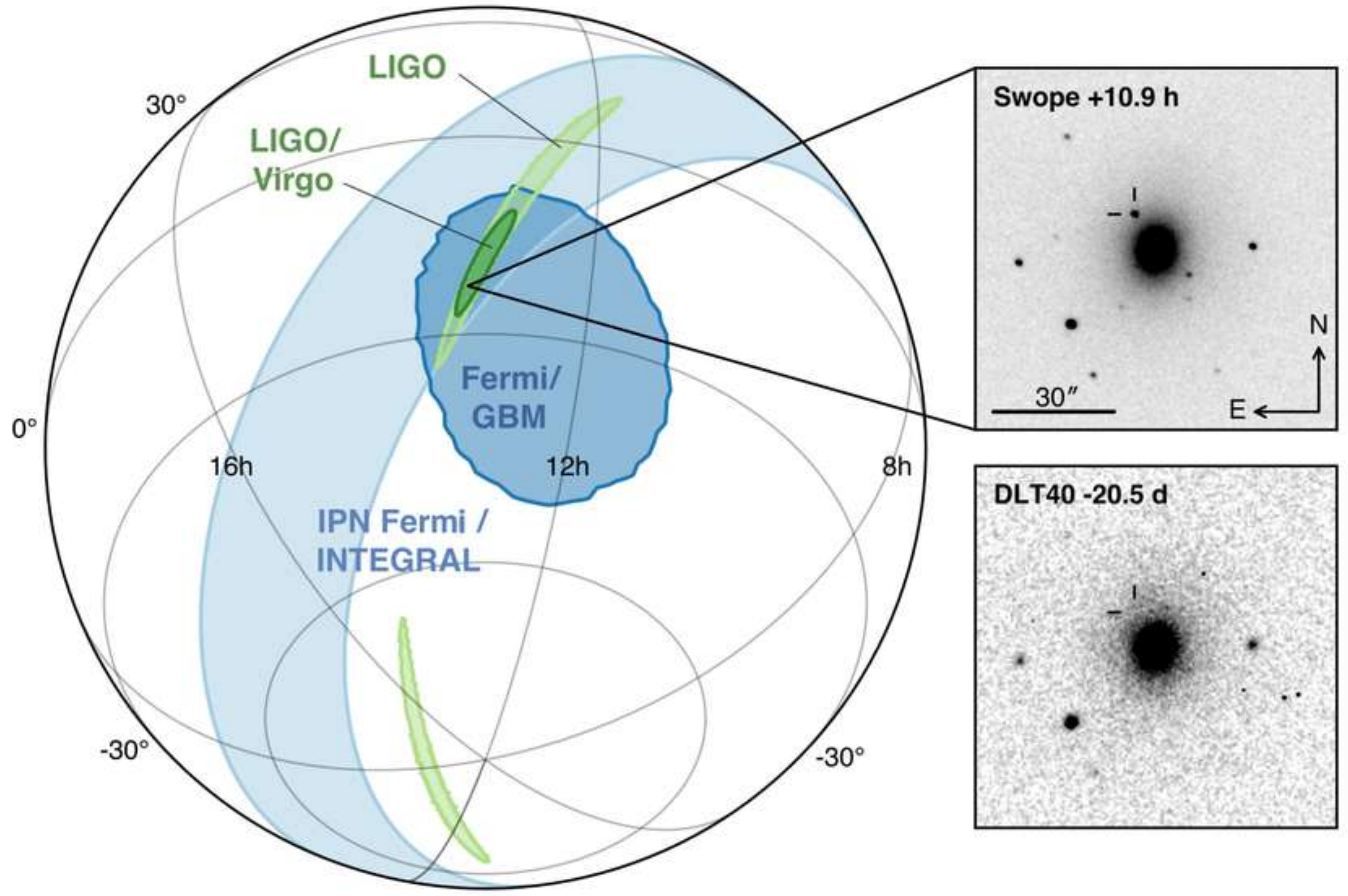


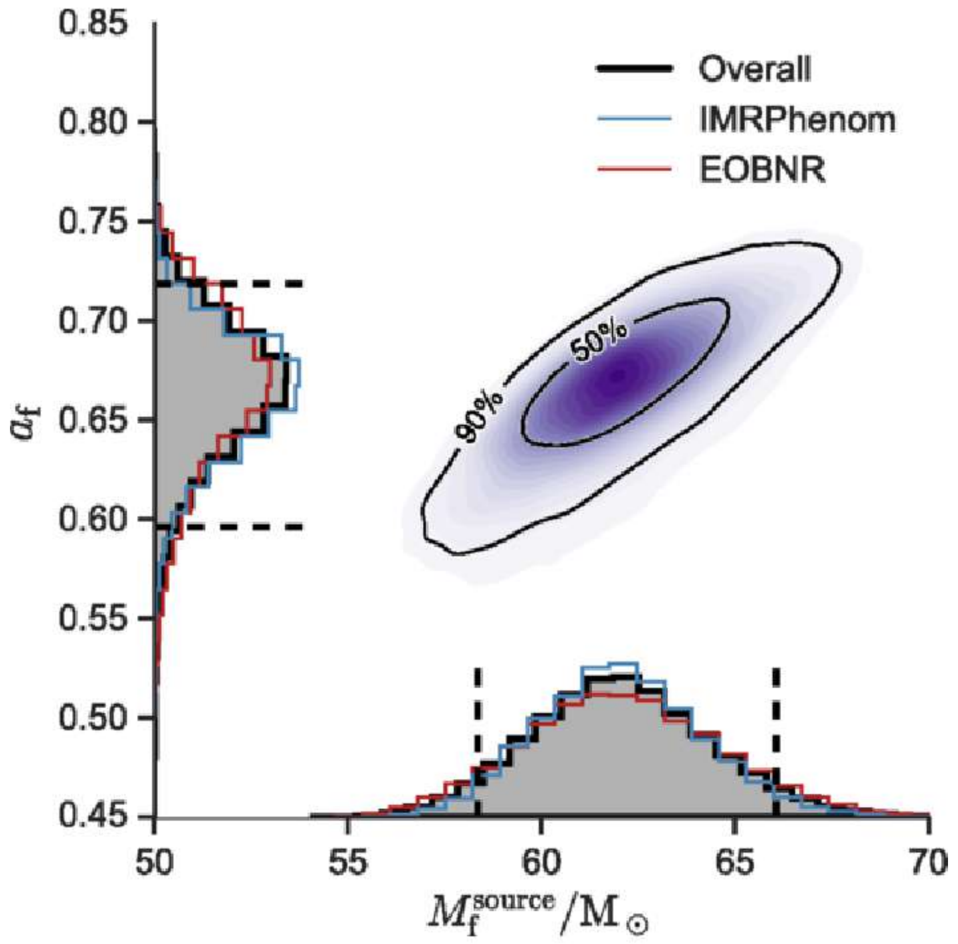
Glitch Classification



Parameter Estimation

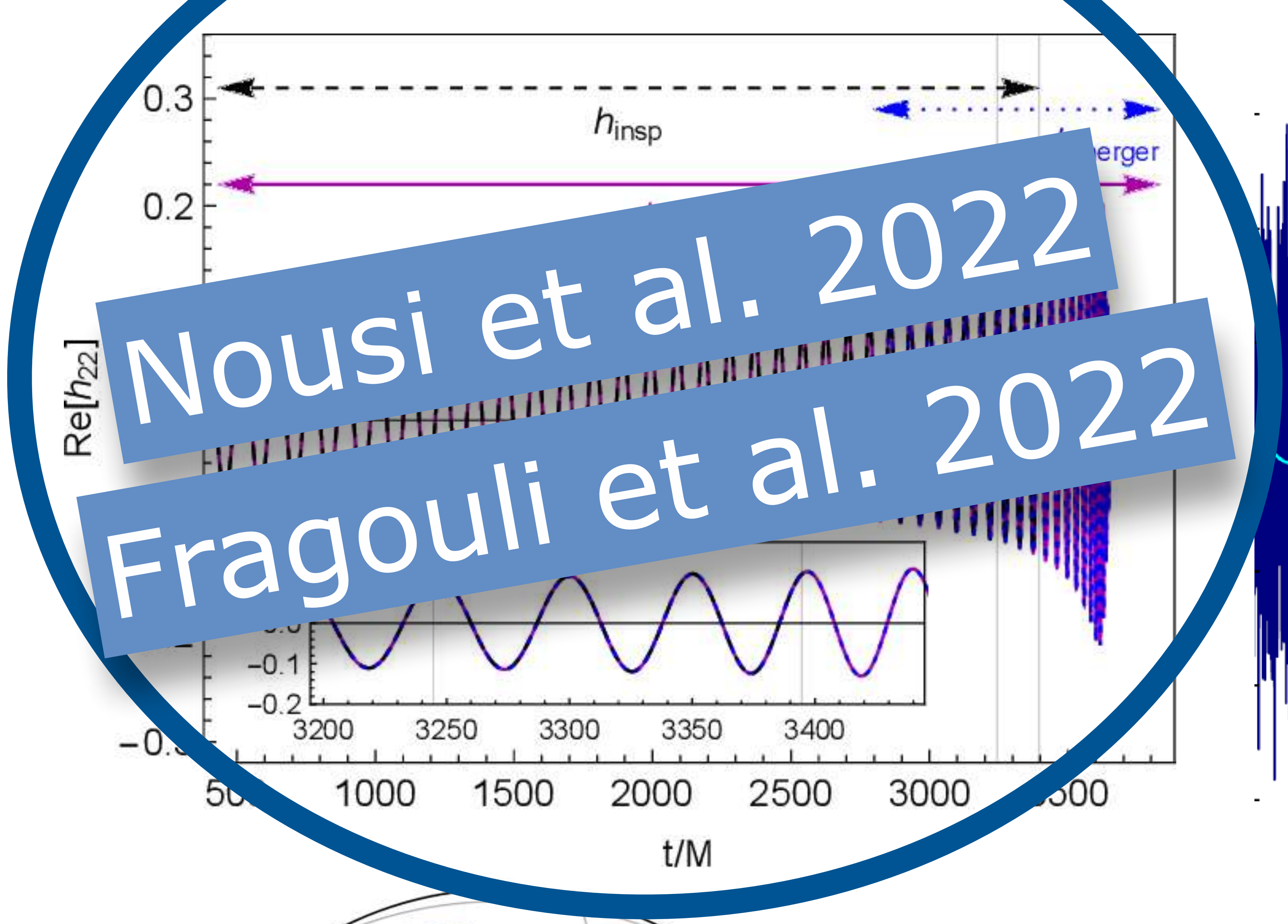
Sky Localization



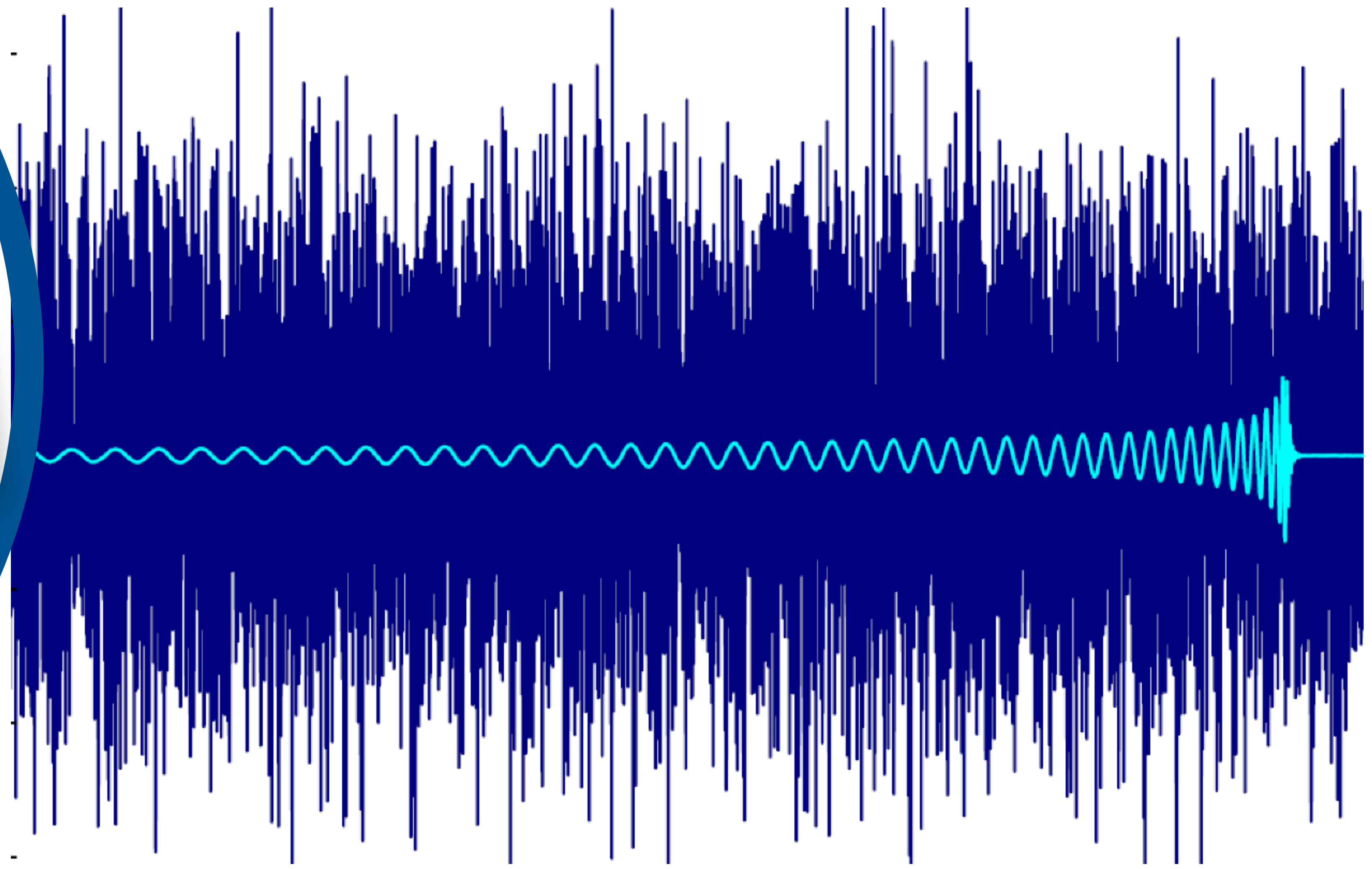


MACHINE LEARNING IN GRAVITATIONAL WAVE ASTRONOMY

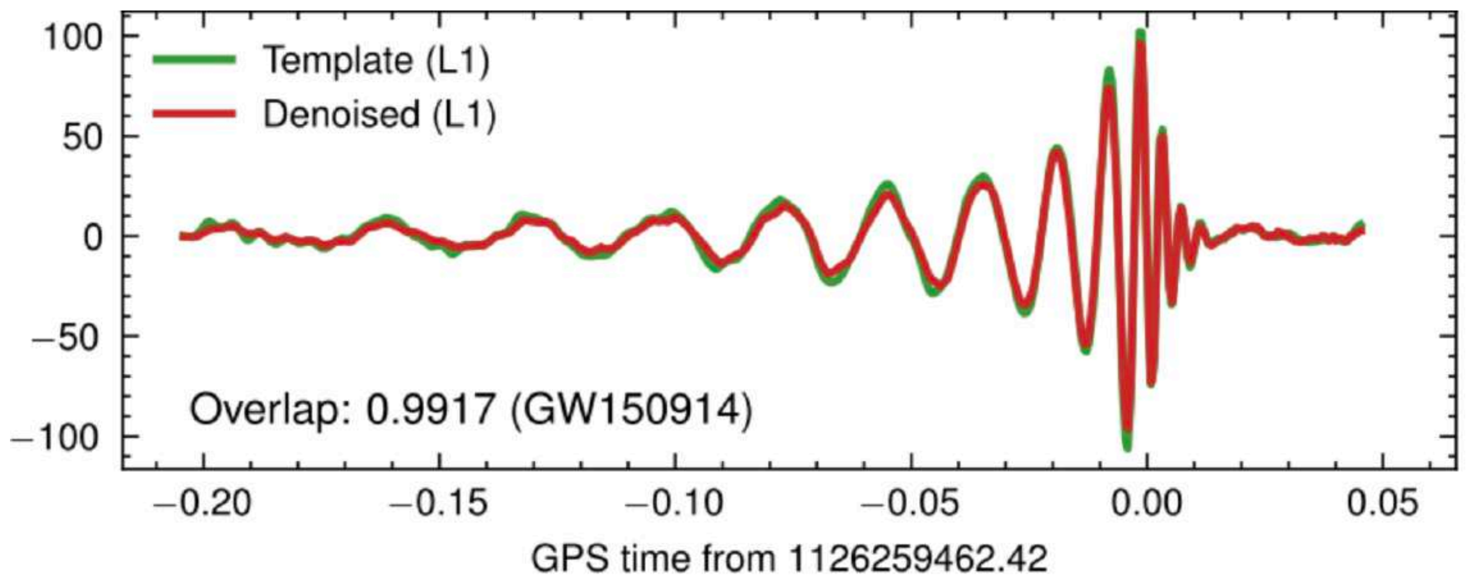
Waveform Modeling



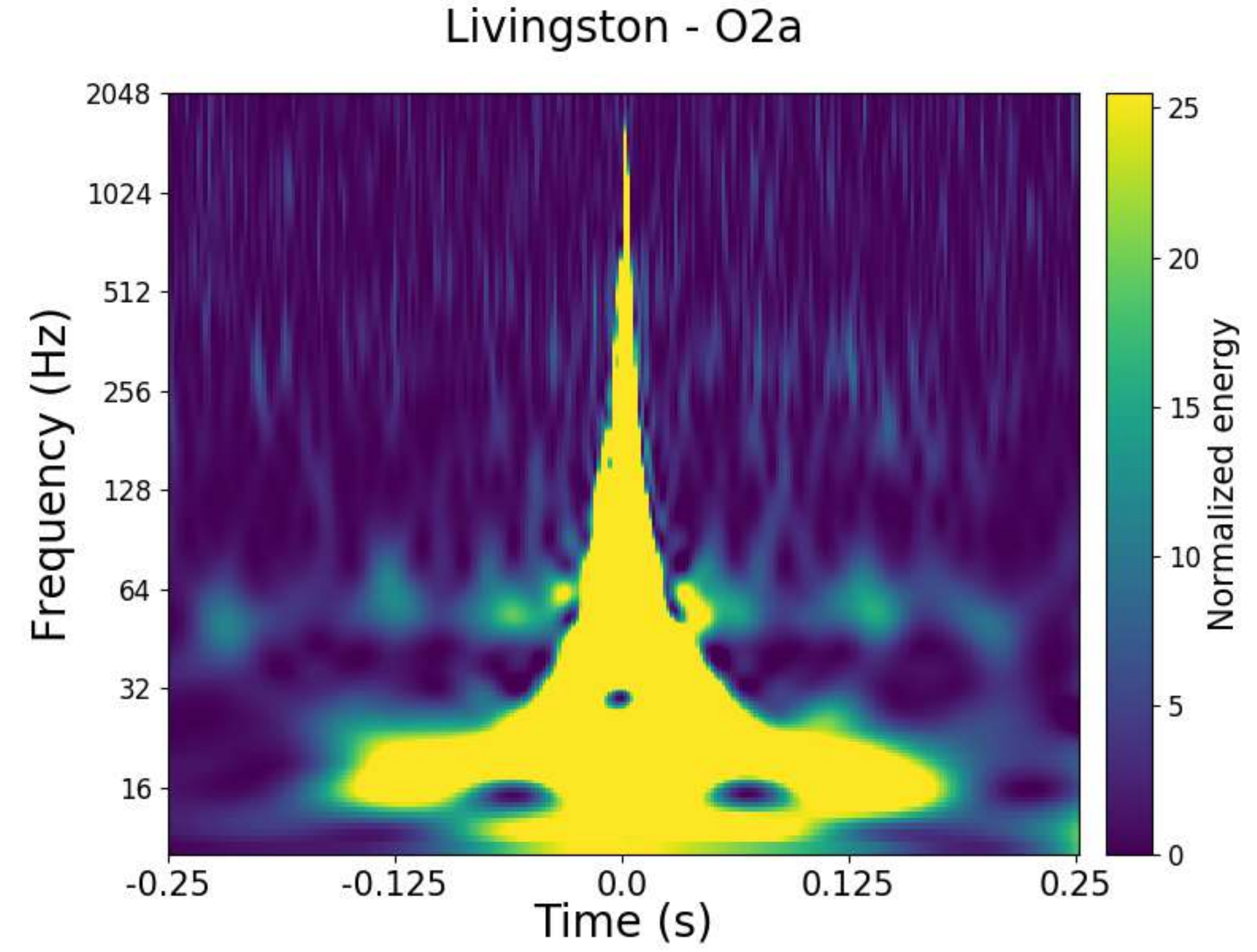
Detection



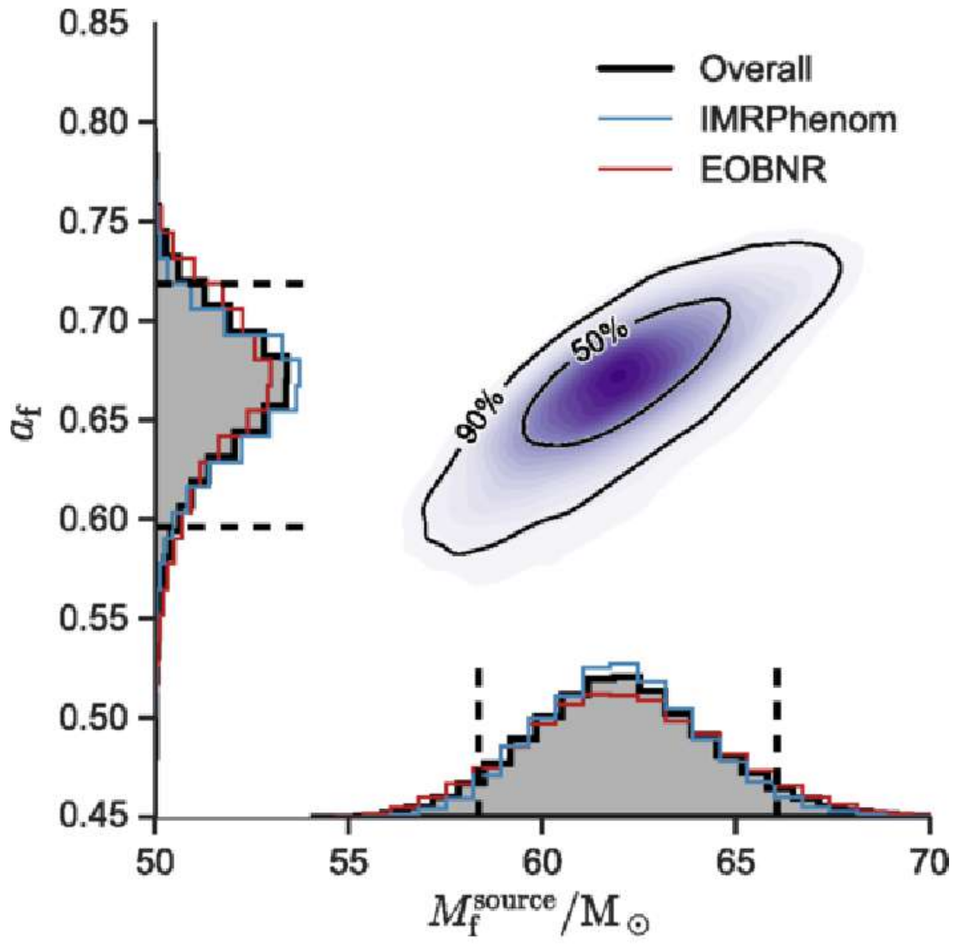
Denoising



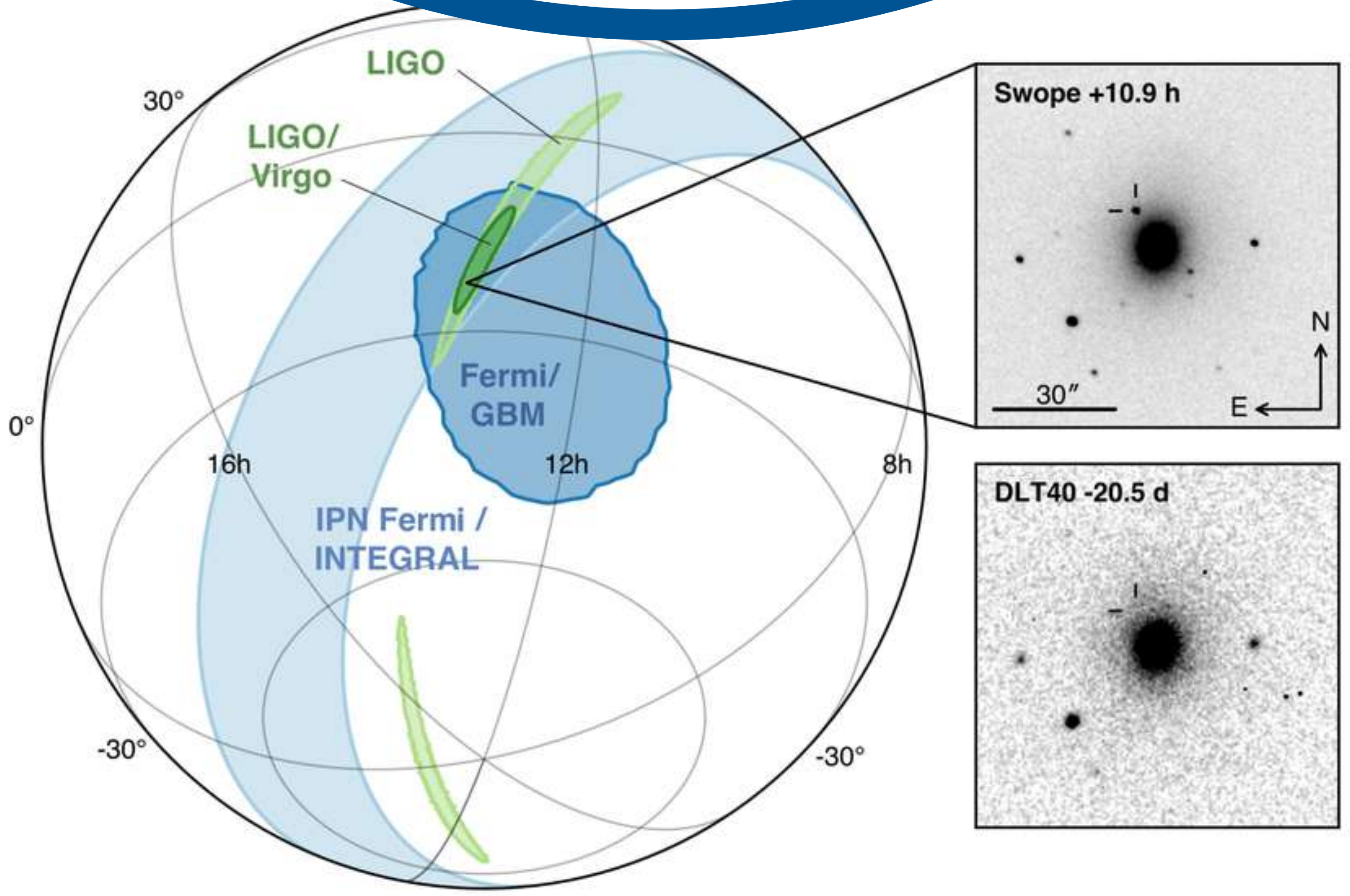
Glitch Classification



Parameter Estimation

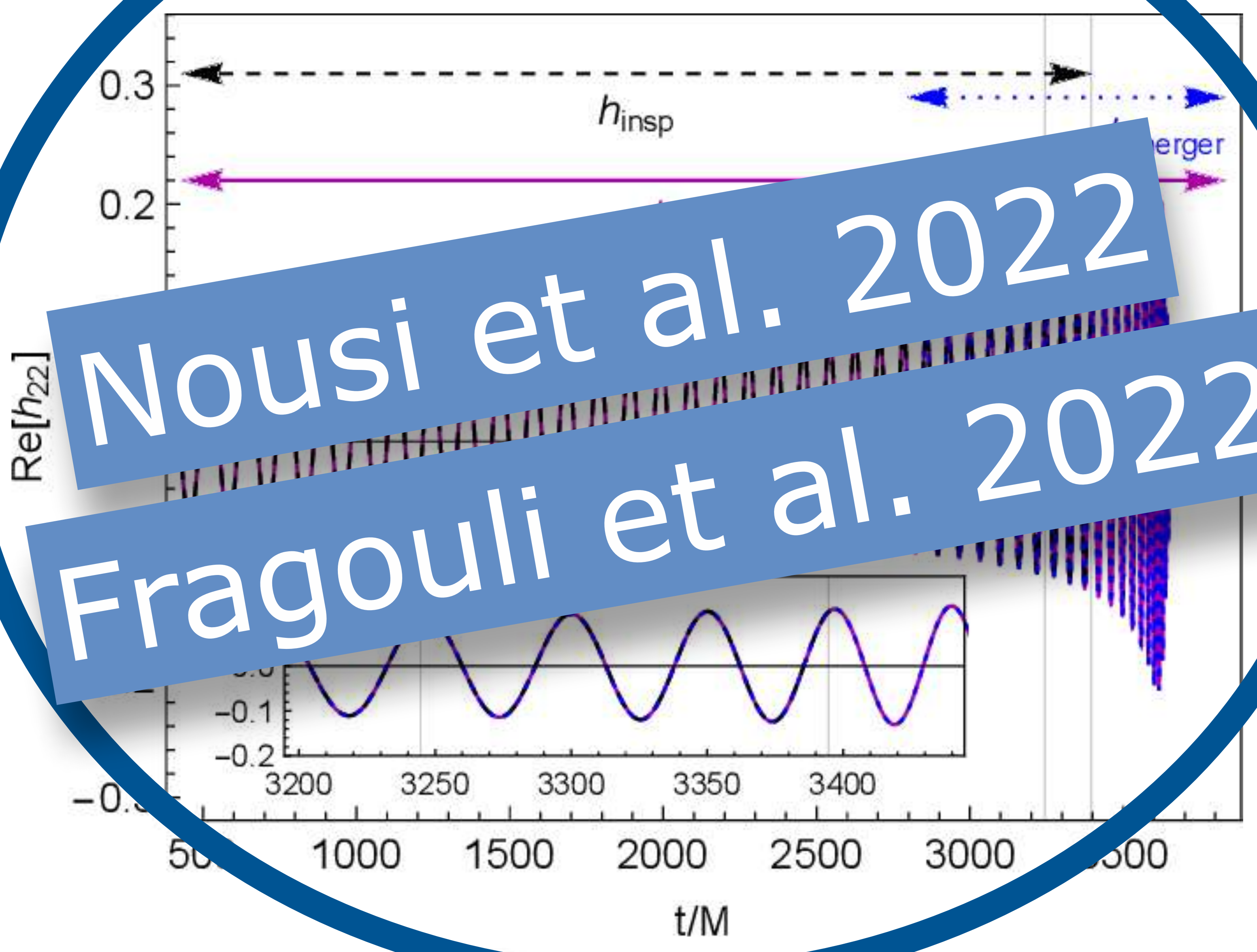


Sky Localization



MACHINE LEARNING IN GRAVITATIONAL WAVE ASTRONOMY

Waveform Modeling



Nousi et al. 2022

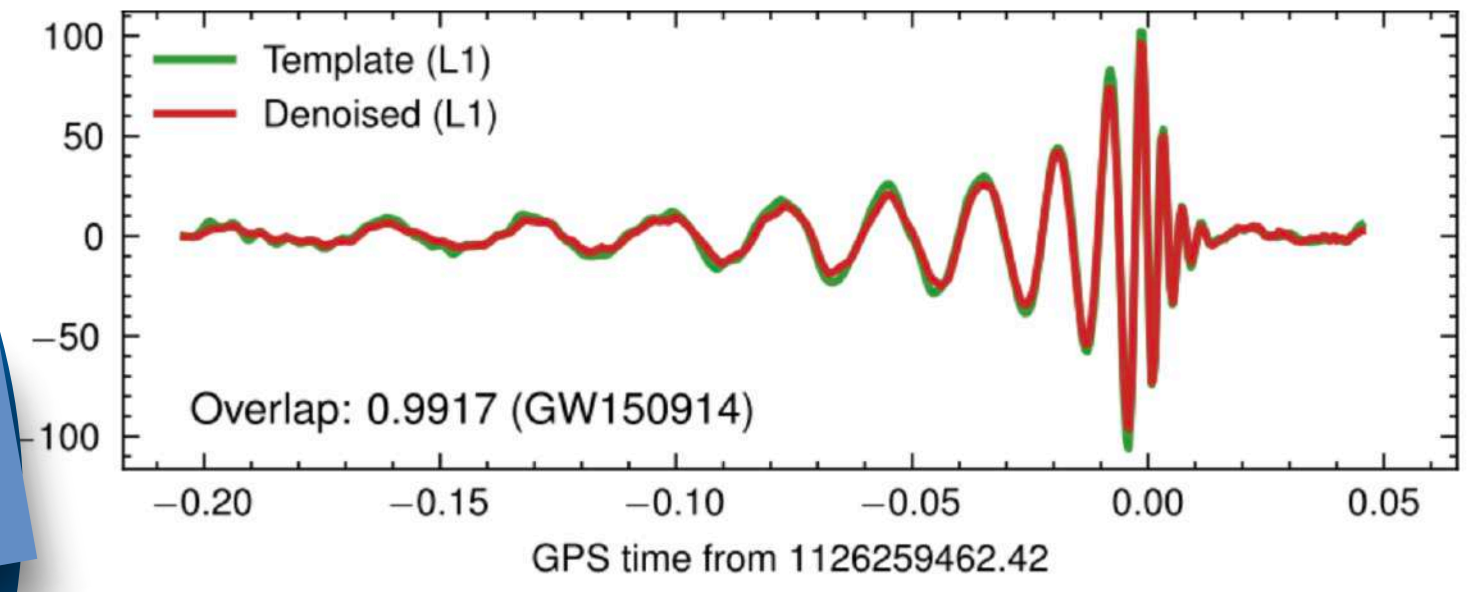
Fragouli et al. 2022

Detection

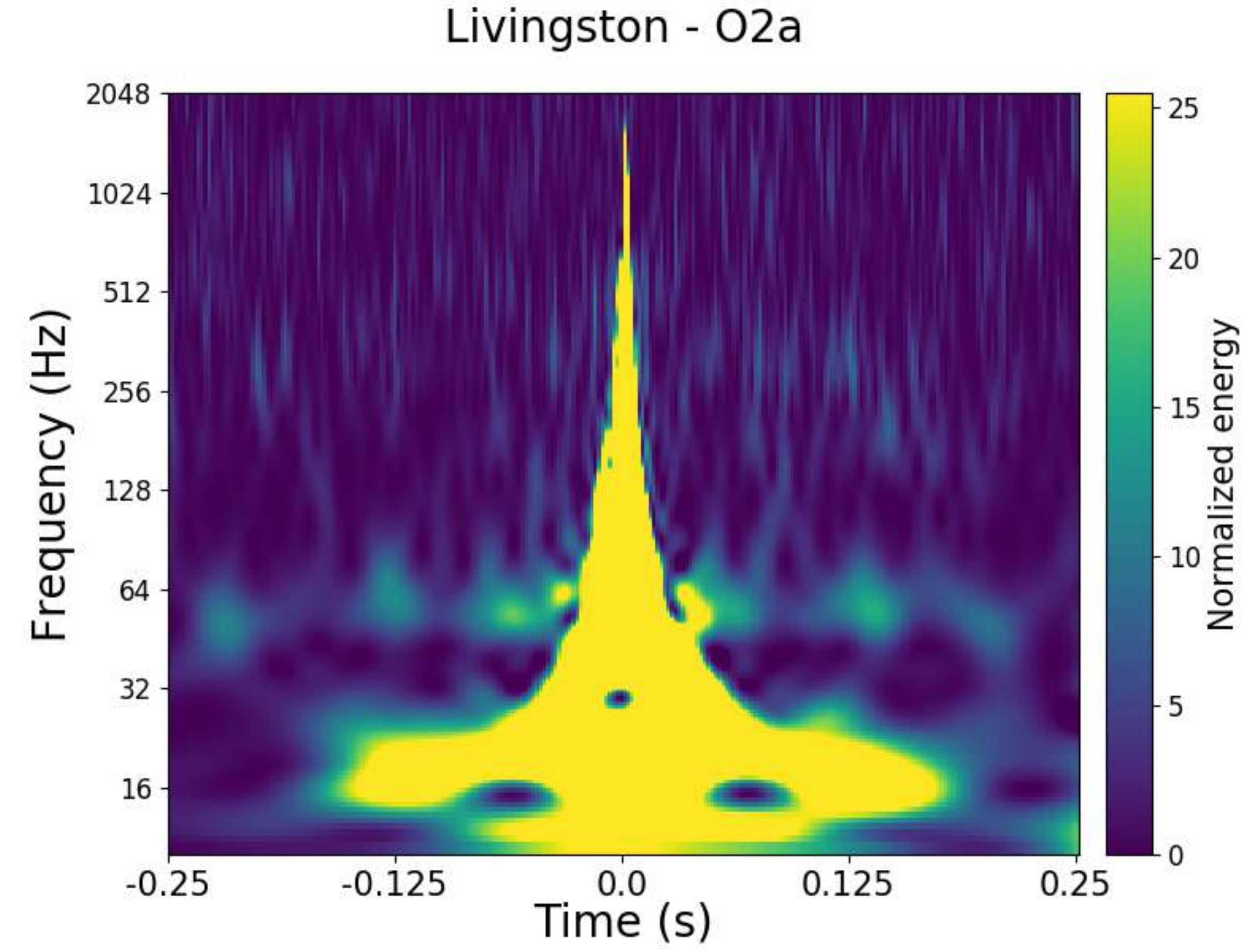
Nousi et al. 2023

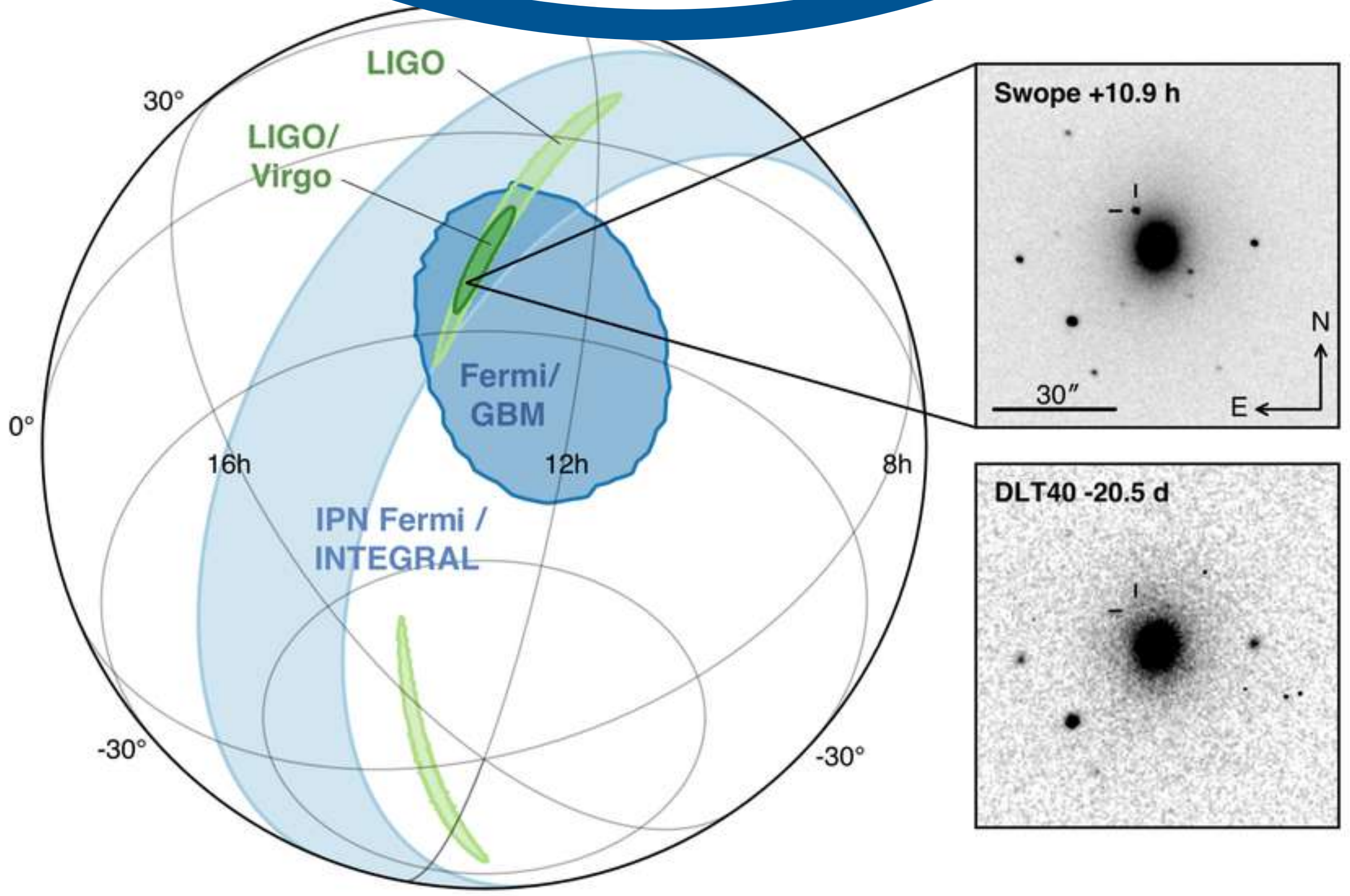
Koloniari et al. 2025

Denoising



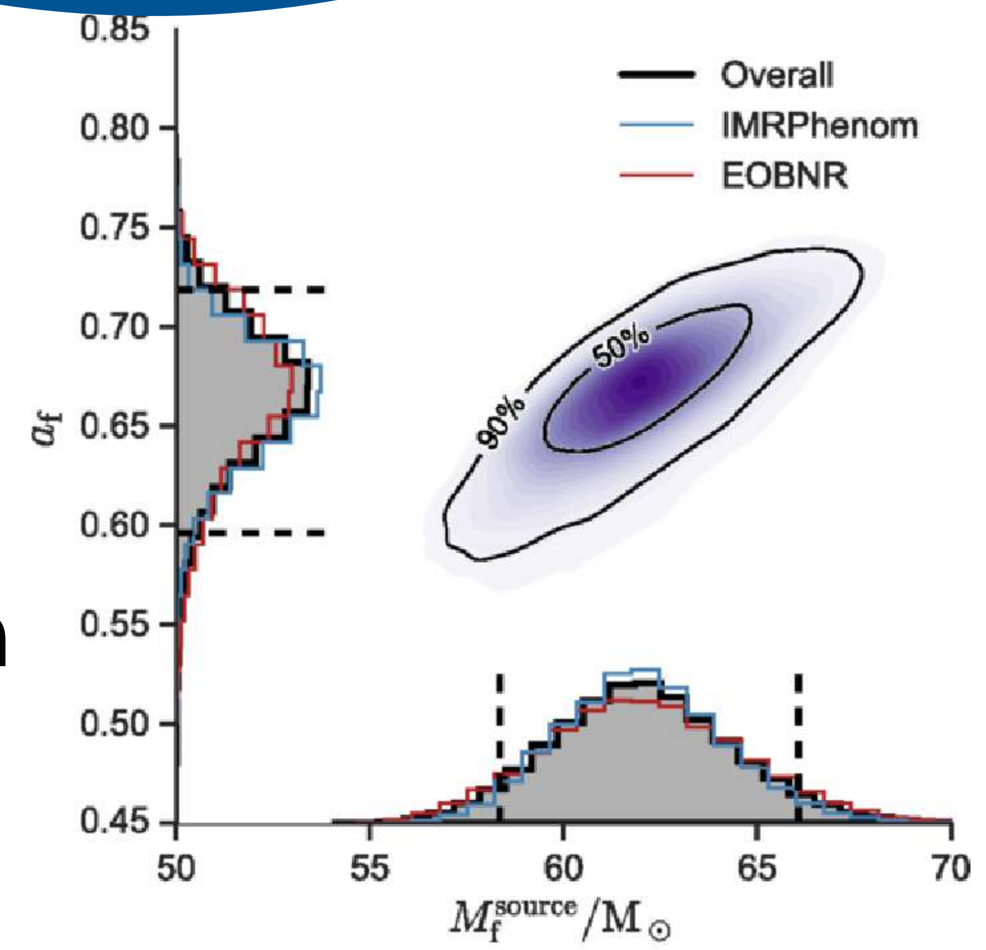
Glitch Classification





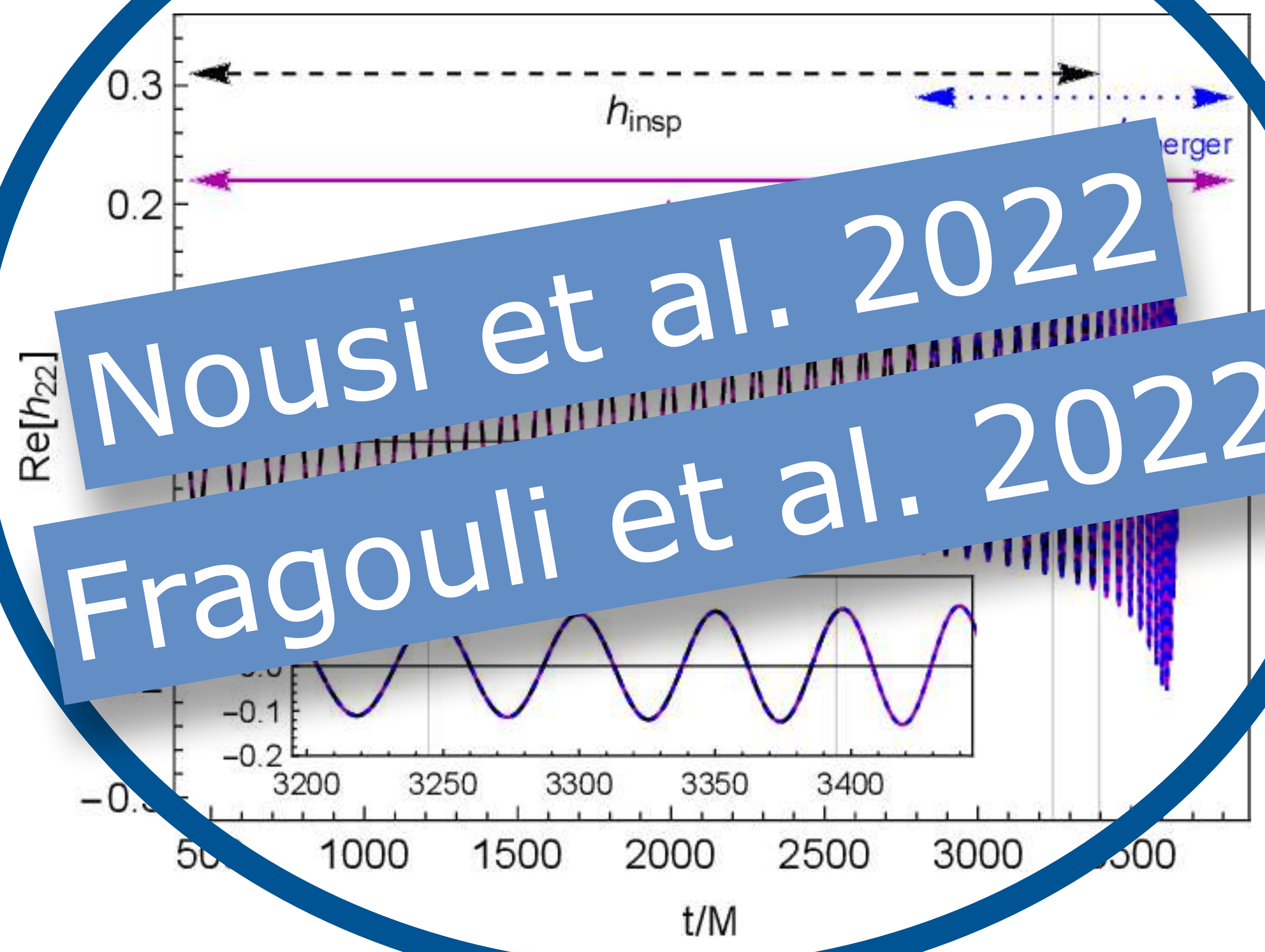
Parameter Estimation

Sky Localization



MACHINE LEARNING IN GRAVITATIONAL WAVE ASTRONOMY

Waveform Modeling



Nousi et al. 2022

Fragouli et al. 2022

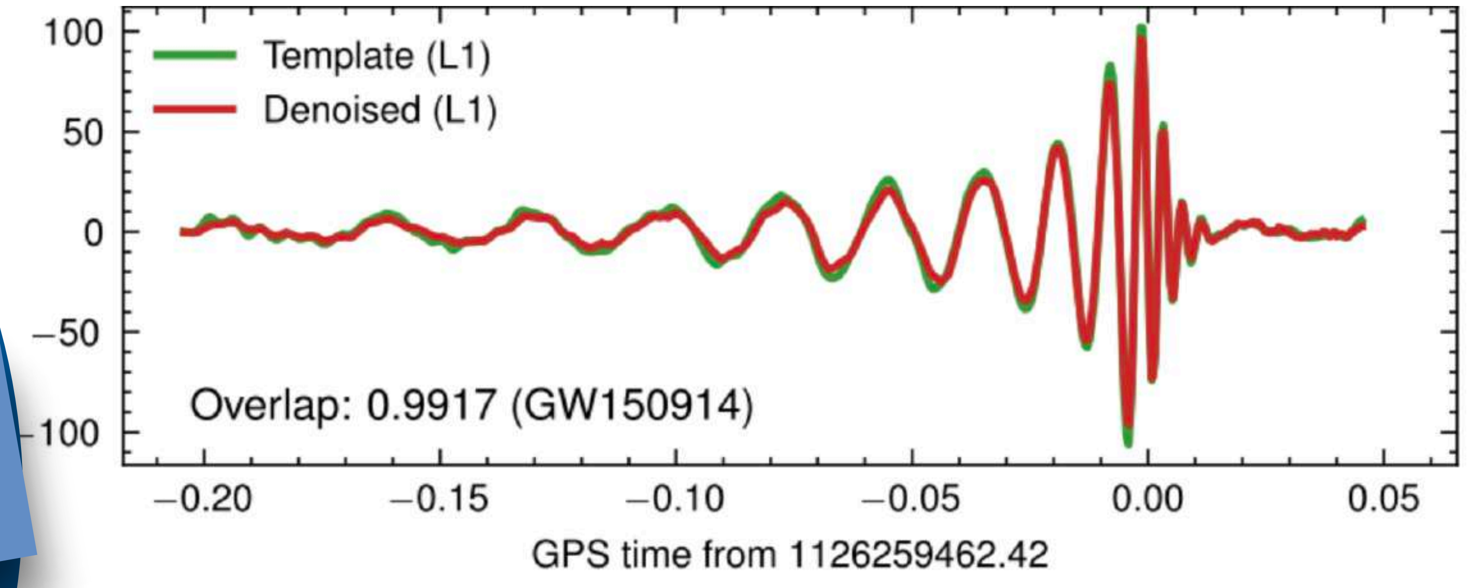
Detection

Nousi et al. 2023

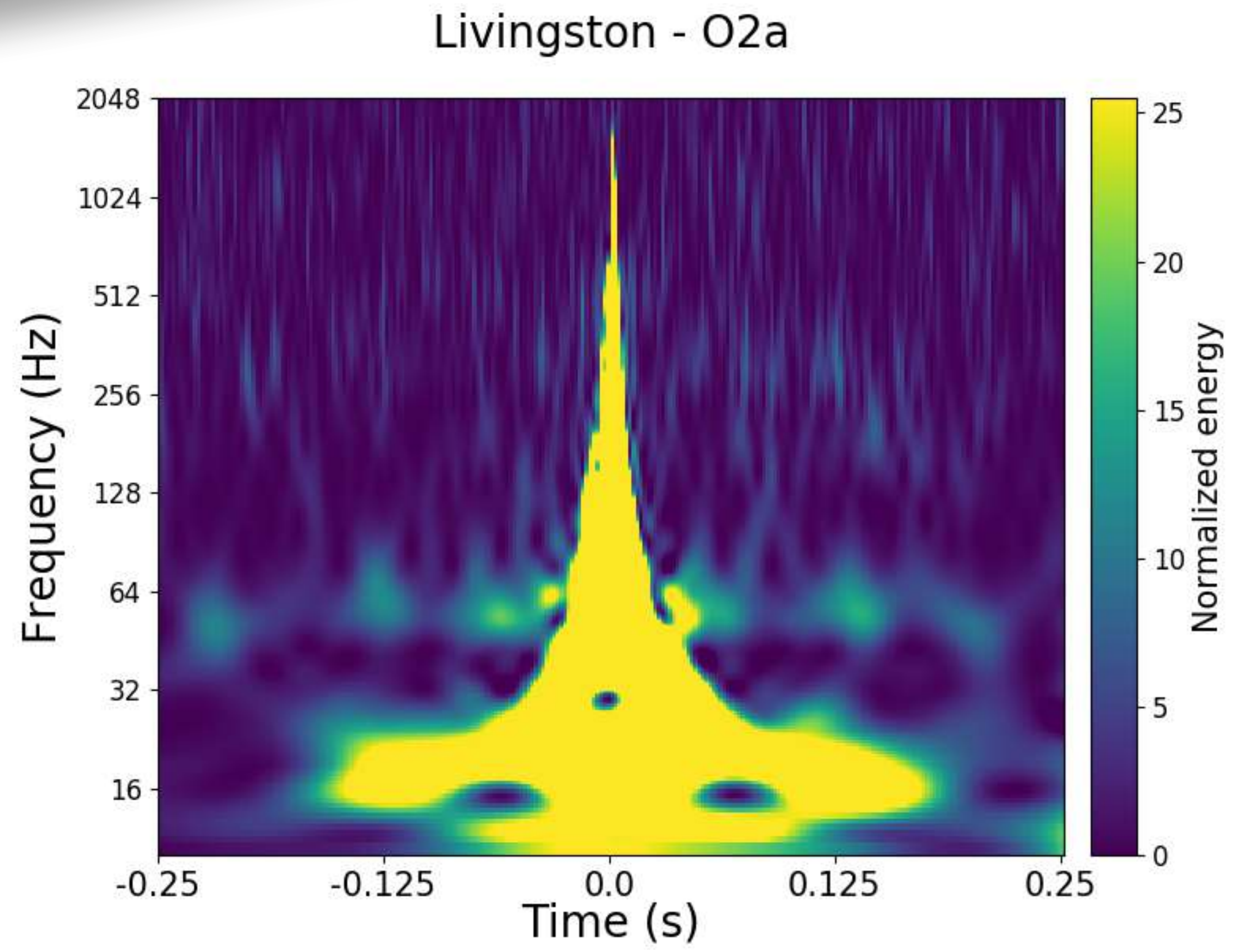
Kolonia ri et al. 2025

See also A. Koloniari and O. Zelenka talks

Denoising

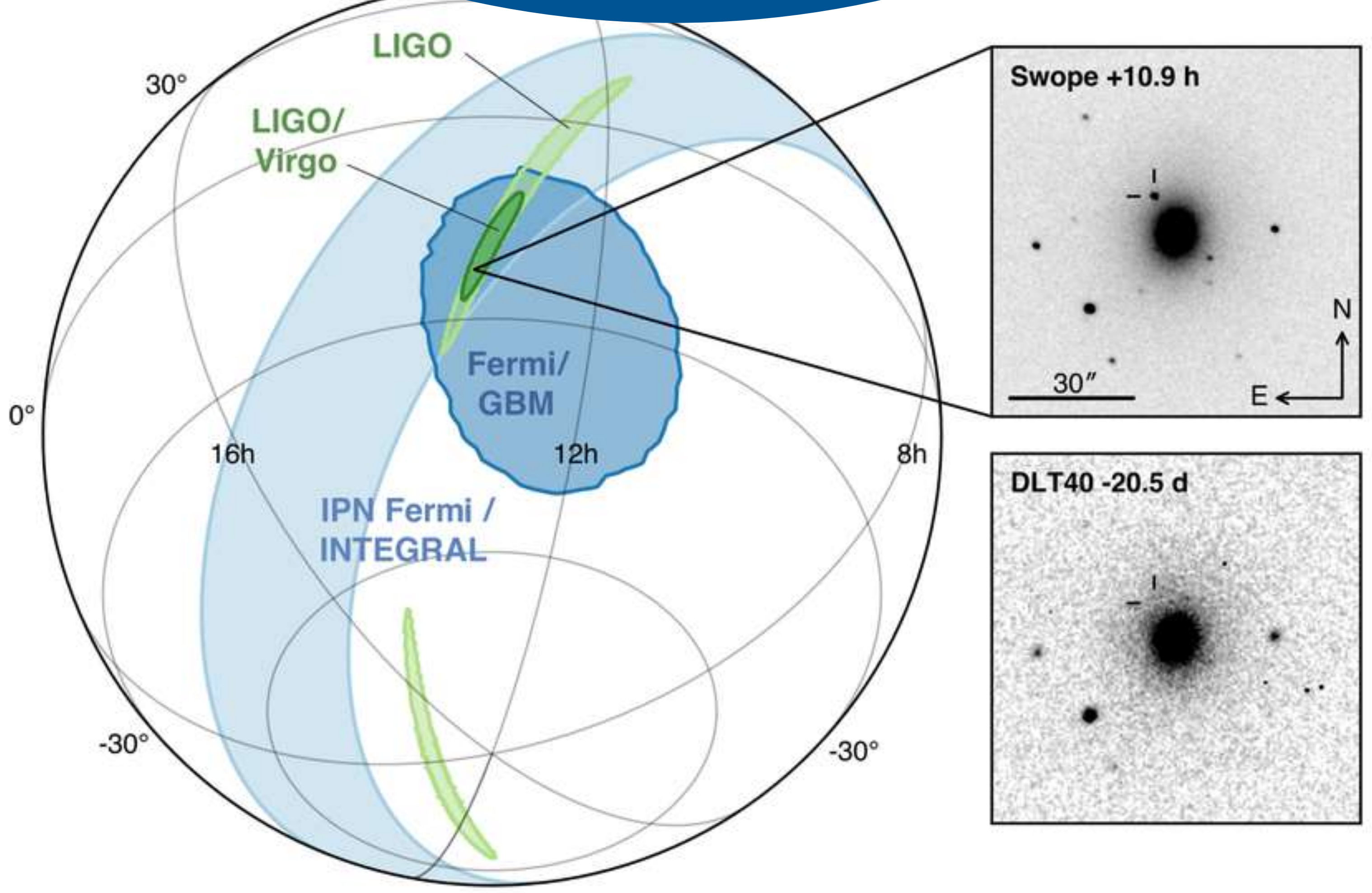


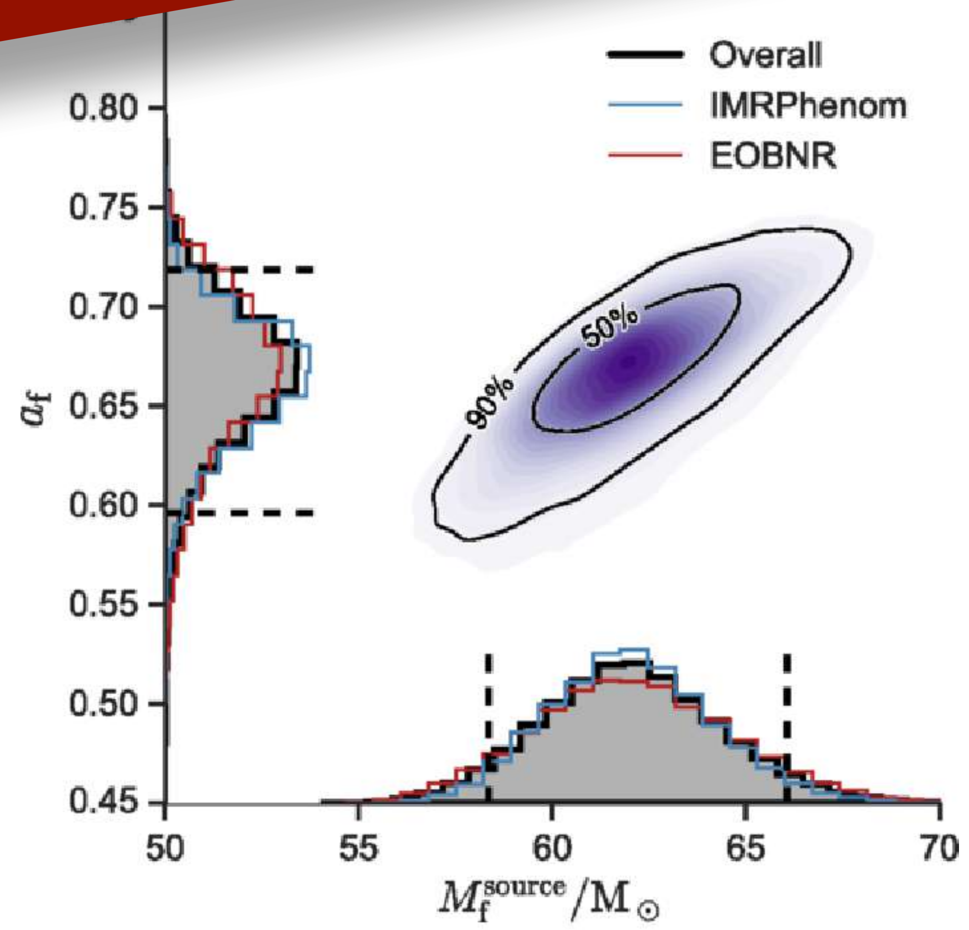
Glitch Classification



Parameter Estimation

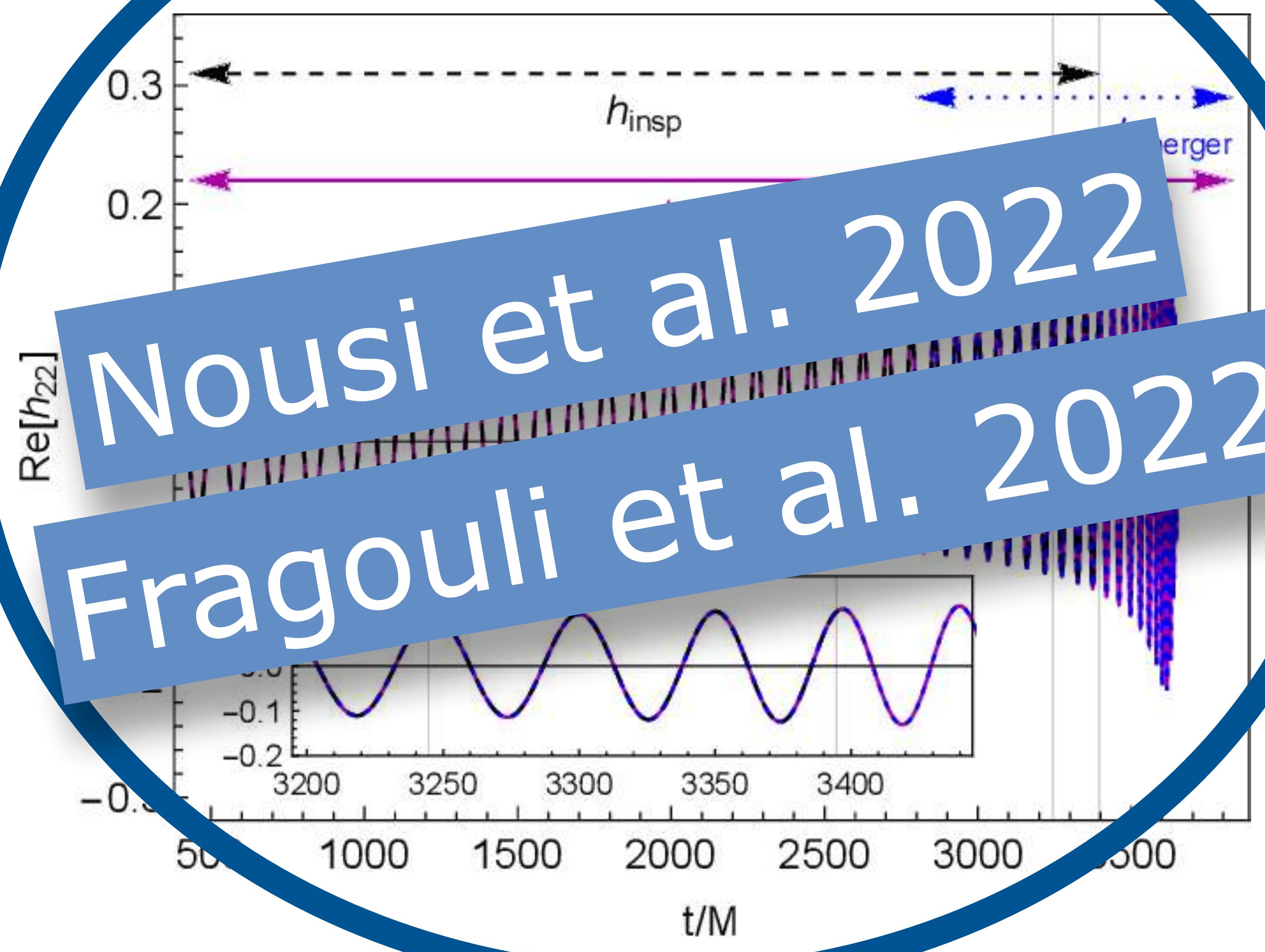
Sky Localization





MACHINE LEARNING IN GRAVITATIONAL WAVE ASTRONOMY

Waveform Modeling



Nousi et al. 2022

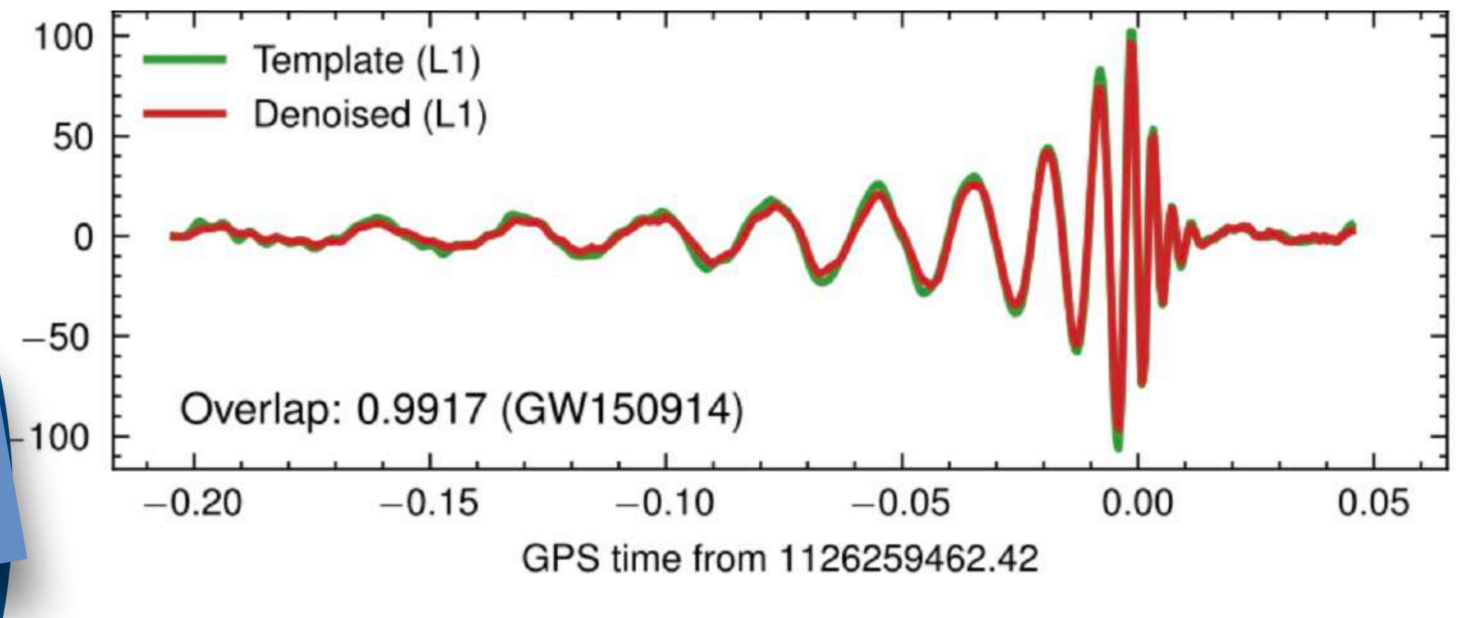
Fragouli et al. 2022

Detection

Nousi et al. 2023

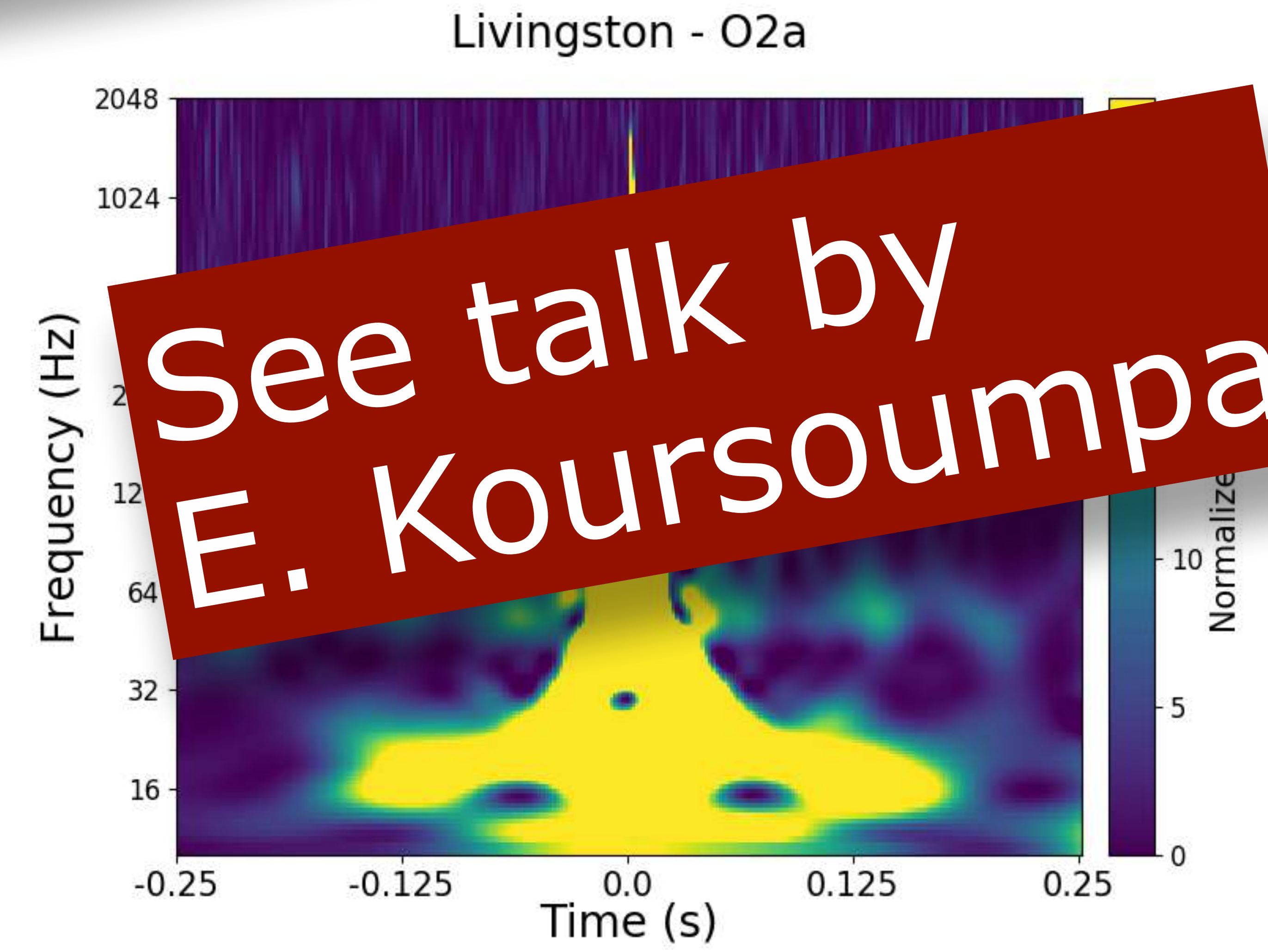
Koloniari et al. 2025

Denoising

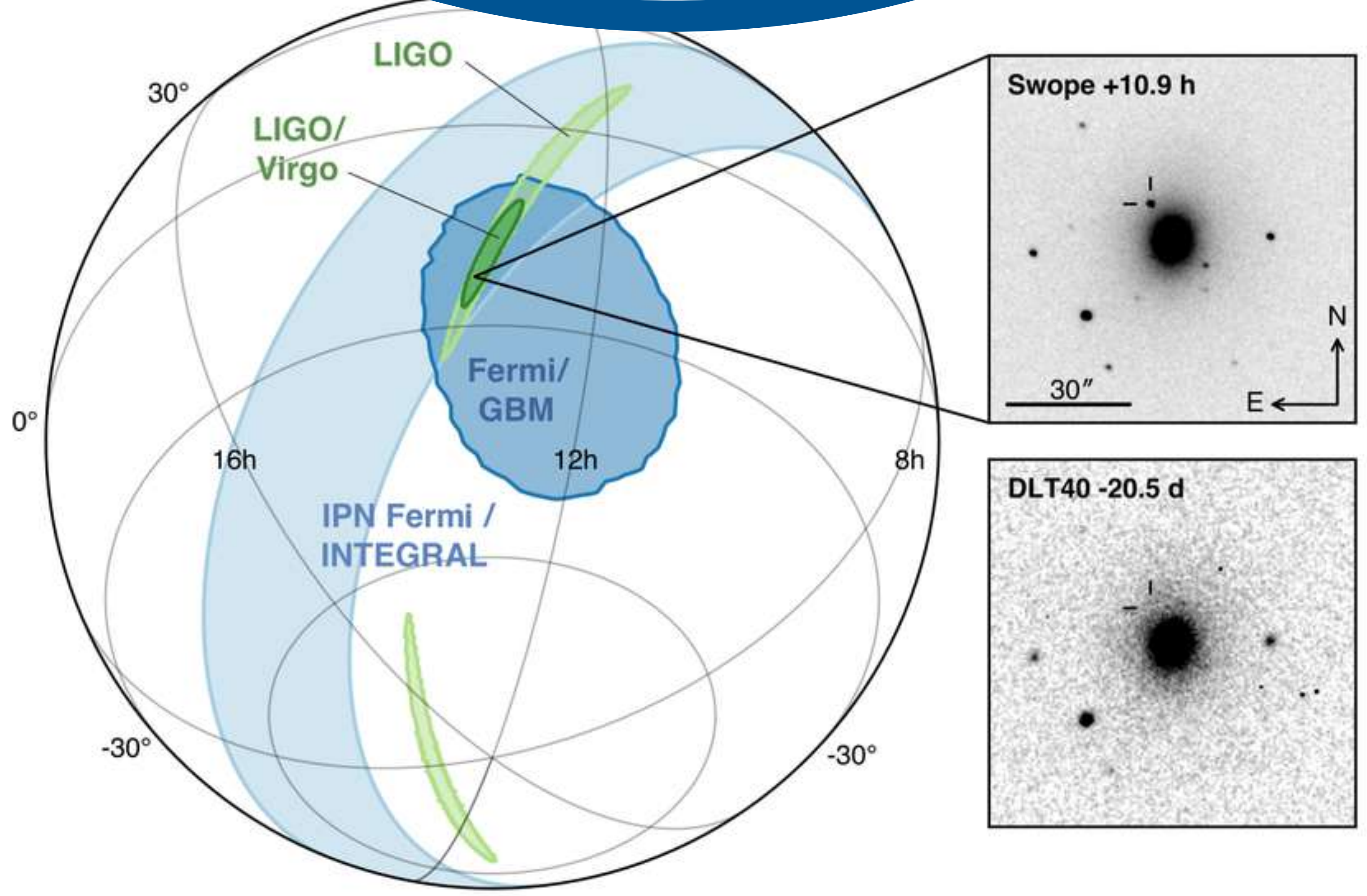


See also A. Koloniari and O. Zelenka talks

Glitch Classification

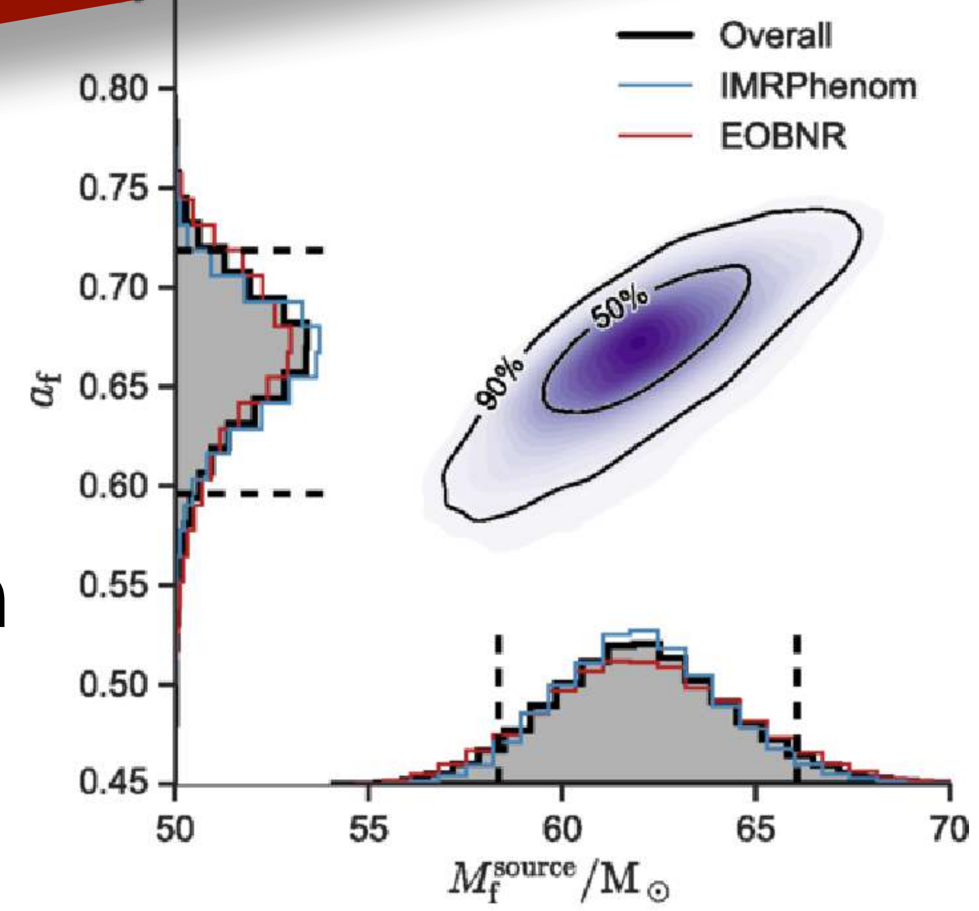


See talk by E. Koursoumpa



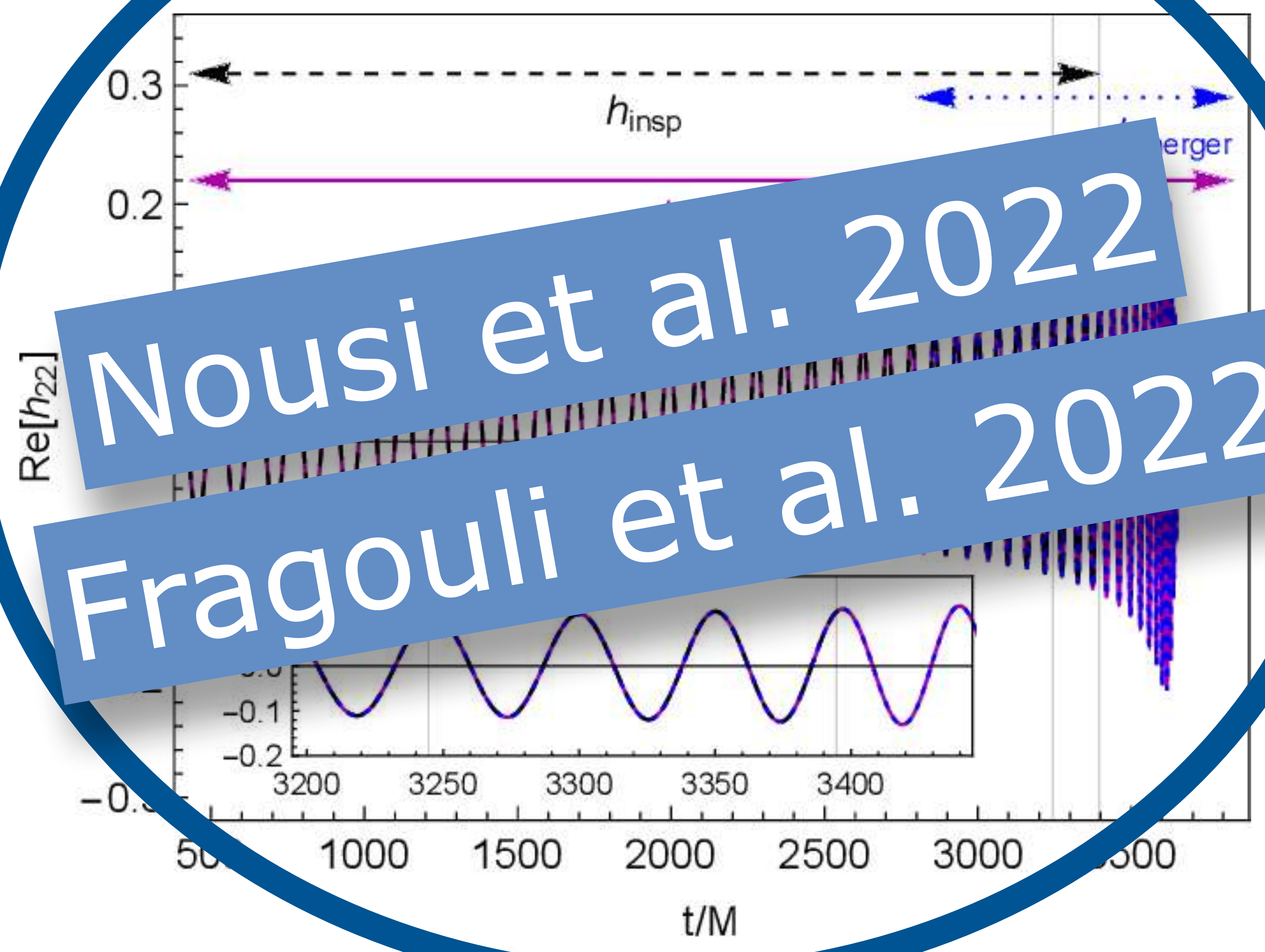
Parameter Estimation

Sky Localization



MACHINE LEARNING IN GRAVITATIONAL WAVE ASTRONOMY

Waveform Modeling



Nousi et al. 2022

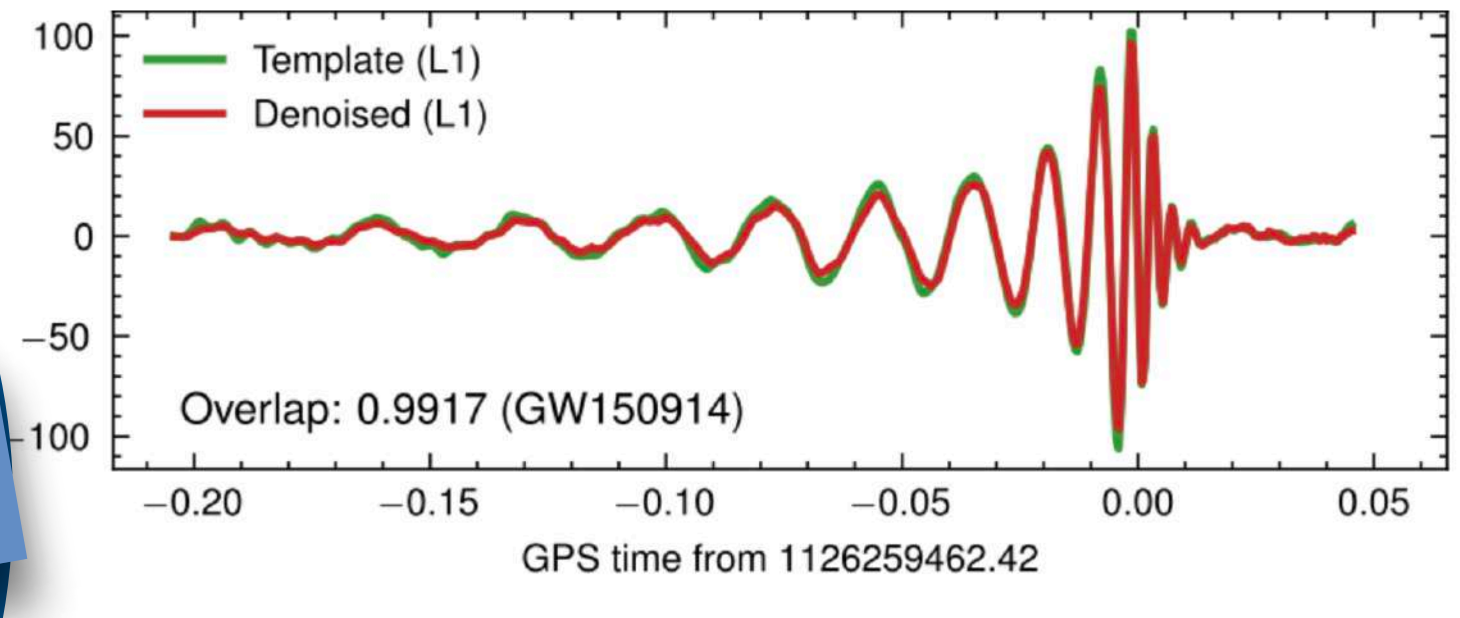
Fragouli et al. 2022

Detection

Nousi et al. 2023

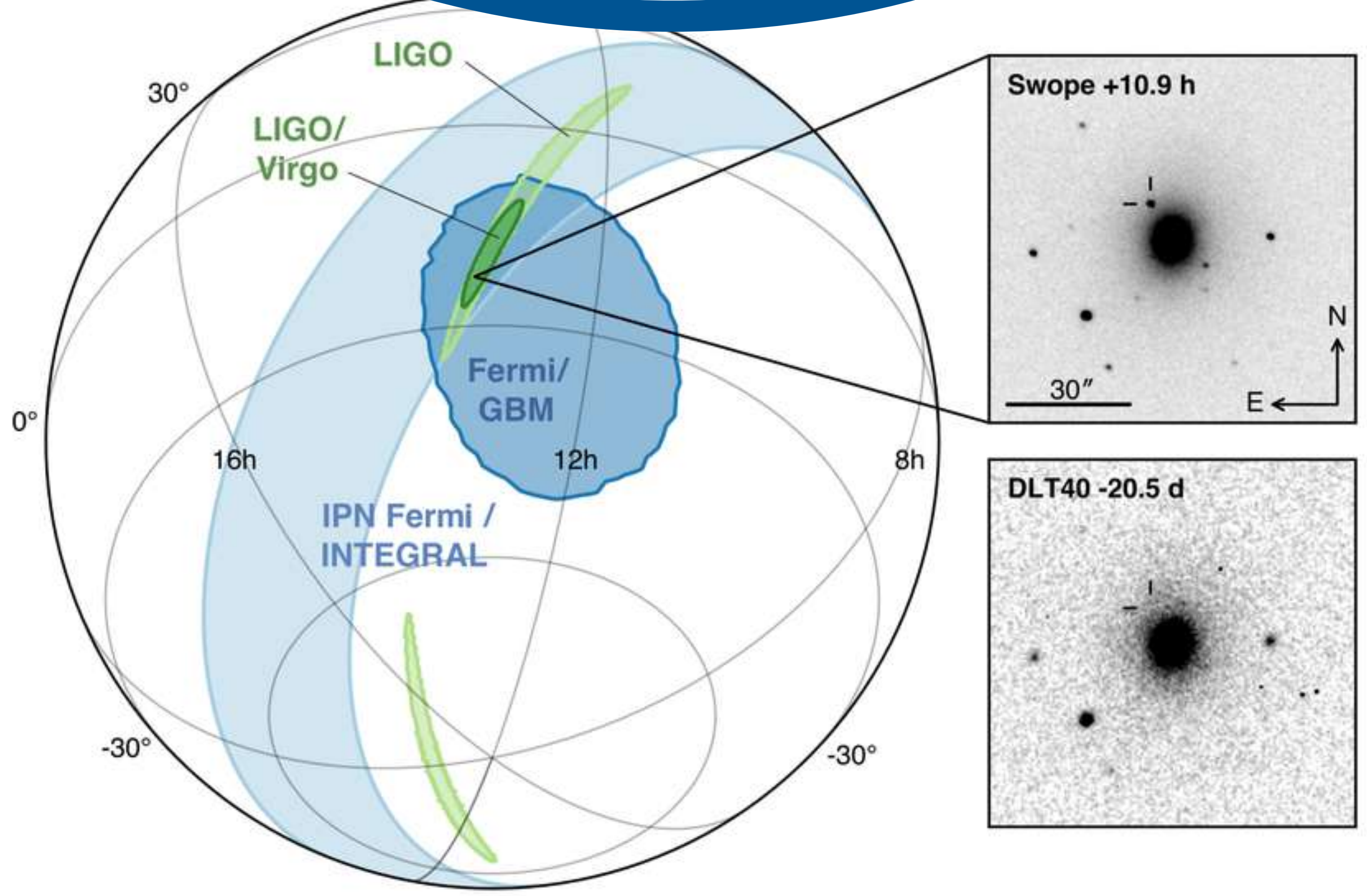
Koloniari et al. 2025

Denoising



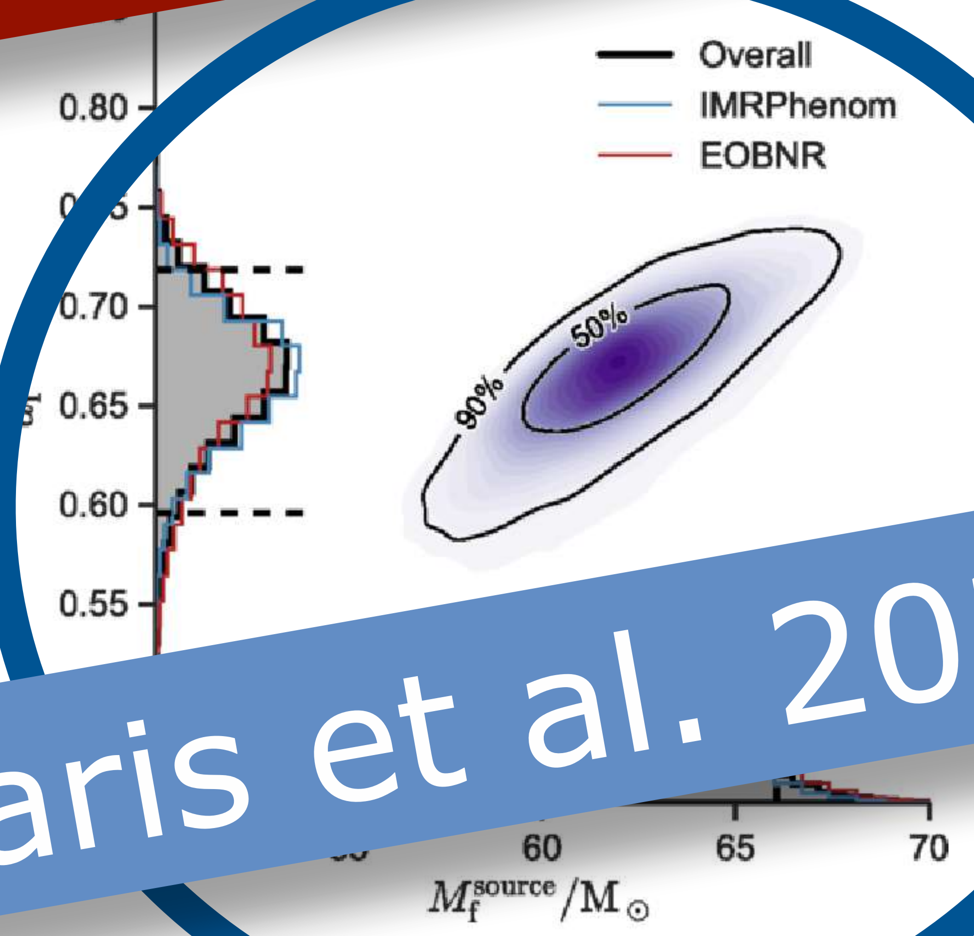
Glitch Classification

See also A. Koloniari and O. Zelenka talks

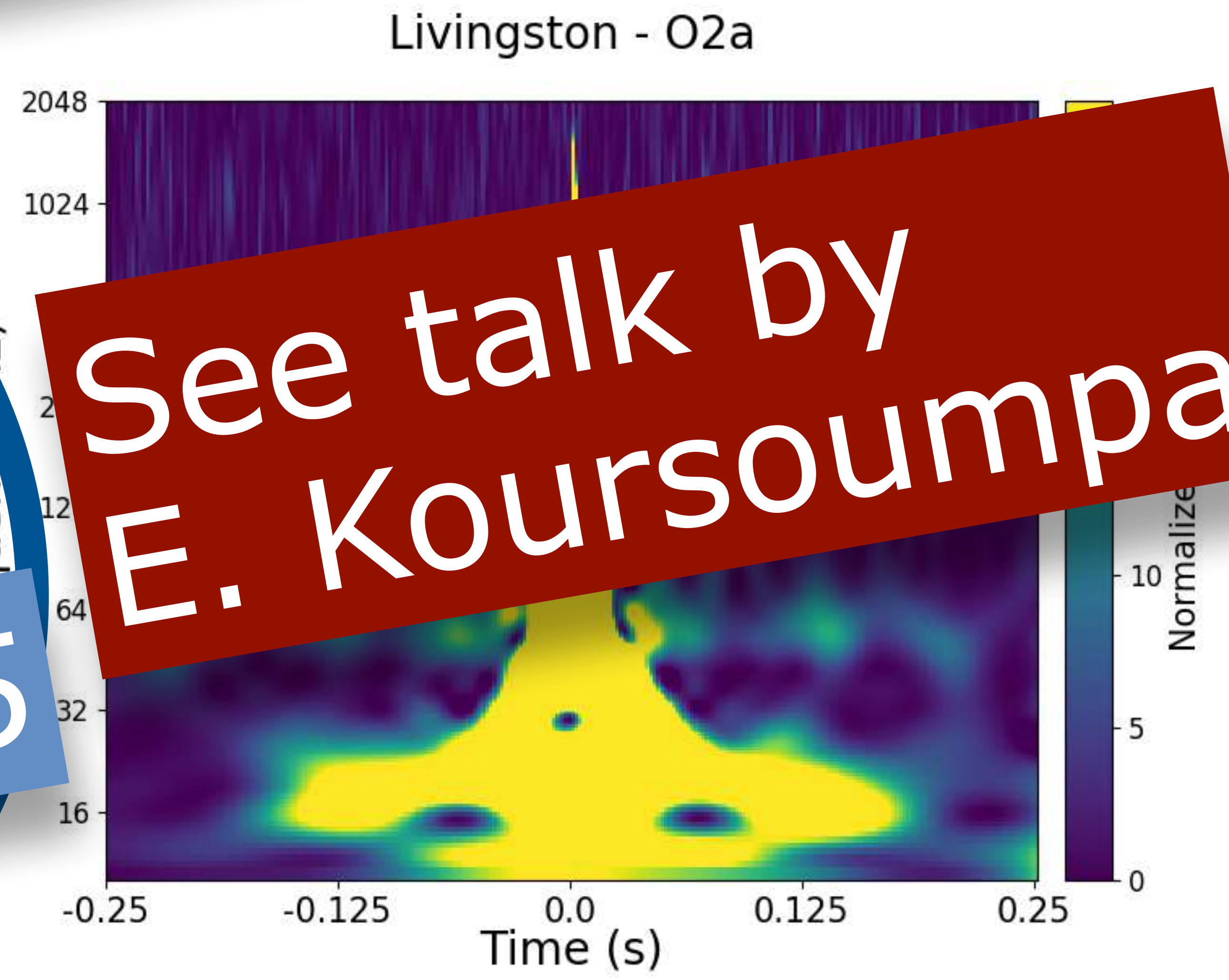


Parameter Estimation

Sky Localization



Vretinaris et al. 2025



See talk by E. Koursoumpa

GRAVITATIONAL WAVE DETECTION

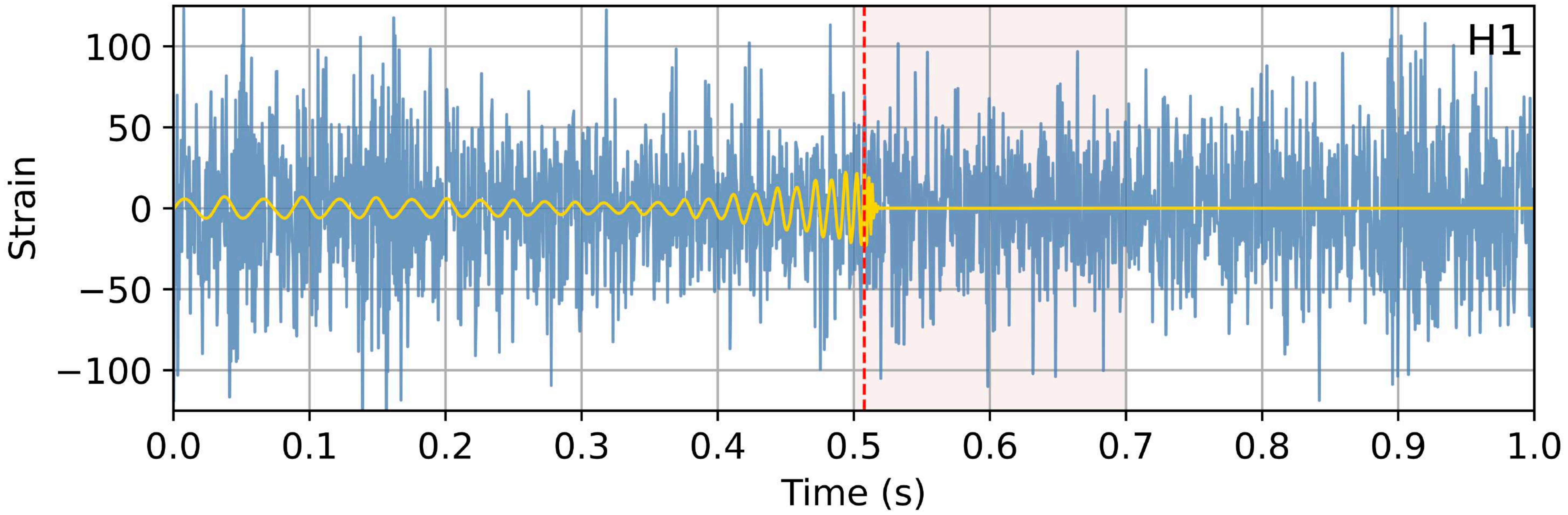
PHYSICAL REVIEW D **108,** 024022 (2023)

Deep residual networks for gravitational wave detection

Paraskevi Nousi,[1] Alexandra E. Koloniari,[2] Nikolaos Passalis,[1] Panagiotis Iosif,[2] Nikolaos Stergioulas,[2] and Anastasios Tefas[1]

[1]*Department of Informatics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece*
[2]*Department of Physics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece*

(Received 17 December 2022; accepted 14 June 2023; published 11 July 2023)



ARESGW

Architecture:

- 54-layer Resnet-1D
- Deep Adaptive Input Normalization
- SNR-based Curriculum Learning
- 30x faster than PyCBC (using a single GPU card)

ARESGW

Architecture:

- 54-layer Resnet-1D
- Deep Adaptive Input Normalization
- SNR-based Curriculum Learning
- 30x faster than PyCBC (using a single GPU card)

Training: 1-second segments @2kHz of BBH injections with IMRPhenomXPHM in real O3 noise from L1 and H1

Mass range: $7M_\odot \leq M \leq 50M_\odot$

ARESGW

Architecture:

- 54-layer Resnet-1D
- Deep Adaptive Input Normalization
- SNR-based Curriculum Learning
- 30x faster than PyCBC (using a single GPU card)

Training: 1-second segments @2kHz of BBH injections with IMRPhenomXPHM in real O3 noise from L1 and H1

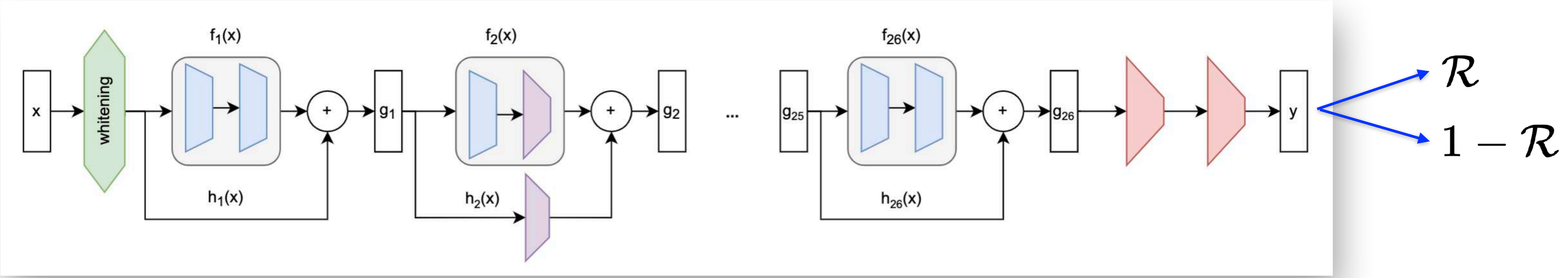
Mass range: $7M_{\odot} \leq M \leq 50M_{\odot}$

Leading algorithm (Virgo-AUTH) in the **1st ML GW search challenge** in the most demanding dataset.

`https://github.com/gwastro/ml-mock-data-challenge-1`

NETWORK ARCHITECTURE OF ARES-GW

1-D ResNet-54 (27 residual blocks with 2 convolutional layers each and skip connections)!

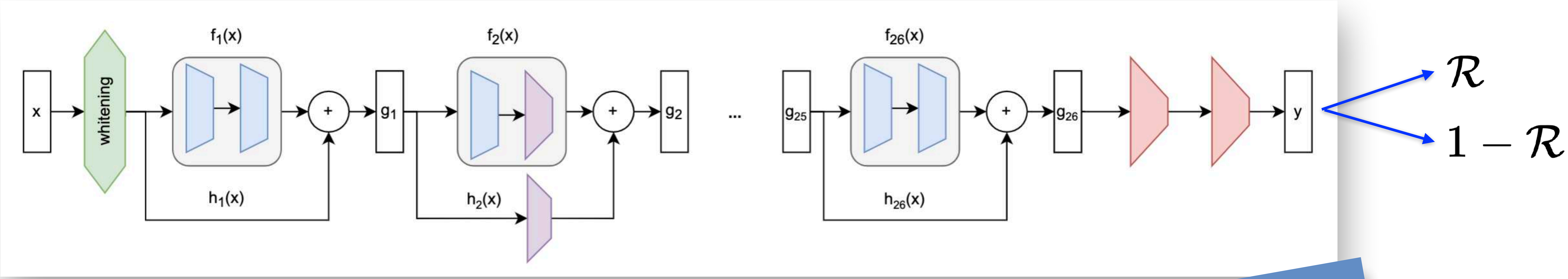


- Residual blocks with skip connections: g(**x**) = f(**x**) +h(**x**)
- h(**x**) = **x** *or* a strided convolutional layer
- After each convolutional layer: batch normalization + ReLU activation
- Mini-batch size of 400 segments
- Adam optimizer for back propagation
- Objective function = regularized binary cross entropy

Residual blocks	Filters	Strided	Input D
4	8		2×2048
1	16	✓	8×2048
2	16		16×1024
1	32	✓	16×1024
2	32		32×512
1	64	✓	32×512
2	64		64×256
1	64	✓	64×256
2	64		64×128
1	64	✓	64×128
2	64		64×64
5	32		64×64
3	16		32×64

NETWORK ARCHITECTURE OF ARES-GW

1-D ResNet-54 (27 residual blocks with 2 convolutional layers each and skip connections)!



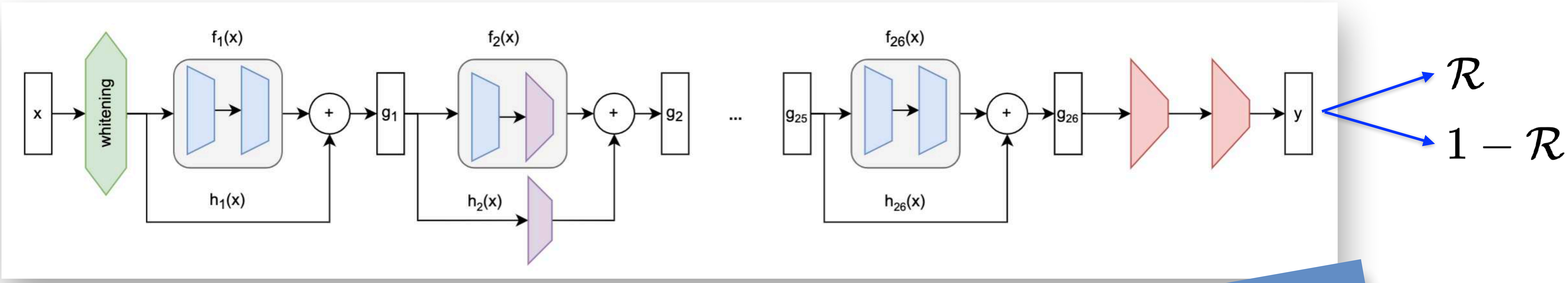
Gradient problem solved!
-> much deeper networks

- Residual blocks with skip connections: g(**x**) = f(
- h(**x**) = **x** *or* a strided convolutional layer
- After each convolutional layer: batch nor
  ReLU activation
- Mini-batch size of 400 segments
- Adam optimizer for back propagation
- Objective function = regularized binary cross entropy

			Input D
			2×2048
	[illegible]	✓	8×2048
	16		16×1024
1	32	✓	16×1024
2	32		32×512
1	64	✓	32×512
2	64		64×256
1	64	✓	64×256
2	64		64×128
1	64	✓	64×128
2	64		64×64
5	32		64×64
3	16		32×64

NETWORK ARCHITECTURE OF ARES-GW

1-D ResNet-54 (27 residual blocks with 2 convolutional layers each and skip connections)!



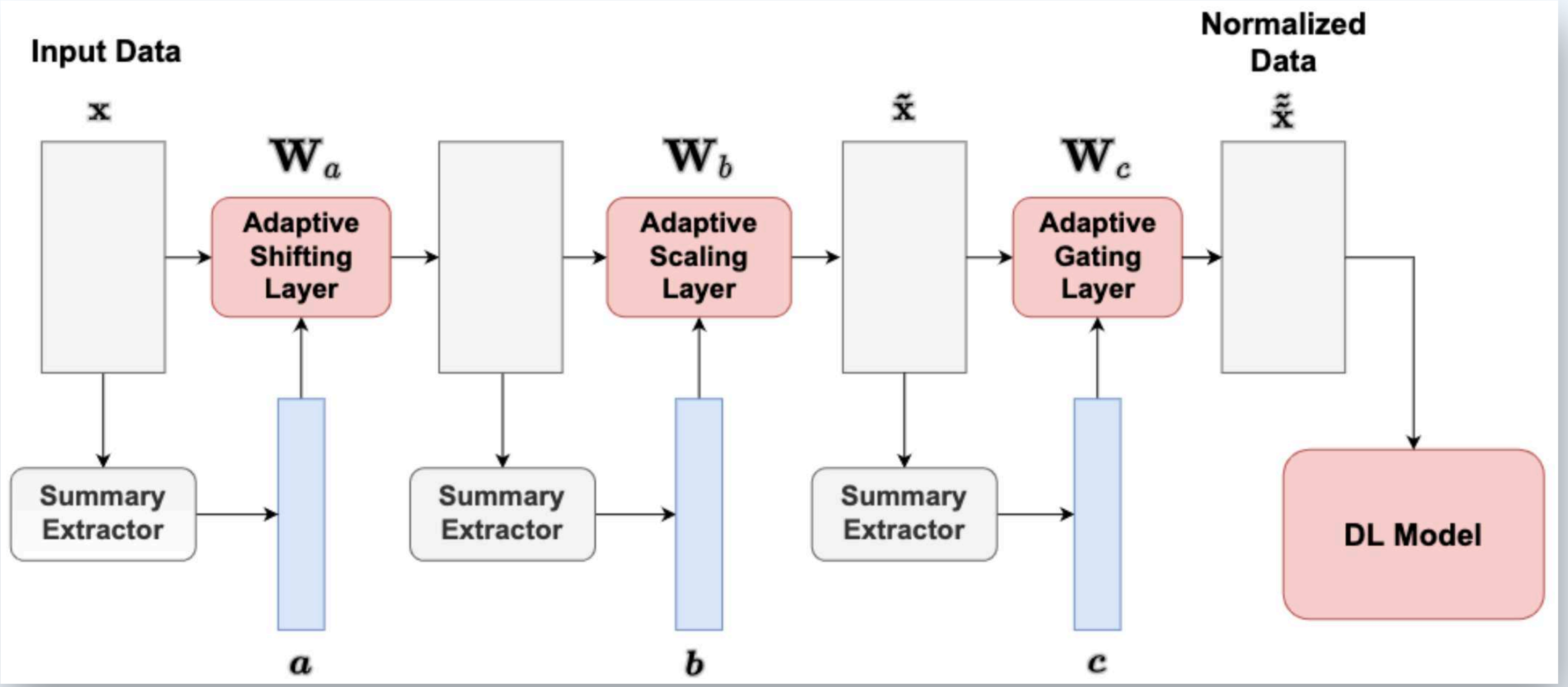
- Residual blocks with skip connections: g(**x**) = f(
- h(**x**) = **x** *or* a strided convolutional layer
- After each convolutional layer: batch nor
  ReLU activation
- Mini-batch size of 400 segments
- Adam optimizer for back propagatio
- Objective function = regularized binar

Gradient problem solved!
-> much deeper networks

1-D ResNet-54 now used by most
ML GW detection codes

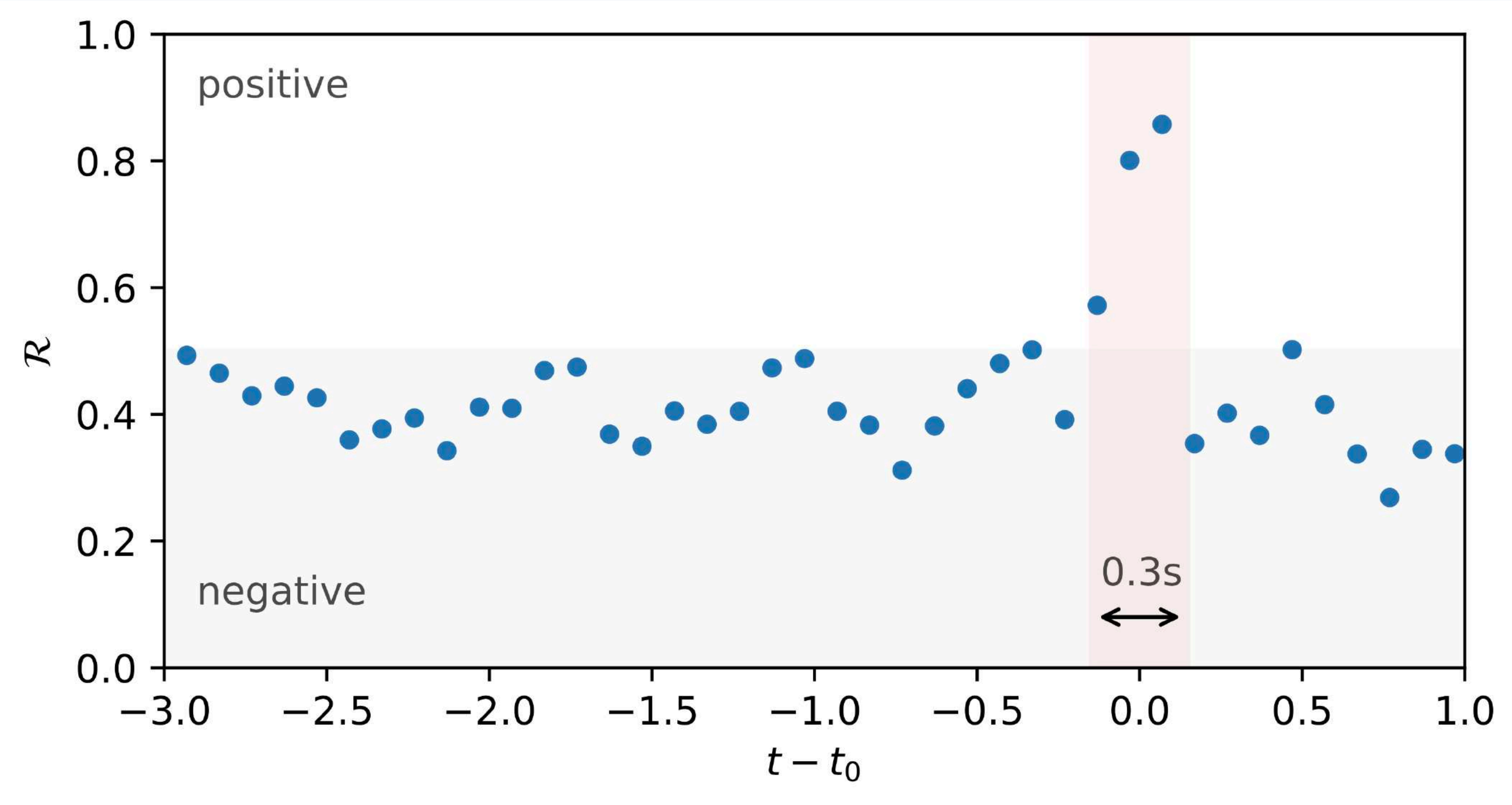
Deep Adaptive Input Normalization (DAIN) (Passalis et al. 2019)

Is applied during training - weights **W**a, **W**b, **W**c are learnable and adapt to input data!



RANKING STATISTIC

Neural Network Ranking Statistic R is reported every 0.1 s of input data

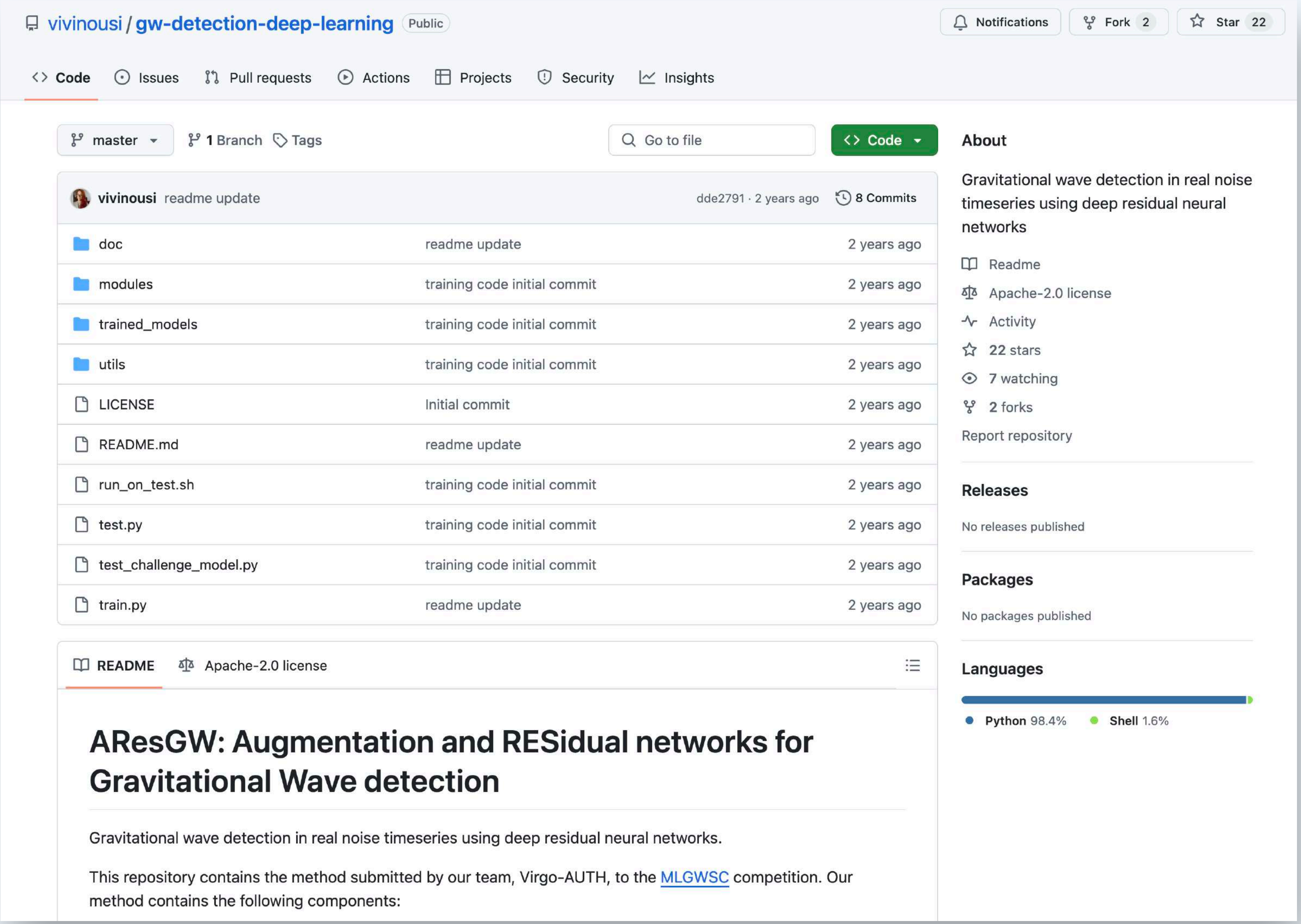


Logarithmic Ranking Statistic

$$\boxed{\mathcal{R}_s = -\log_{10}\left(1 - \mathcal{R} + 10^{-16}\right)}$$

ARESGW ON GITHUB

https://github.com/vivinousi/gw-detection-deep-learning



ENHANCED ARESGW ON REAL O3 DATA

IOP Publishing

Mach. Learn.: Sci. Technol. **6** (2025) 015054

https://doi.org/10.1088/2632-2153/adb5ed

MACHINE
LEARNING
Science and Technology

OPEN ACCESS

RECEIVED
20 September 2024

REVISED
16 January 2025

ACCEPTED FOR PUBLICATION
13 February 2025

PUBLISHED
27 February 2025

PAPER

New gravitational wave discoveries enabled by machine learning

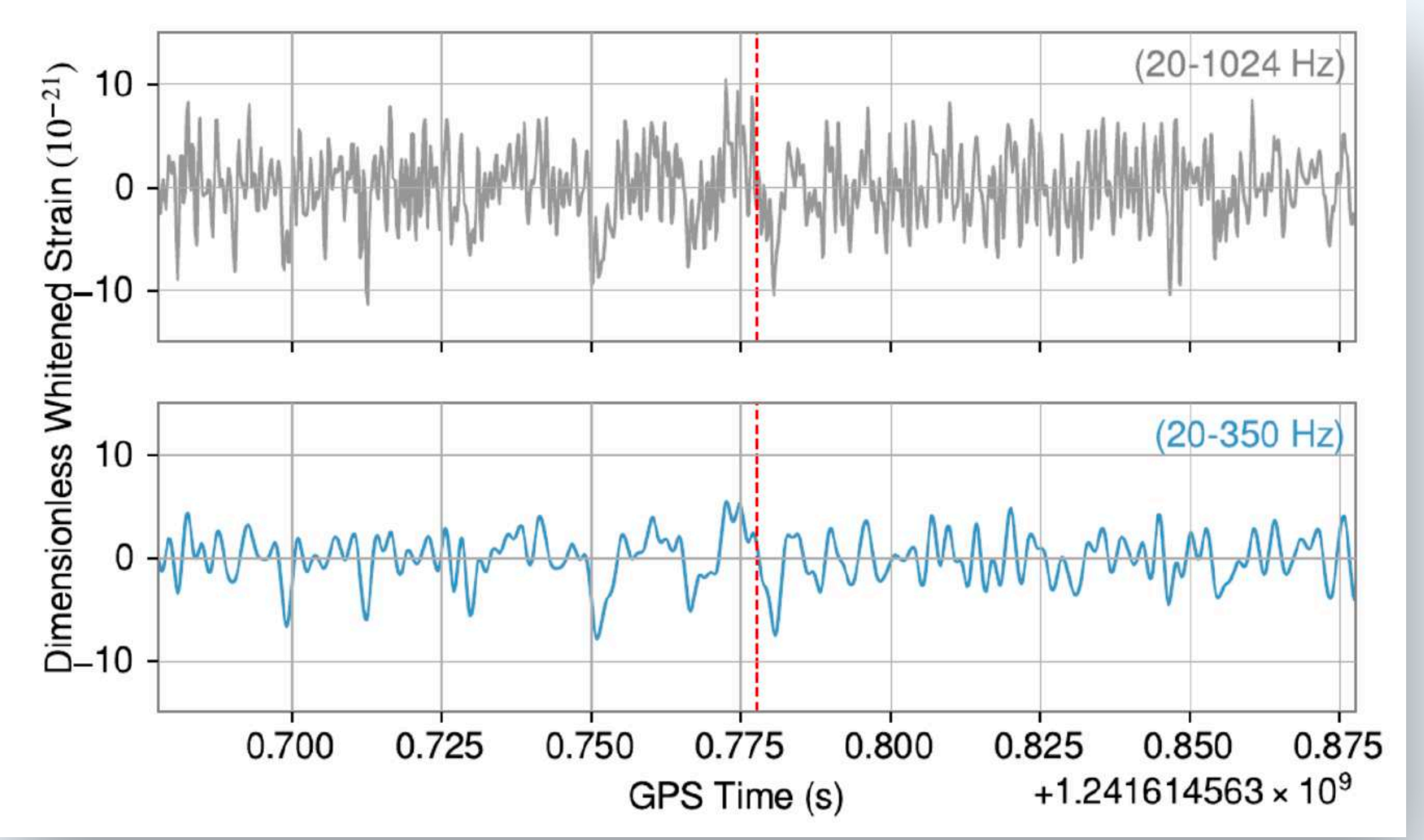
Alexandra E Koloniari[1,*], Evdokia C Koursoumpa[1], Paraskevi Nousi[2], Paraskevas Lampropoulos[1], Nikolaos Passalis[3], Anastasios Tefas[4] and Nikolaos Stergioulas[1]

[1] Department of Physics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
[2] Swiss Data Science Center, ETH, Zürich, Switzerland
[3] Department of Chemical Engineering, Faculty of Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
[4] Department of Informatics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
[*] Author to whom any correspondence should be addressed.

E-mail: akolonia@auth.gr

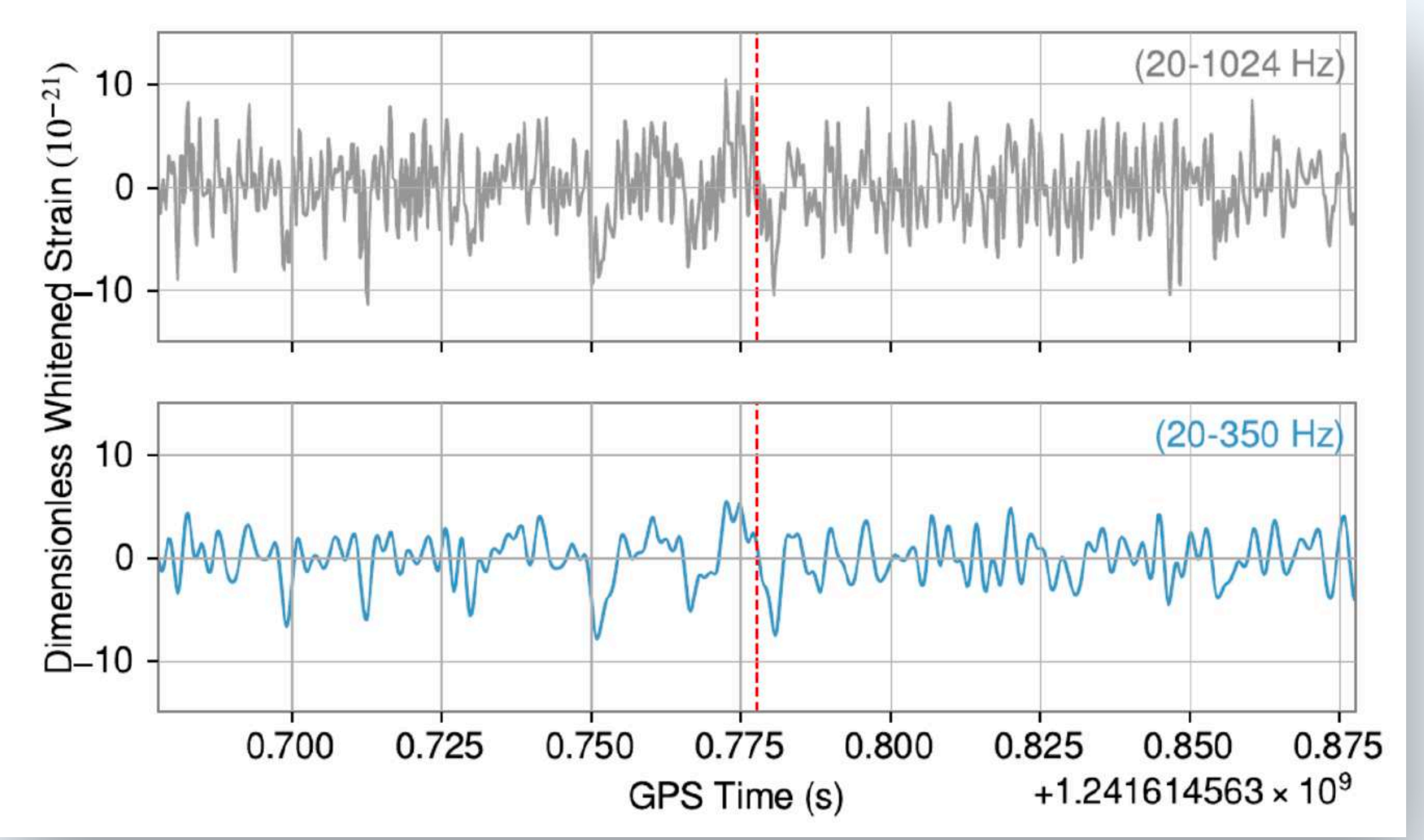
MAIN ENHANCEMENTS:

1) **Training** on **low-pass filtered** data

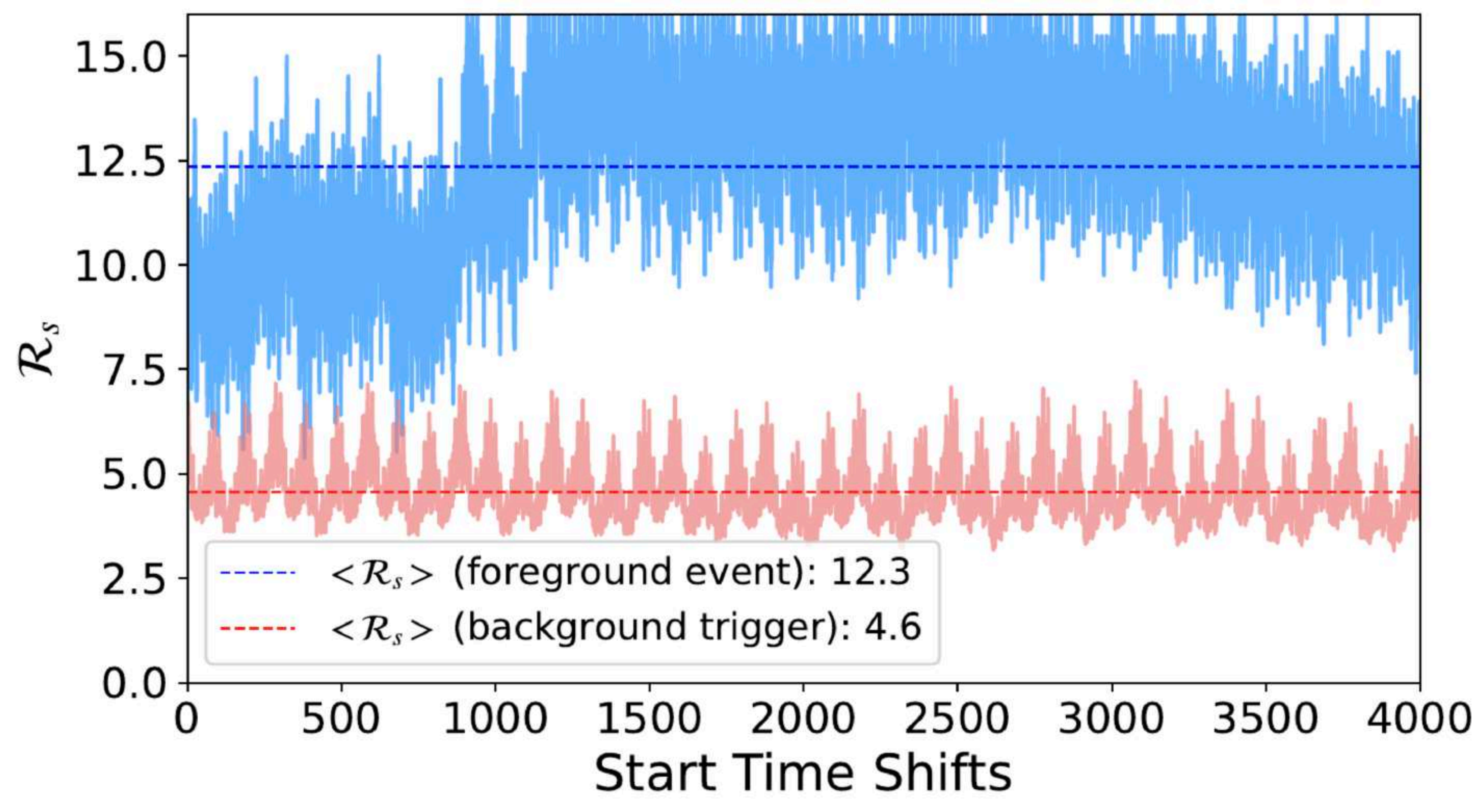


MAIN ENHANCEMENTS:

1) **Training** on **low-pass filtered** data

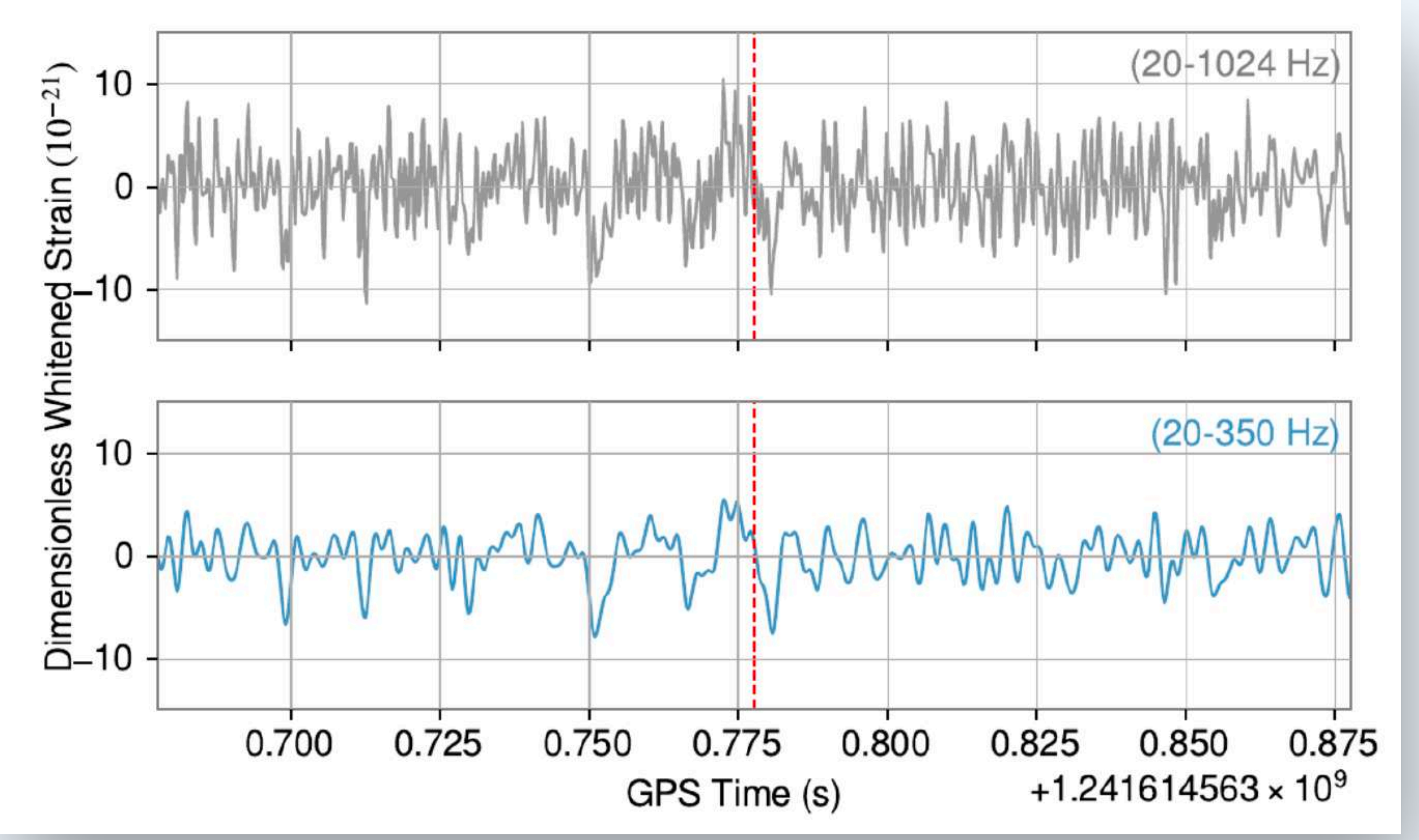


2) **Ensemble-averaged** ranking statistic

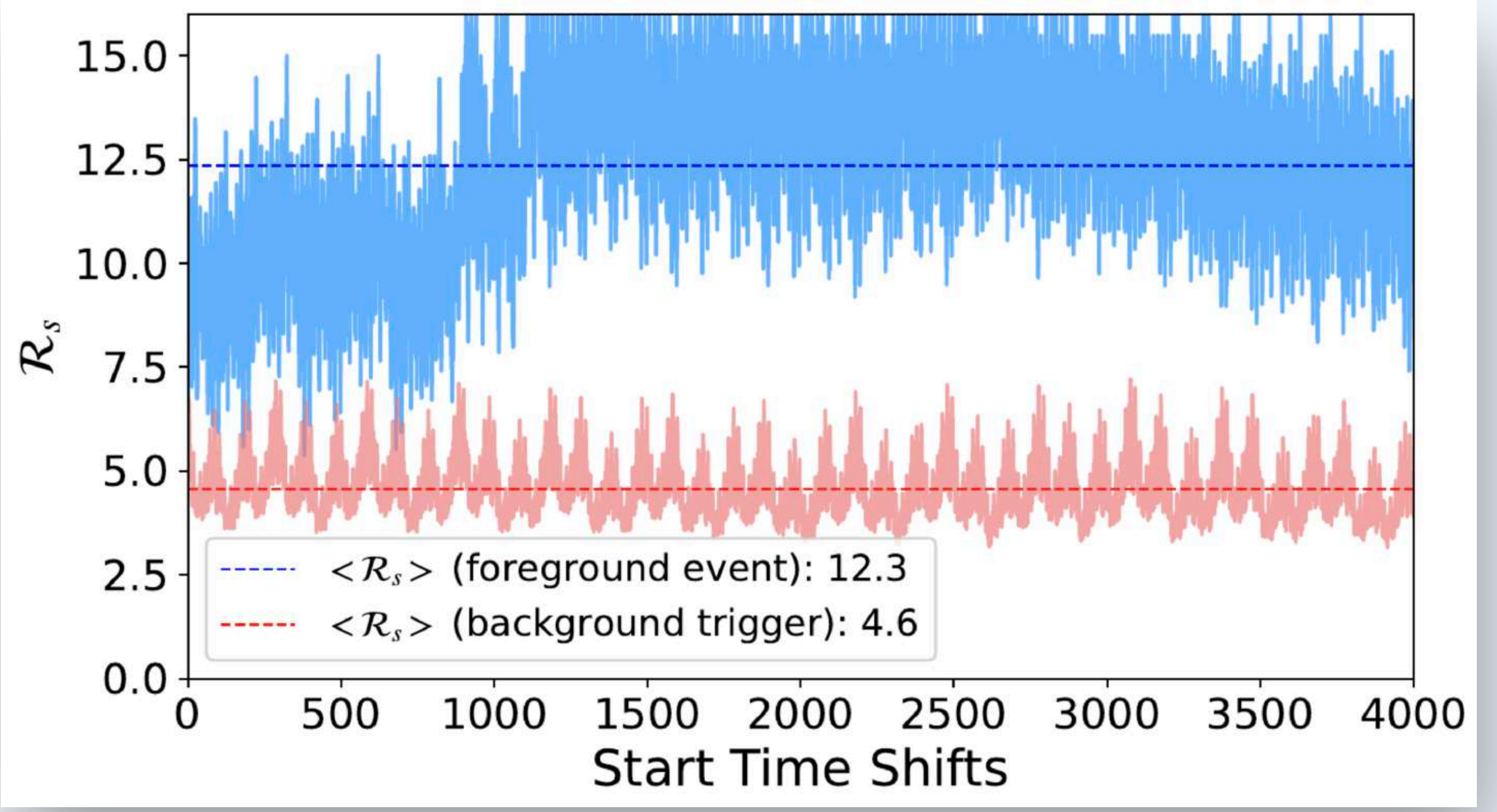


MAIN ENHANCEMENTS:

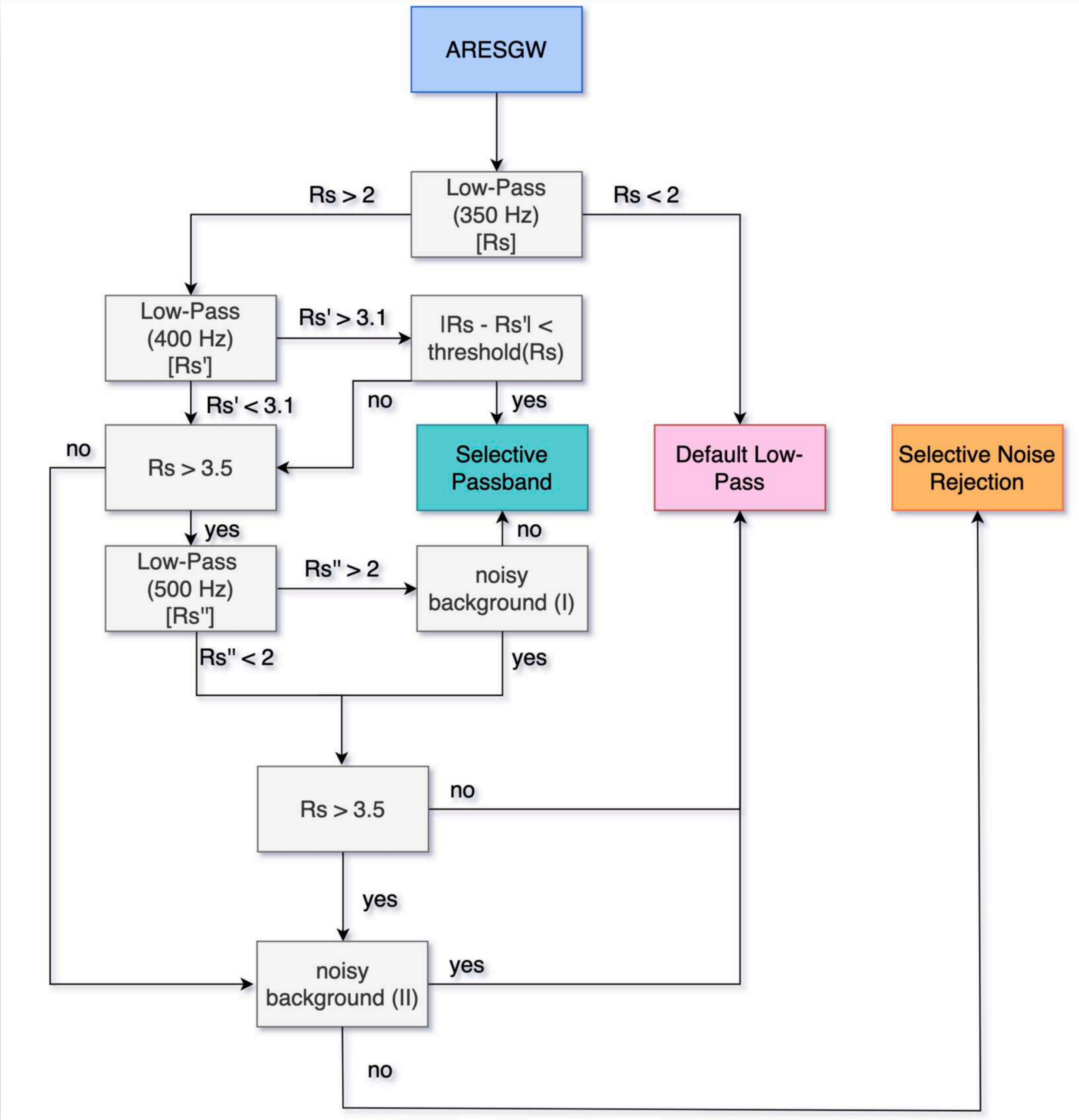
1) **Training** on **low-pass filtered** data



2) **Ensemble-averaged** ranking statistic



3) Application of **noise filters** to **reduce background FAR**.



ASTROPHYSICAL PROBABILITY

p_astro is calculated as

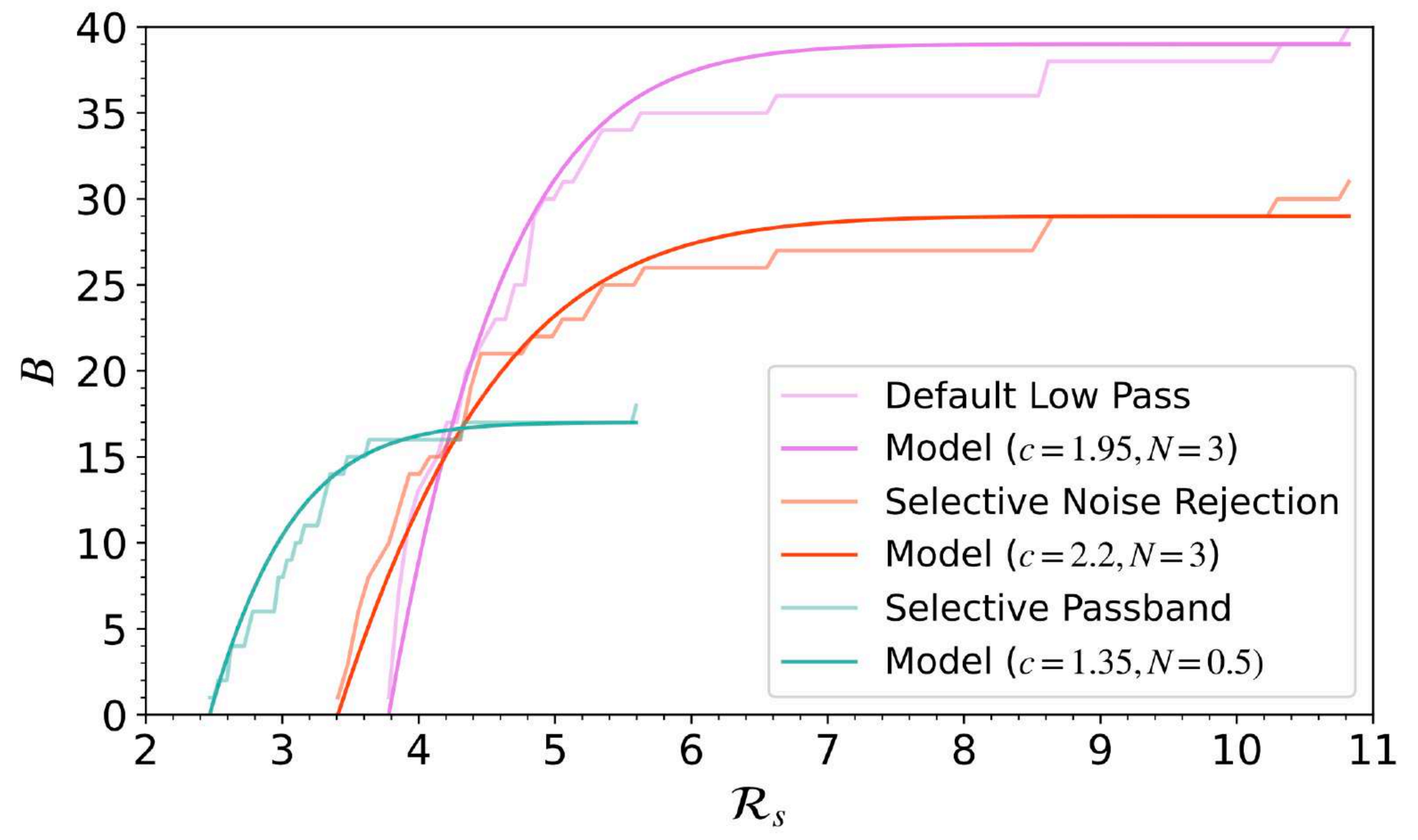
$$p_{\rm astro} = \frac{f(\mathcal{R}_s)}{b(\mathcal{R}_s) + f(\mathcal{R}_s)}$$

where the **background** and **foreground differential rates** are

$$b(\mathcal{R}_s) = \frac{dB}{d\mathcal{R}_s} \qquad f(\mathcal{R}_s) = \frac{dF}{d\mathcal{R}_s}$$

The cumulative distribution of the **O3 background** is modeled analytically as

$$\hat{B}(x) = \frac{\left(1+\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right)^N - \left(1+\mathrm{erf}\left(\frac{x_{\min}}{\sqrt{2}}\right)\right)^N}{2^N - \left(1+\mathrm{erf}\left(\frac{x_{\min}}{\sqrt{2}}\right)\right)^N}$$

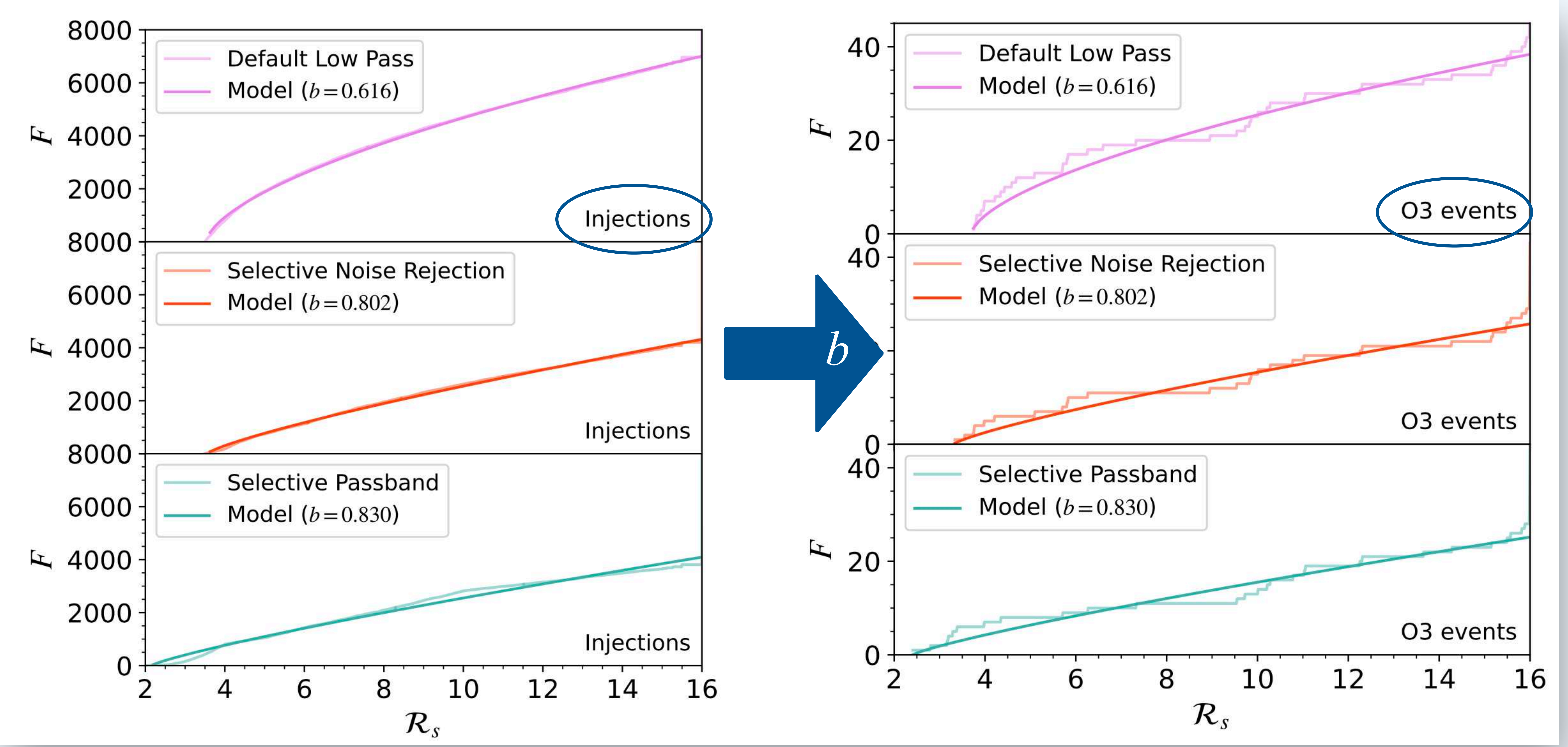


ASTROPHYSICAL PROBABILITY

Analytic model of cumulative **foreground** distribution:

$$F(x) = a(x - x_{\min})^b$$

Coefficient b determined through injections in O3 noise:



NEW GW DETECTIONS WITH AresGW IN O3 DATA

AresGW detects **42 out of 51** known O3 events (*most sensitive pipeline in this mass range*)

We found **8 new gravitational wave candidates** with p_{astro} > 50% (3 p_{astro} >99%).

TABLE VII: New candidate events identified by AresGW.

#	Event Name	GPS Time (s)	p_{astro}	FAR (1/yr)	$\langle \mathcal{R}_s \rangle$	Time delay (s)	χ^2_L	χ^2_H	Class
1	GW190511_125545	1241614563.77	1.00	0.27	9.54	0.0027	1.16	1.46	Selective Passband
2	GW190614_134749	1244555287.93	0.99	4.6	5.80	0.0012	0.65	0.80	Selective Passband
3	GW190607_083827	1243931925.99	0.99	6.5	8.95	0.0056	0.90	0.48	Selective Noise Rejection
4	GW190904_104631	1251629209.01	0.72	14	4.35	0.0002	0.38	0.71	Selective Passband
5	GW190523_085933	1242637191.44	0.68	20	6.60	0.0054	0.75	1.39	Selective Noise Rejection
6	GW200208_211609	1265231787.68	0.55	18	4.0	0.0063	0.69	0.98	Selective Passband
7	GW190705_164632	1246380410.88	0.51	49	5.82	0.0103	1.05	0.98	Default Low-Pass*
8	GW190426_082124	1240302101.93	0.50	20	3.91	0.0007	1.48	0.53	Selective Passband

NEW GW DETECTIONS WITH AresGW IN O3 DATA

AresGW detects **42 out of 51** known O3 events (*most sensitive pipeline in this mass range*)

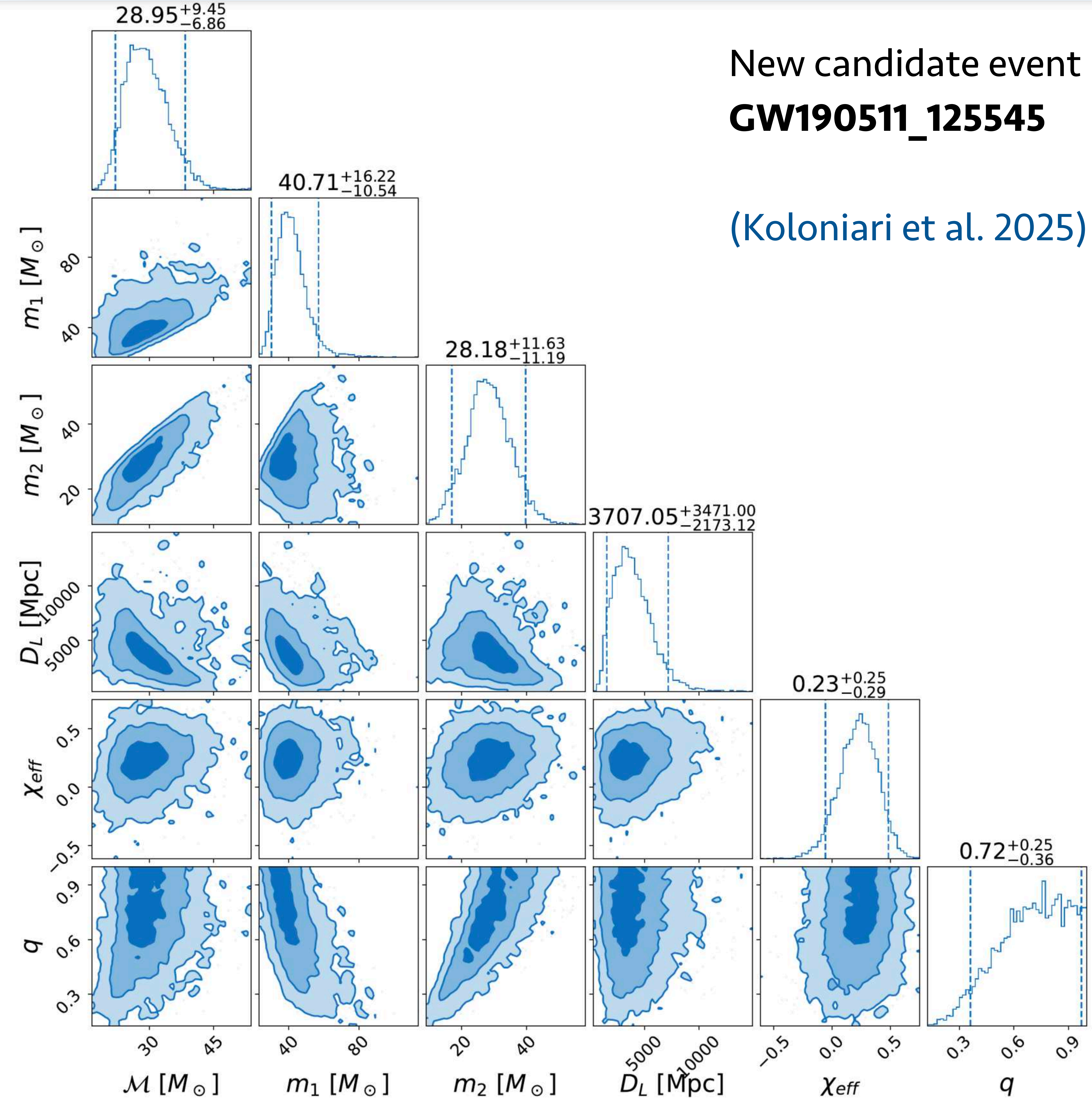
We found **8 new gravitational wave candidates** with p_{astro} > 50% (3 p_{astro} >99%).

TABLE VII: New candidate events identified by AresGW.

#	Event Name	GPS Time (s)	p_{astro}	FAR (1/yr)	$\langle \mathcal{R}_s \rangle$	Time delay	χ^2	[illegible]	Class
1	GW190511_125545	1241614563.77	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Selective Passband
2	GW190614_1347[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.80	Selective Passband
3	GW190607_08[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0056	0.90	0.48	Selective Noise Rejection
4	GW190904_104[illegible]	[illegible]	0.72	14	4.35	0.0002	0.38	0.71	Selective Passband
5	GW190523_085[illegible]	1242637191.44	0.68	20	6.60	0.0054	0.75	1.39	Selective Noise Rejection
6	GW200208_211609	1265231787.68	0.55	18	4.0	0.0063	0.69	0.98	Selective Passband
7	GW190705_164632	1246380410.88	0.51	49	5.82	0.0103	1.05	0.98	Default Low-Pass*
8	GW190426_082124	1240302101.93	0.50	20	3.91	0.0007	1.48	0.53	Selective Passband

First GW events identified using a machine-learning algorithm

PARAMETER ESTIMATION USING BILBY

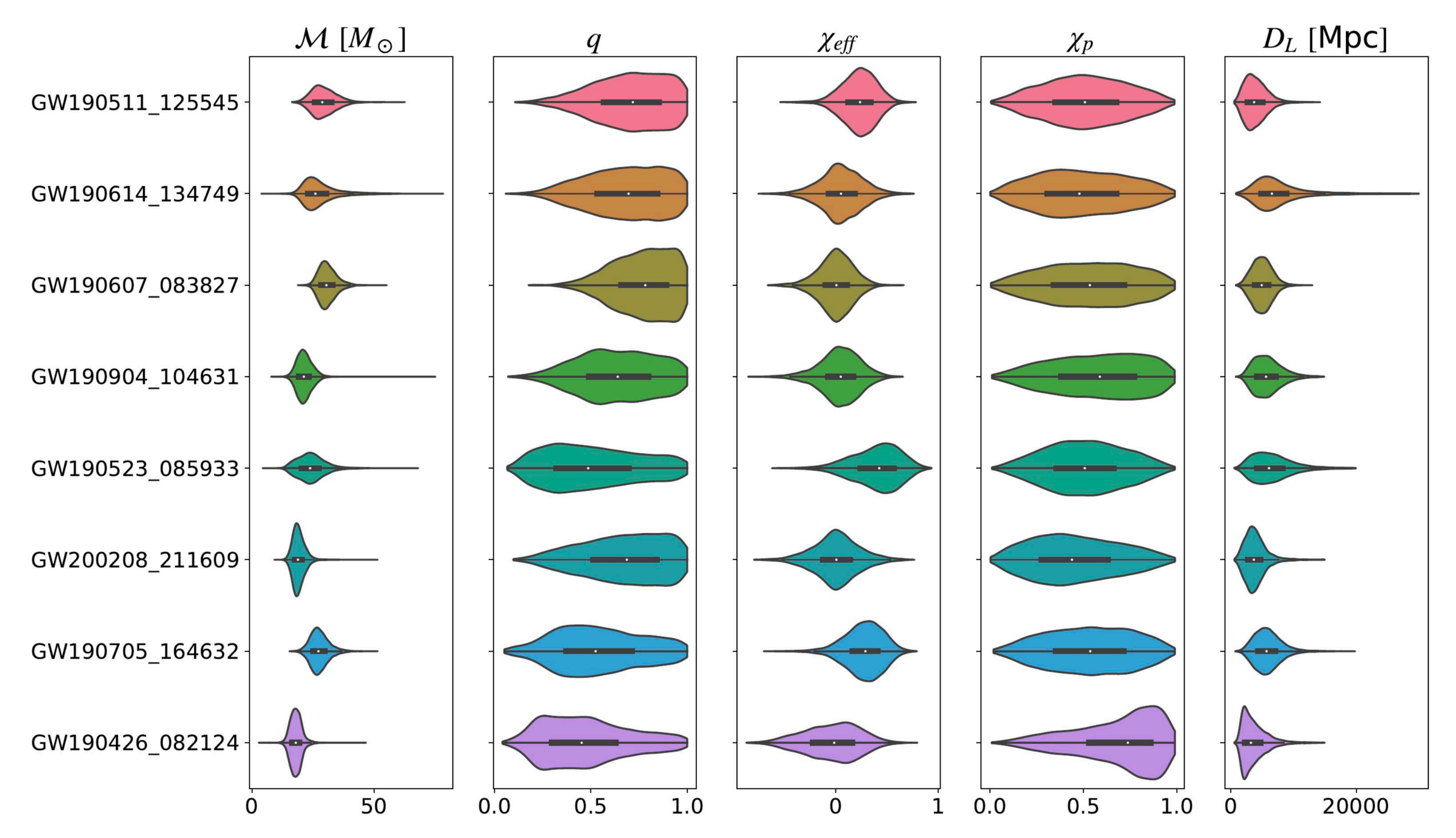


New candidate event

GW190511_125545

(Koloniari et al. 2025)

PARAMETER ESTIMATION USING BILBY



RECONSTRUCTED WAVEFORMS FOR NEW EVENTS

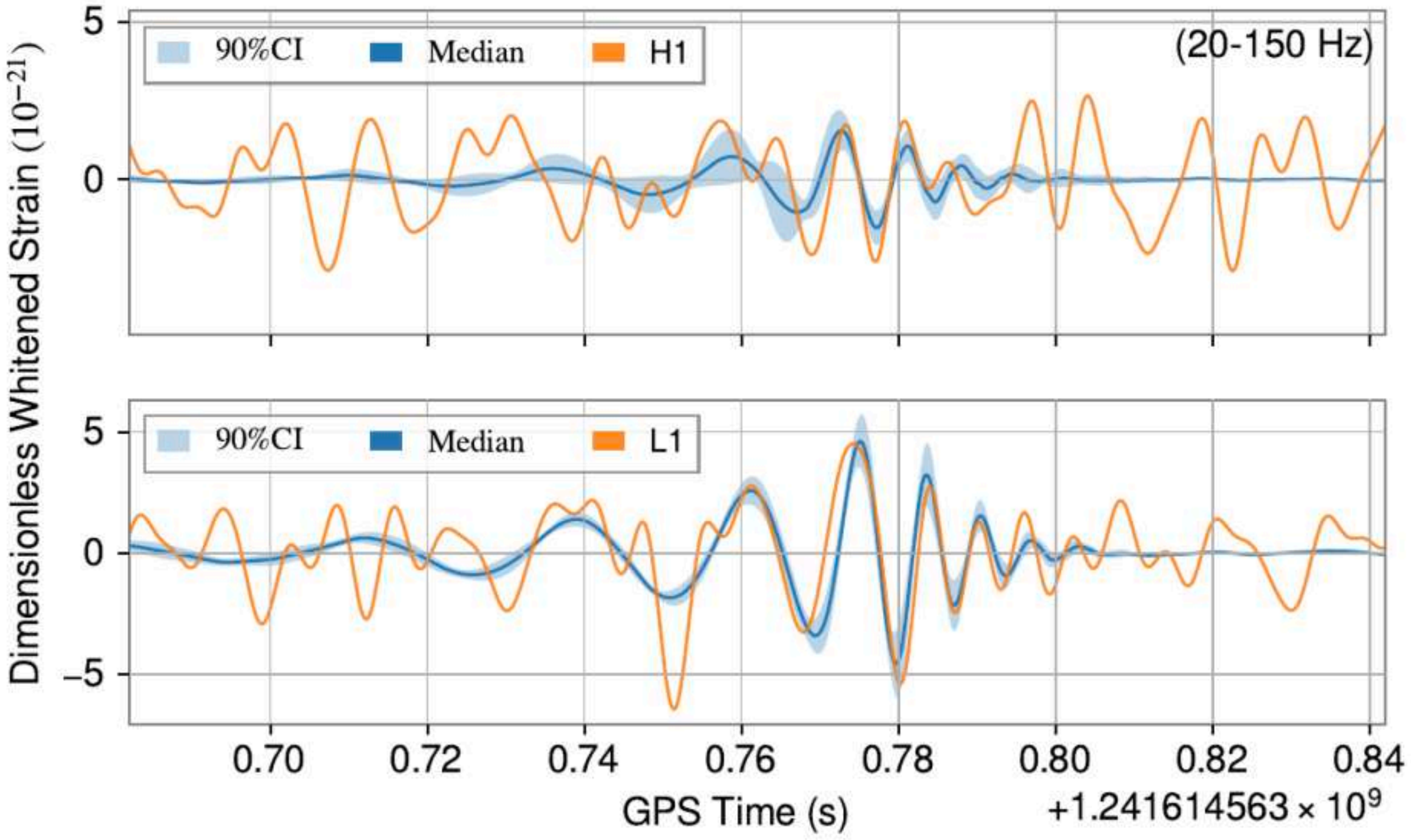


FIG. 35: Whitened, bandpassed strain data and reconstructed waveform for the new event GW190511_125545 identified by AresGW.

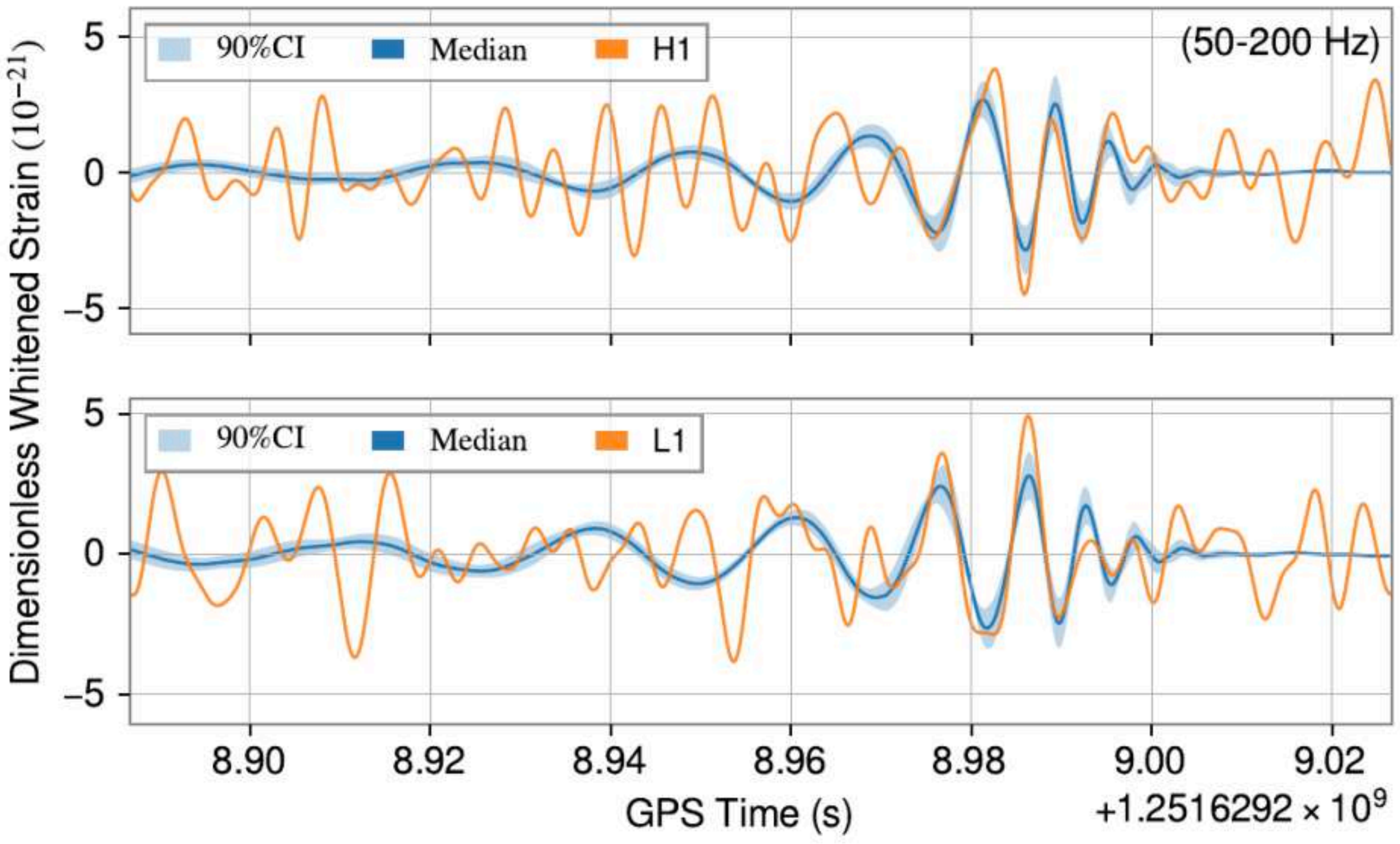
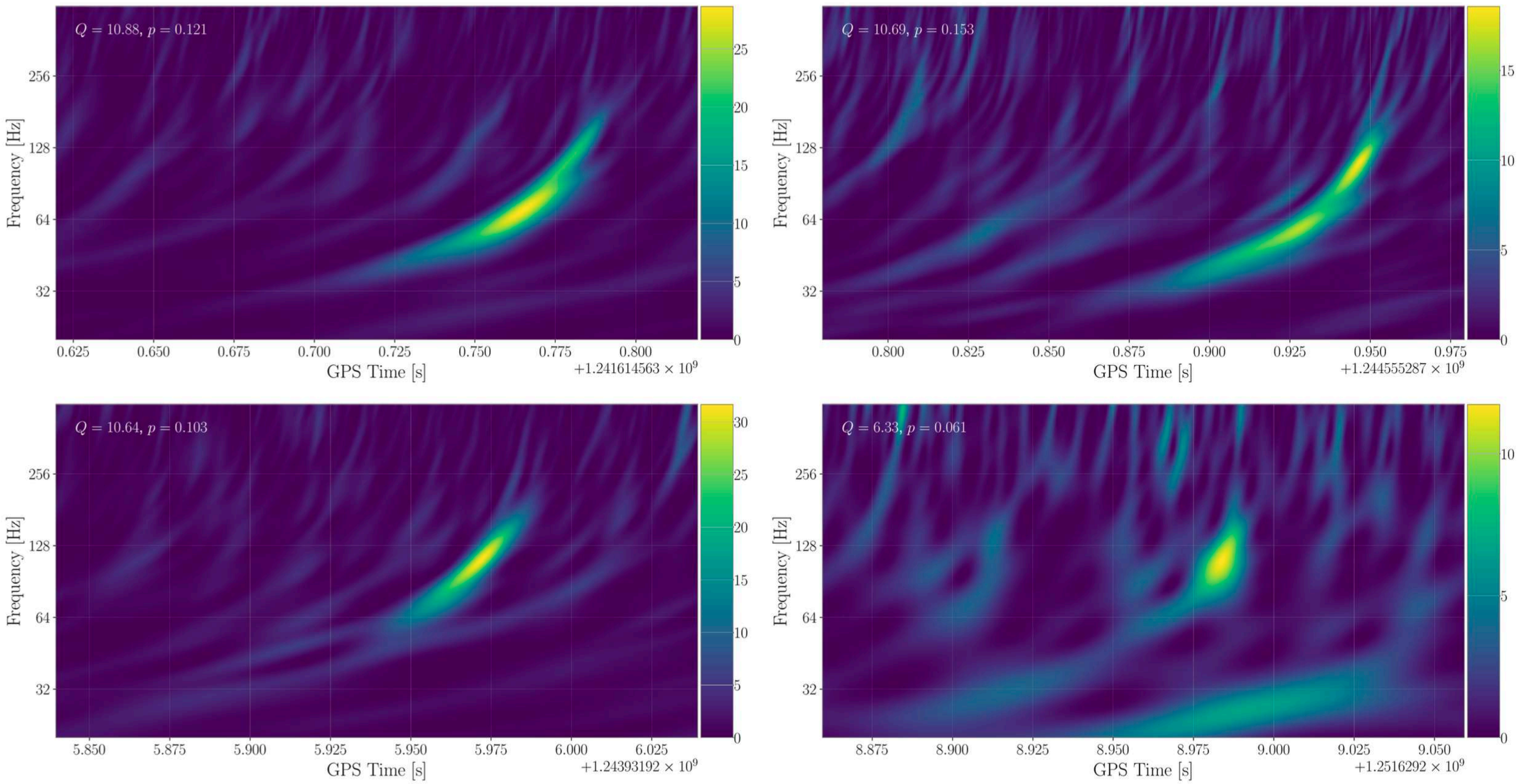


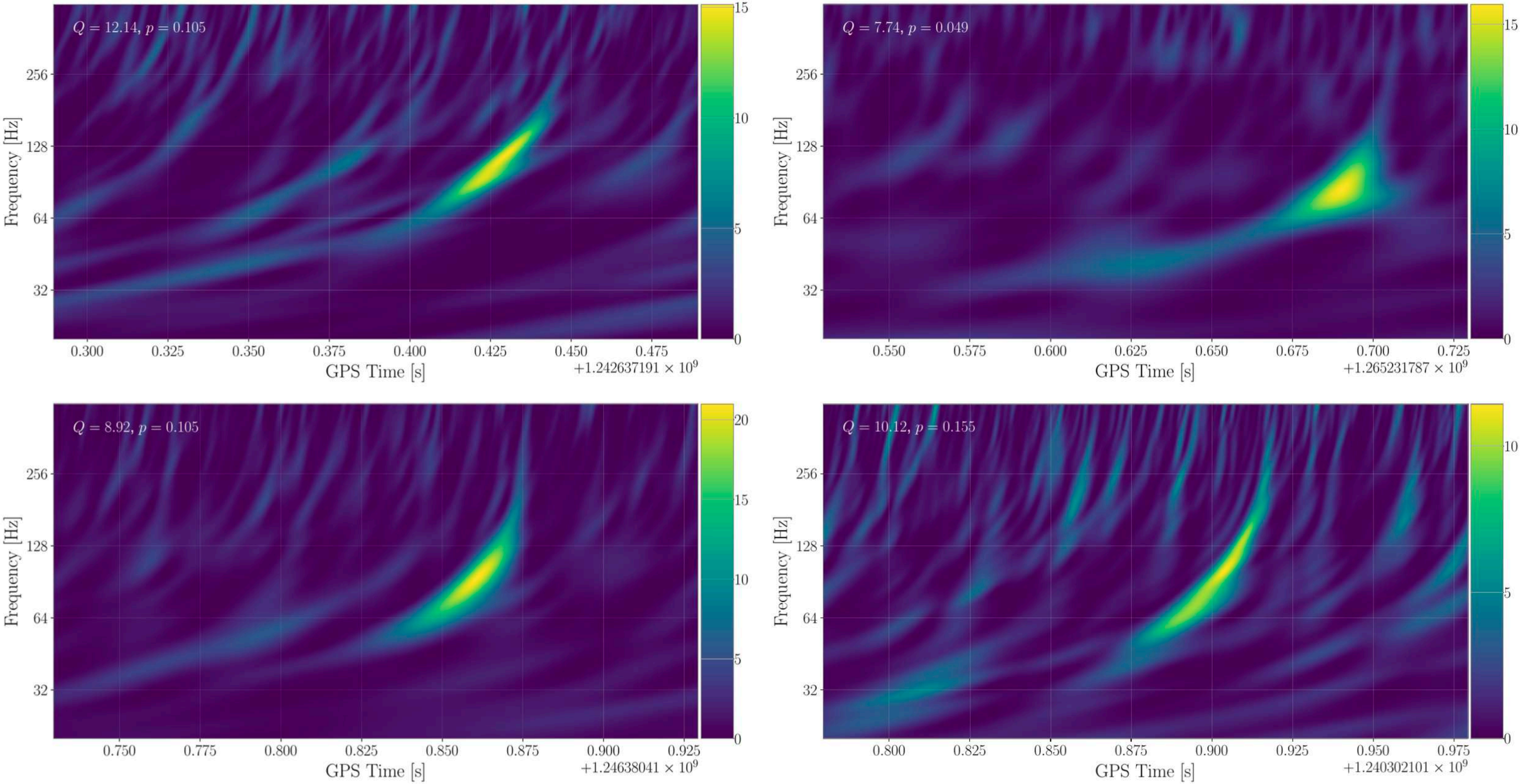
FIG. 38: Same as Fig. 35, but for the new event GW190904_104631.

Koloniari, Koursoumpa, Nousi, Lampropoulos, Passalis, Tefas, Stergioulas (2025)

ARESGW NEW CANDIDATE EVENTS

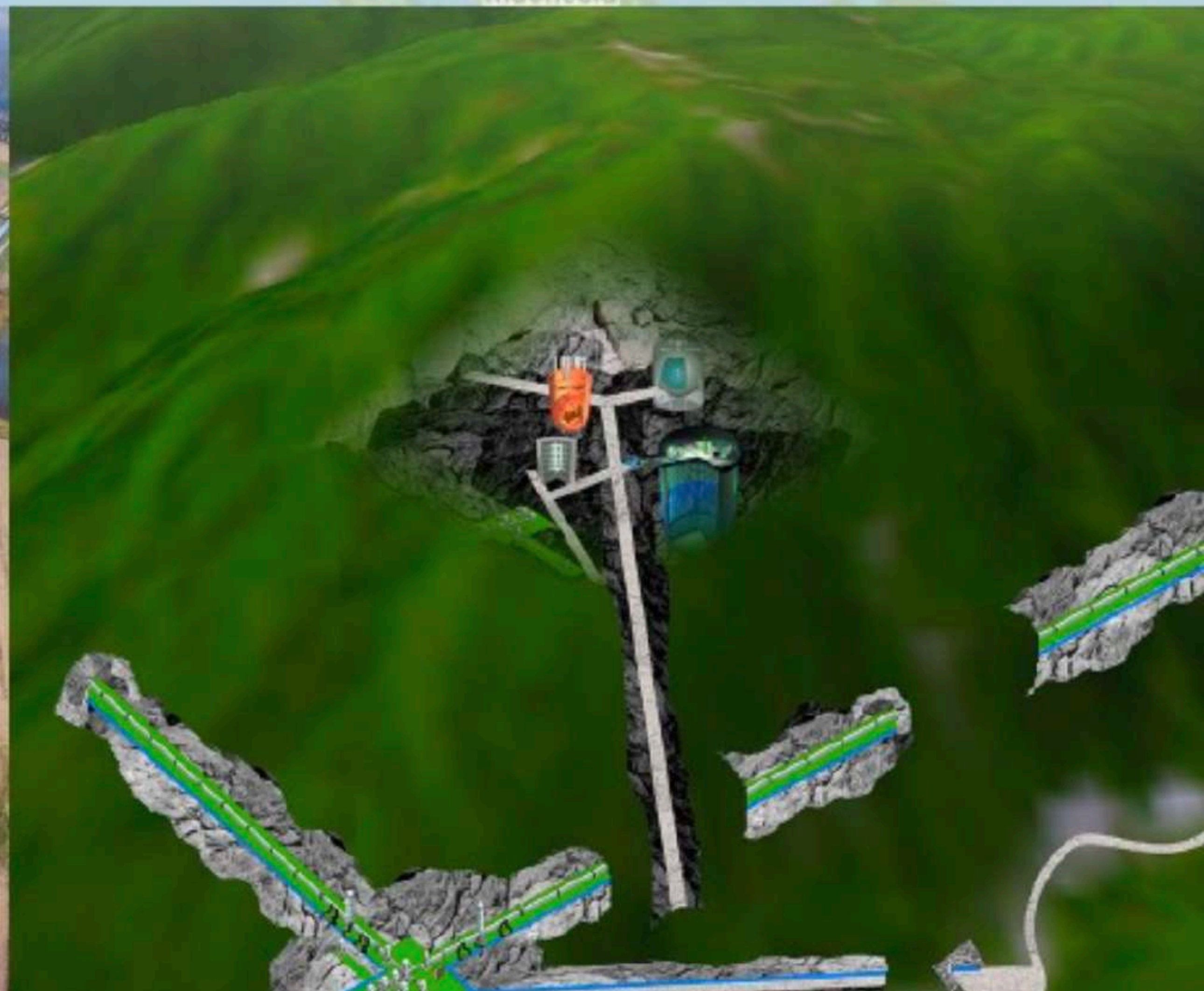


ARESGW NEW CANDIDATE EVENTS

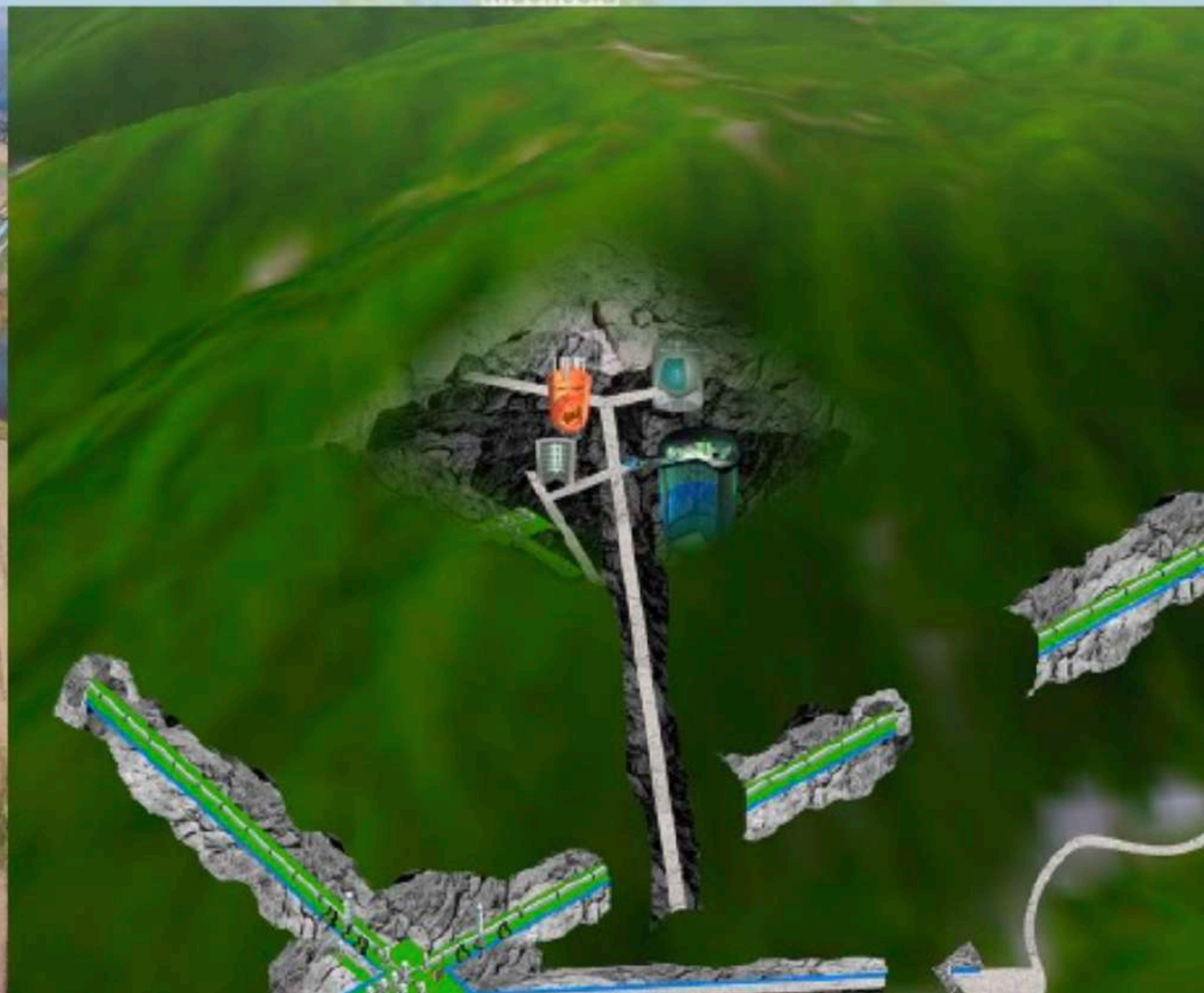


PART B.
BINARY NEUTRON STAR POST-MERGER PHASE

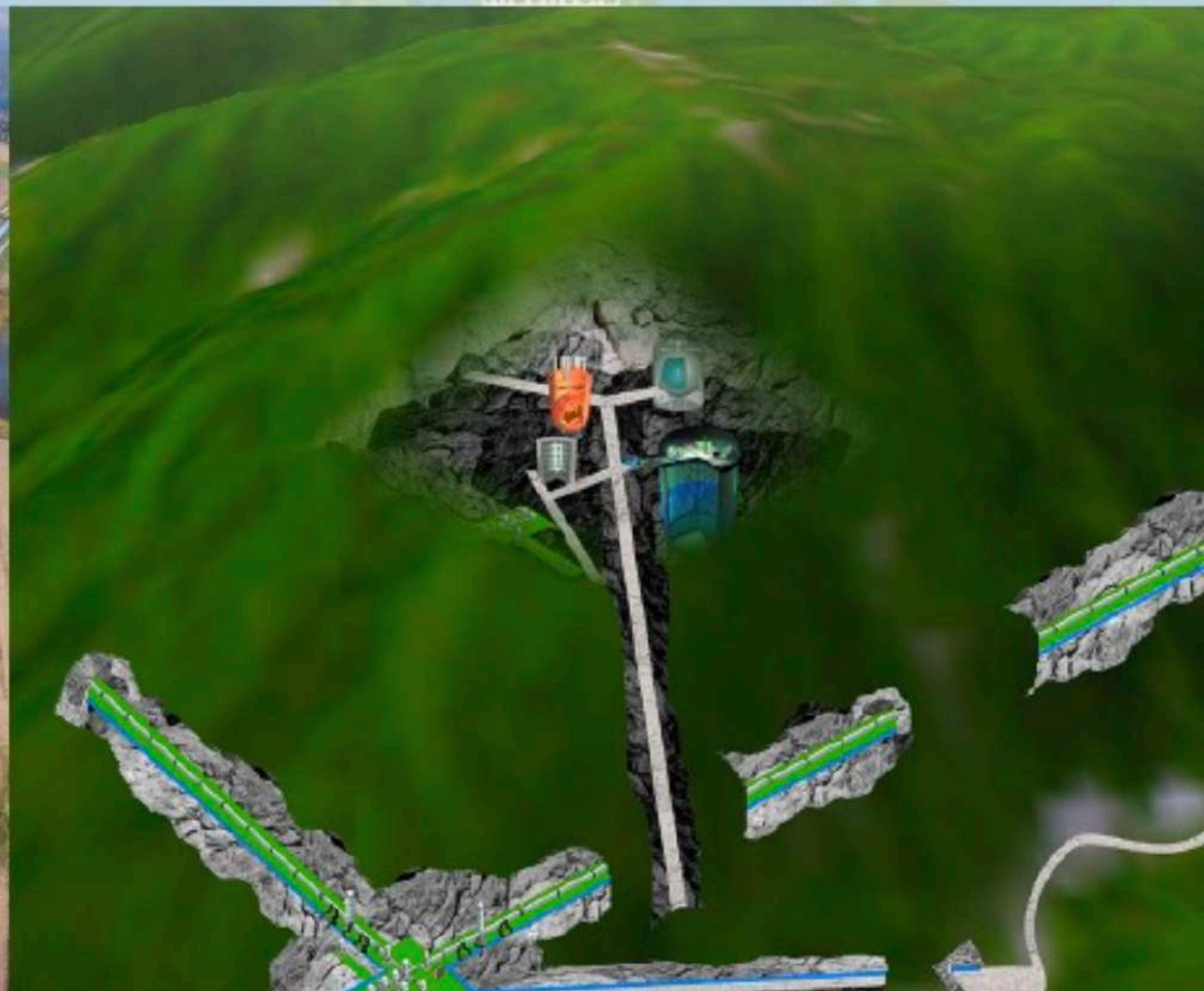
INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



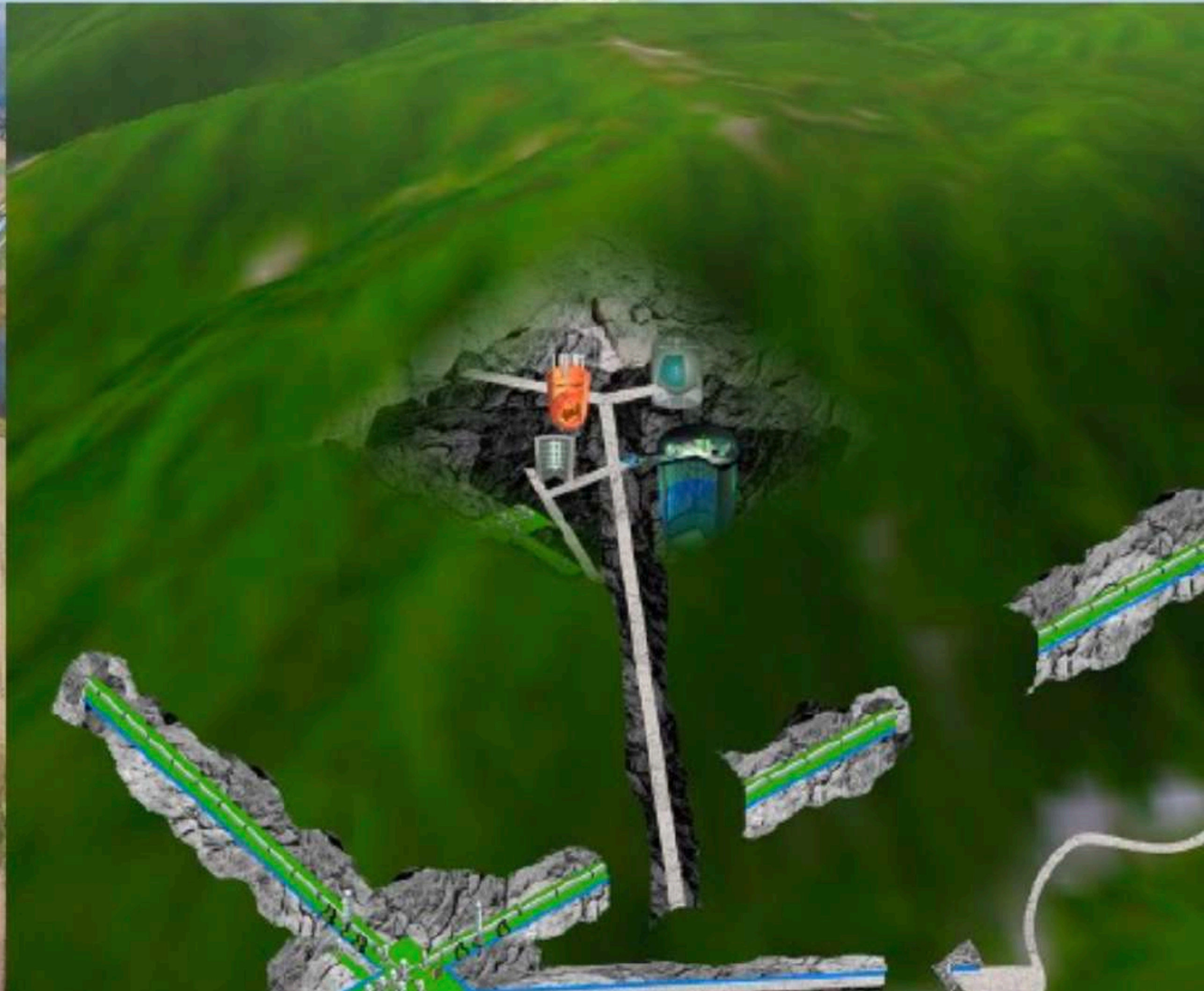
INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



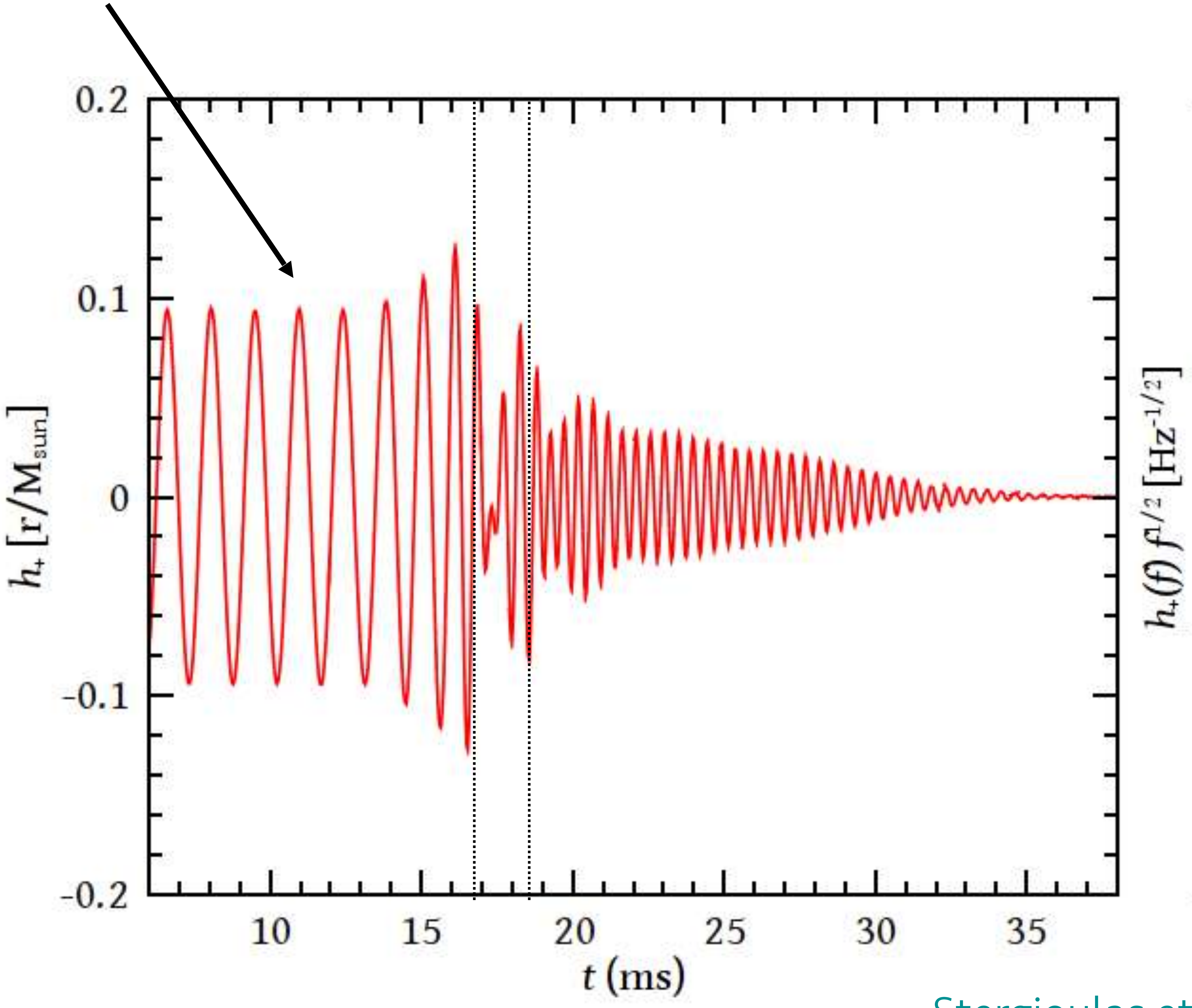
INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



<u>Time domain</u>: three distinct phases of the GW signal:
inspiral, *merger* and *post-merger oscillations*.

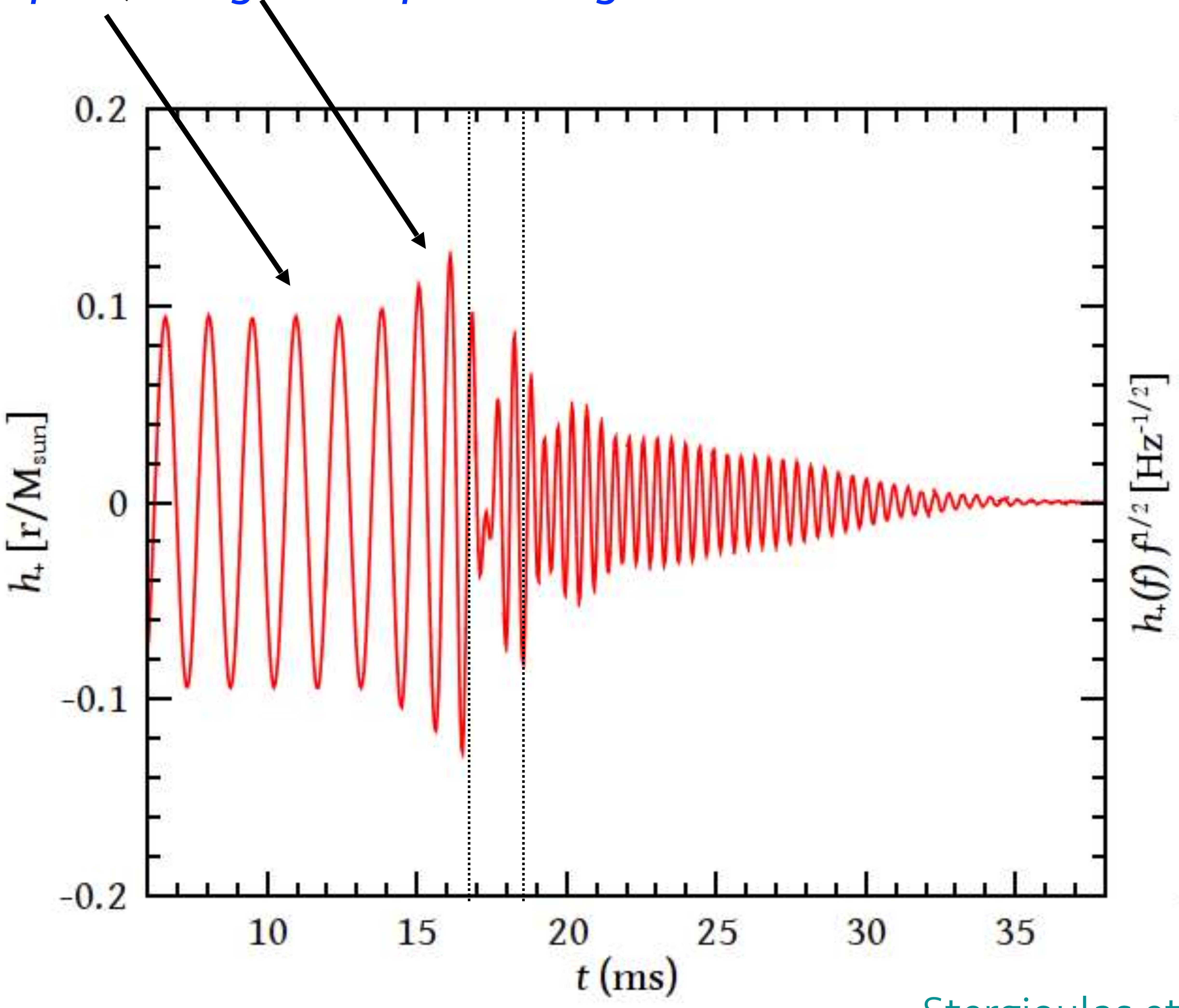


Stergioulas et al. (2011)

POST-MERGER PHASE IN BNS MERGERS

Time domain: three distinct phases of the GW signal:

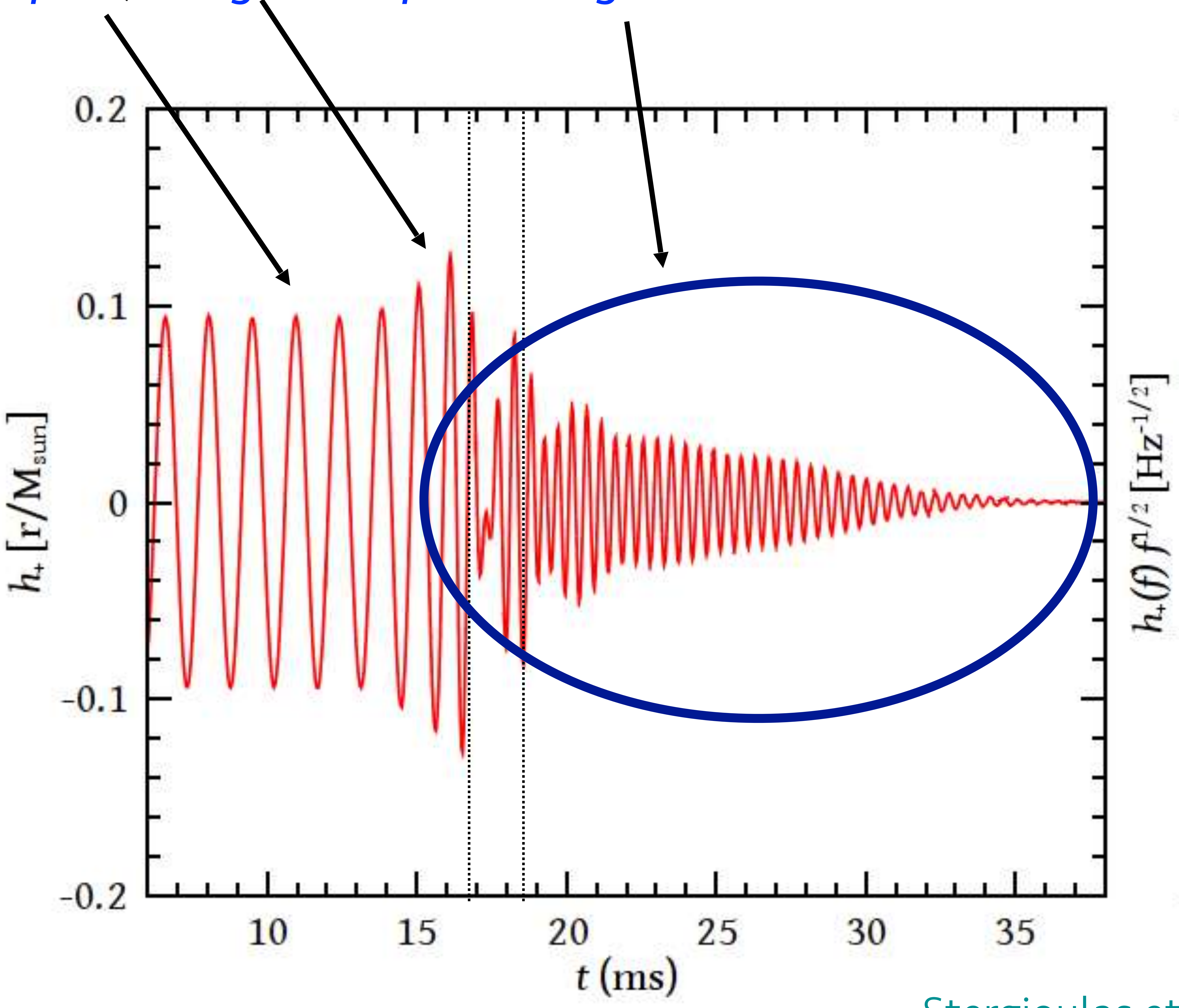
inspiral, *merger* and *post-merger oscillations*.



Stergioulas et al. (2011)

Time domain: three distinct phases of the GW signal:

inspiral, *merger* and *post-merger oscillations*.

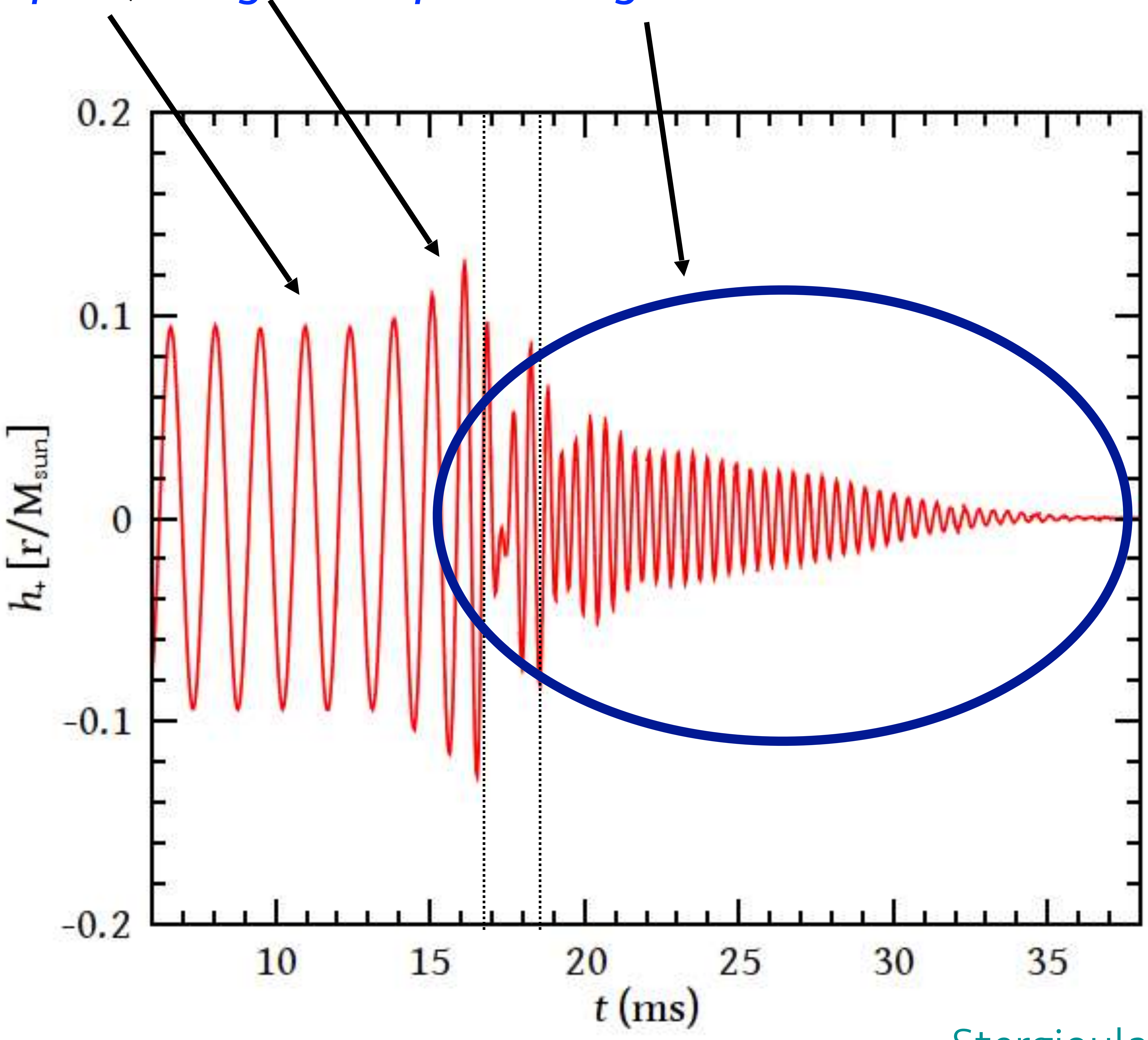


Stergioulas et al. (2011)

POST-MERGER PHASE IN BNS MERGERS

Time domain: three distinct phases of the GW signal:

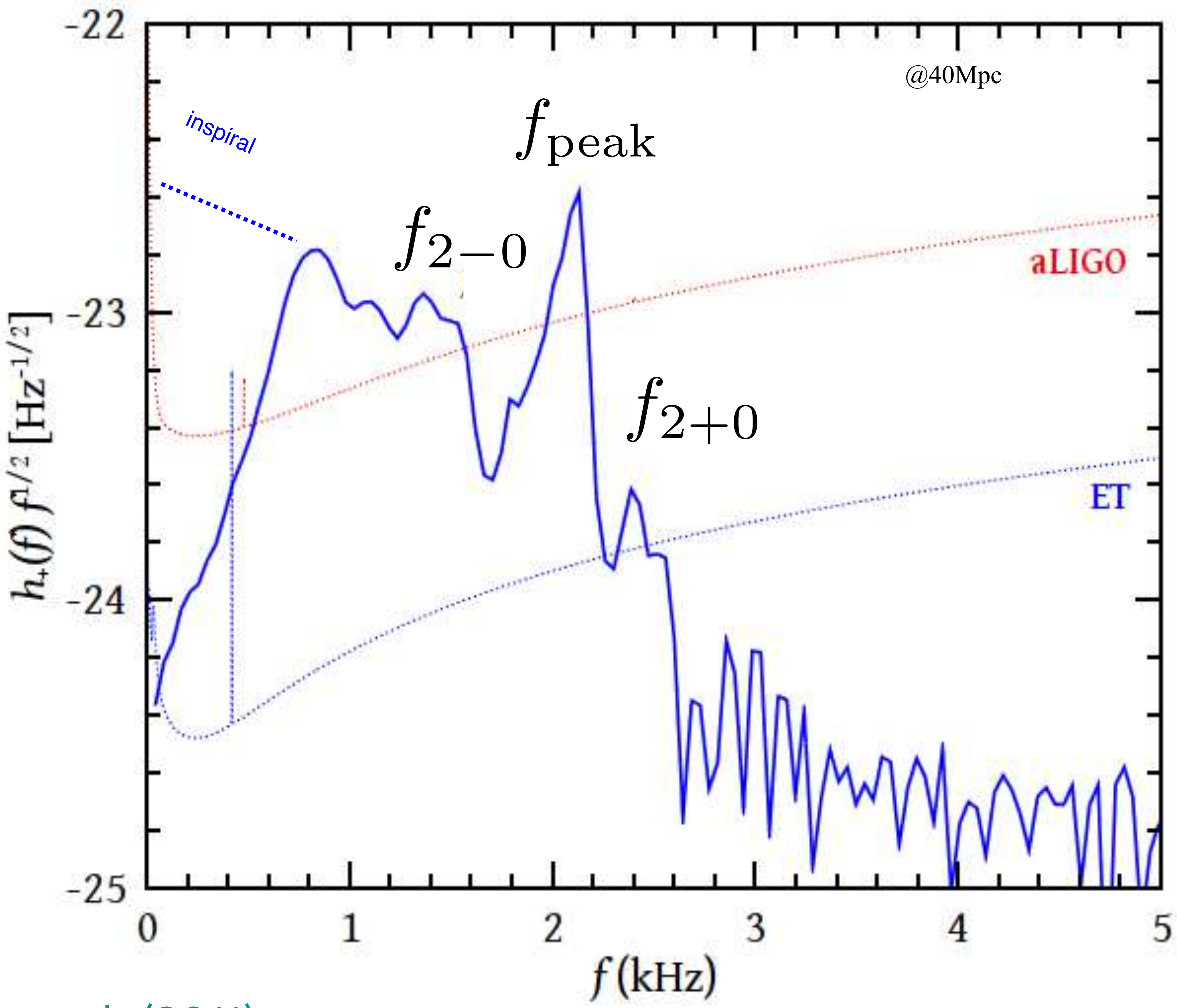
inspiral, *merger* and *post-merger oscillations*.



Frequency domain:

f_{peak} : *l=m=2 fundamental mode*.

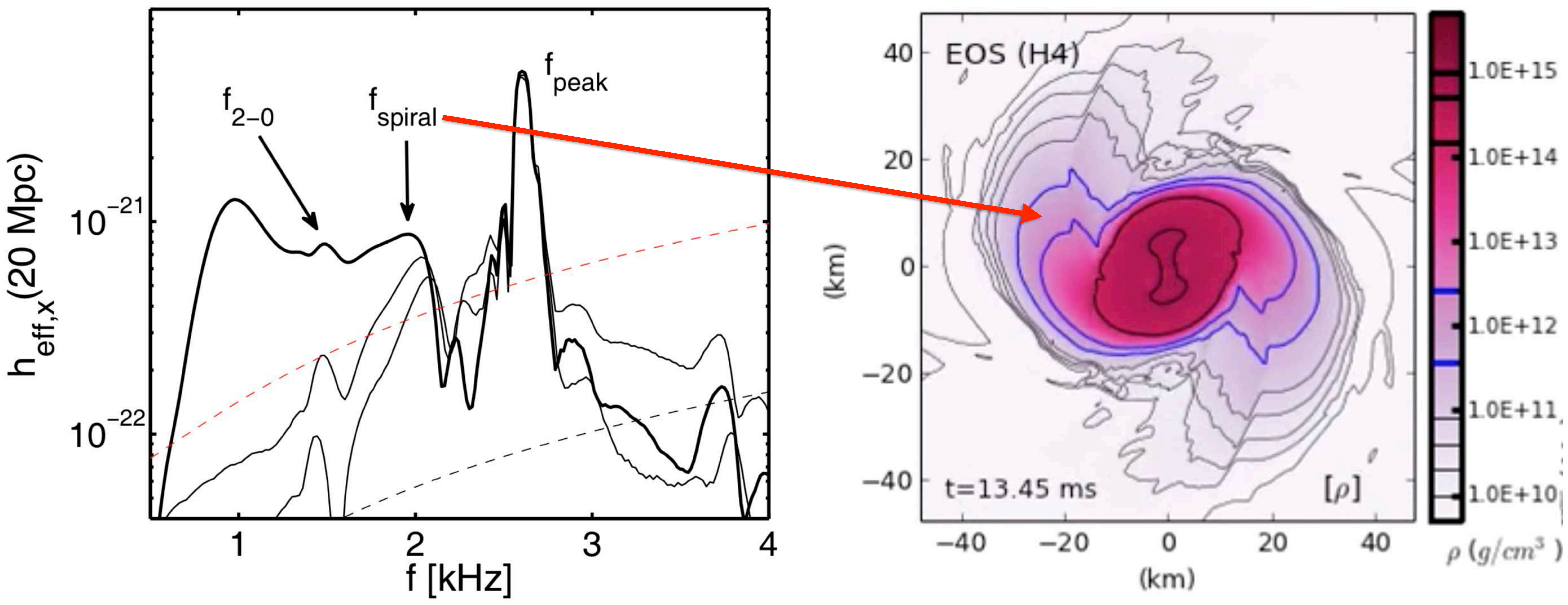
$f_{2\text{-}0}, f_{2+0}$: *nonlinear combination tones*



Stergioulas et al. (2011)

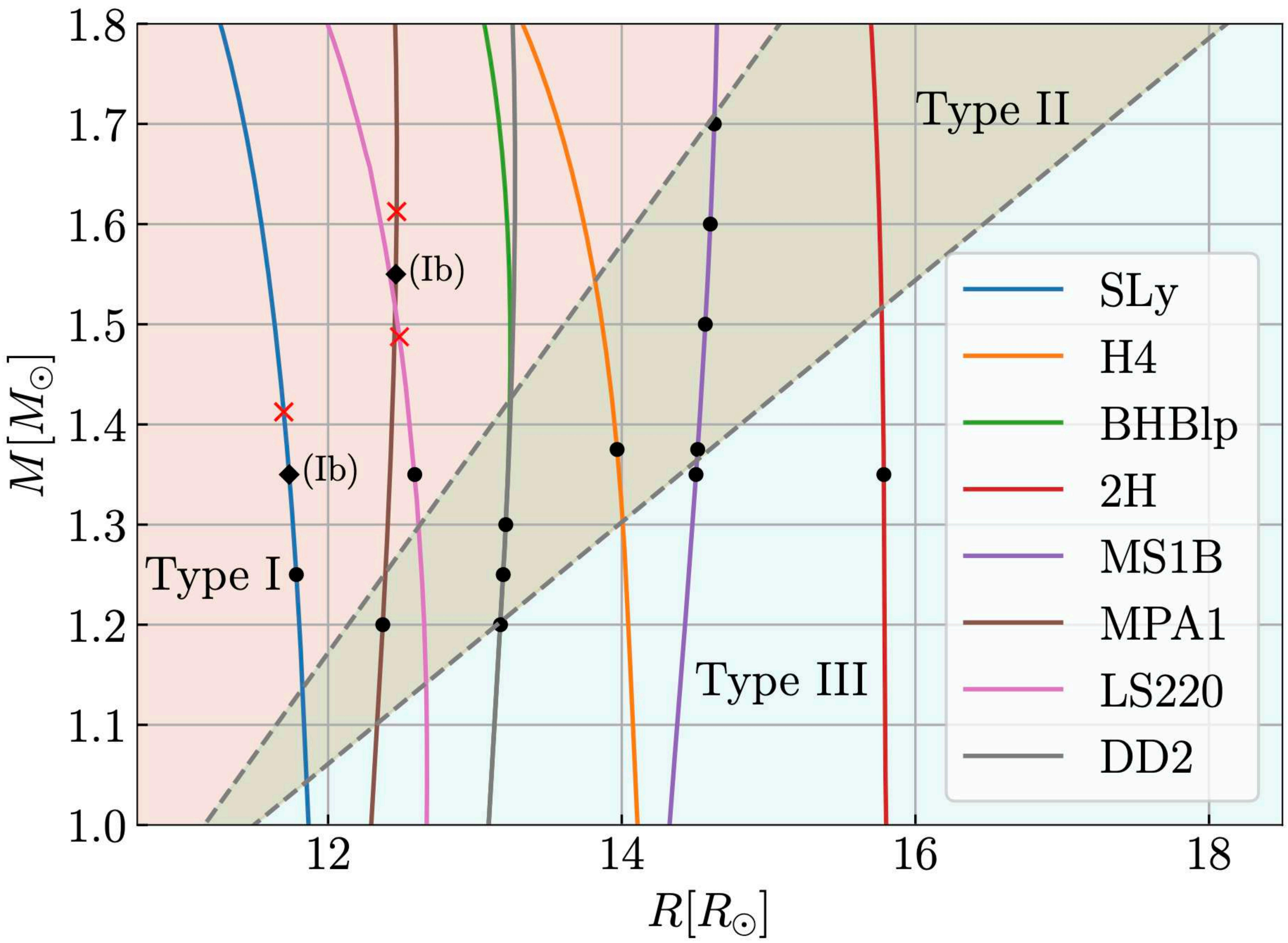
POST-MERGER PHASE IN BNS MERGERS

Orbiting spiral arms also lead to a distinct frequency f_{spiral}



Bauswein & Stergioulas (2015)

CLASSIFICATION OF POST-MERGER WAVEFORMS



Vretinaris et al. (2025)

Bauswein & Stergioulas (2015)

Vretinaris, Bauswein & Stergioulas (2020)

Type I:

$f_{2\text{-}0}$ stronger than f_{spiral}

Type II:

$f_{2\text{-}0}$ comparable to f_{spiral}

Type III:

f_{spiral} stronger than $f_{2\text{-}0}$

Type Ib:

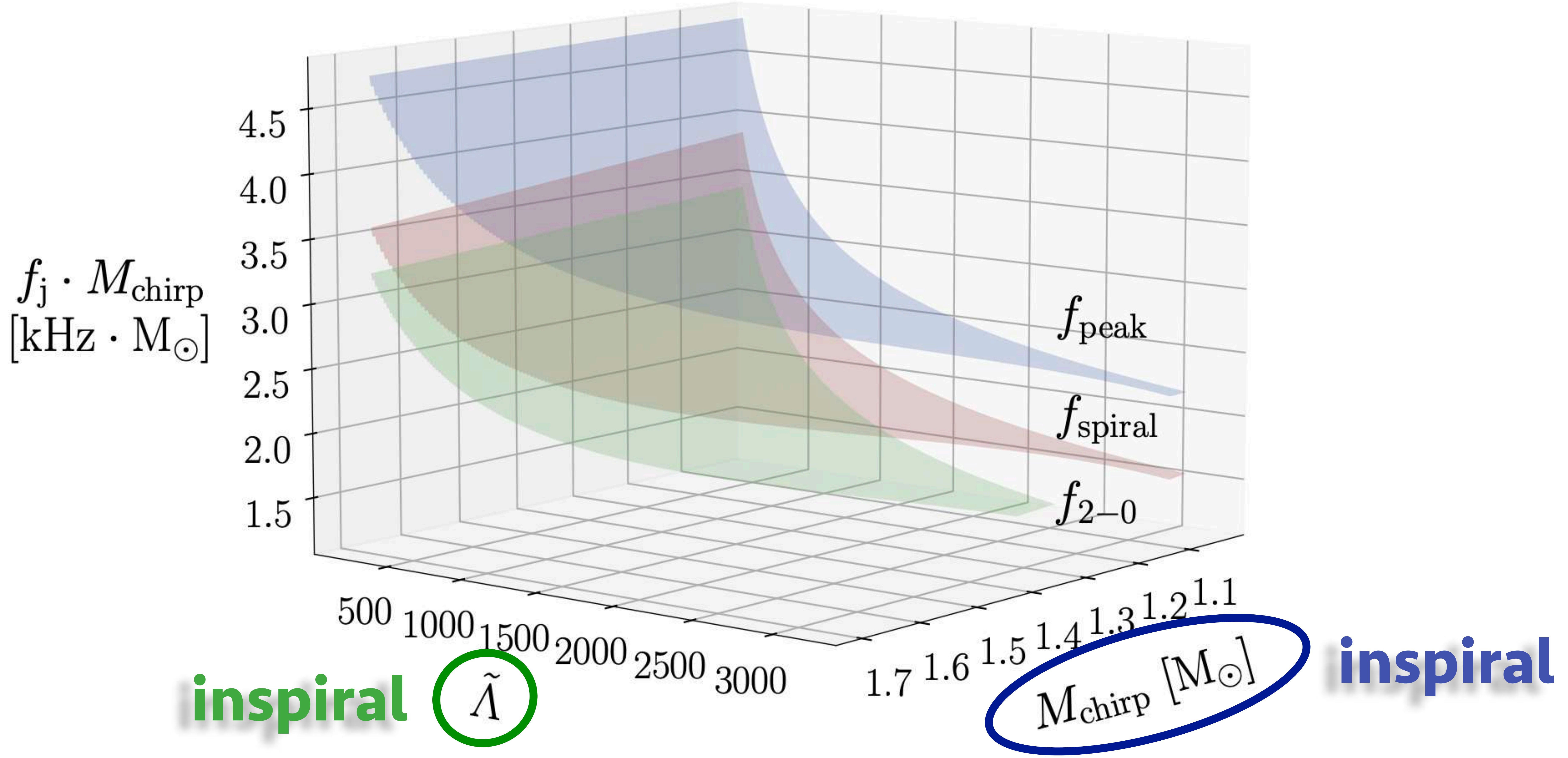
(close to M_{thres})

EMPIRICAL RELATIONS OF POST-MERGER FREQUENCIES

$$f_{\mathrm{peak}}\,M_{\mathrm{chirp}} = 1.392 - 0.108 M_{\mathrm{chirp}} + 51.70\tilde{\Lambda}^{-1/2}$$ Vretinaris, Bauswein & Stergioulas (2020)

$$f_{2-0}\,M_{\mathrm{chirp}} = 0.558 - 0.406 M_{\mathrm{chirp}} + 48.696\tilde{\Lambda}^{-1/2}$$ Vretinaris et al. (2025)

$$f_{\mathrm{spiral}}\,M_{\mathrm{chirp}} = 1.2 - 0.442 M_{\mathrm{chirp}} + 45.819\tilde{\Lambda}^{-1/2}$$ Vretinaris et al. (2025)

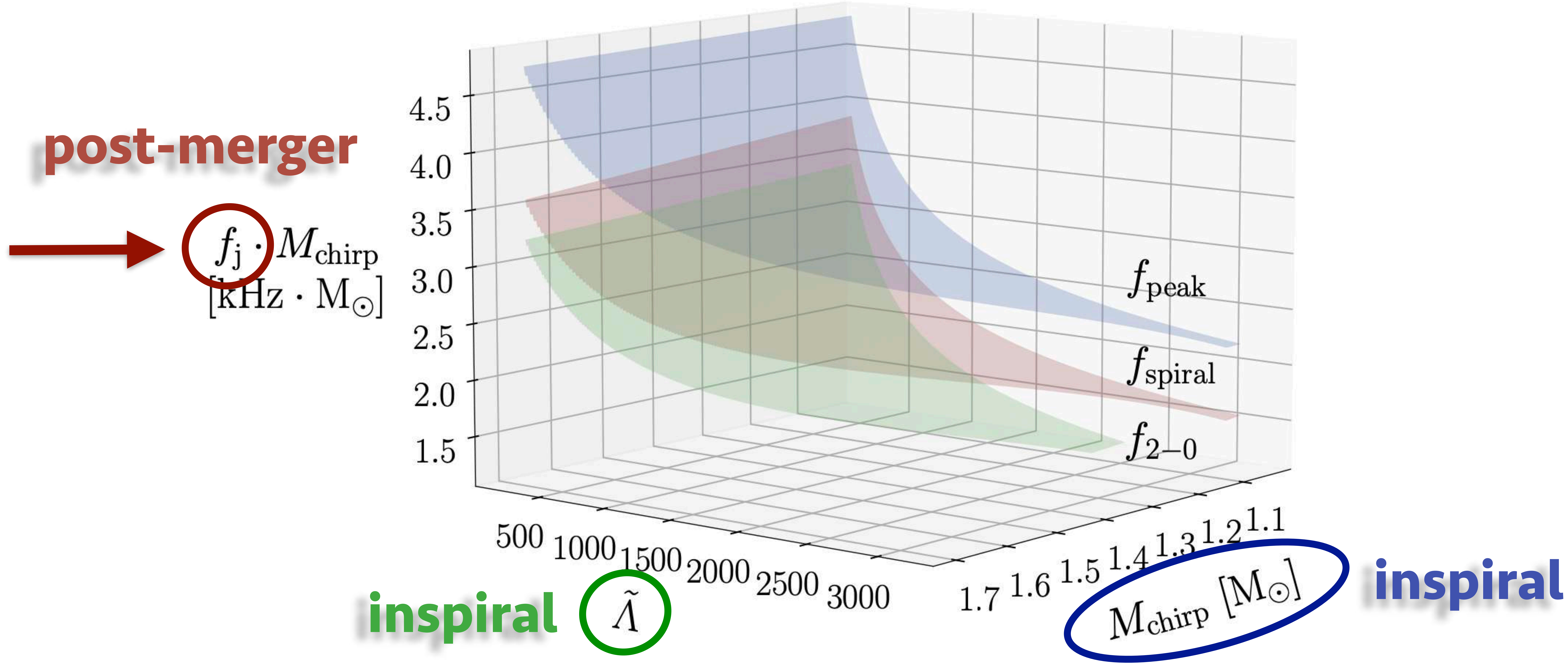


EMPIRICAL RELATIONS OF POST-MERGER FREQUENCIES

$$f_{\mathrm{peak}}\, M_{\mathrm{chirp}} = 1.392 - 0.108 M_{\mathrm{chirp}} + 51.70 \tilde{\Lambda}^{-1/2}$$ Vretinaris, Bauswein & Stergioulas (2020)

$$f_{2-0} M_{\mathrm{chirp}} = 0.558 - 0.406 M_{\mathrm{chirp}} + 48.696 \tilde{\Lambda}^{-1/2}$$ Vretinaris et al. (2025)

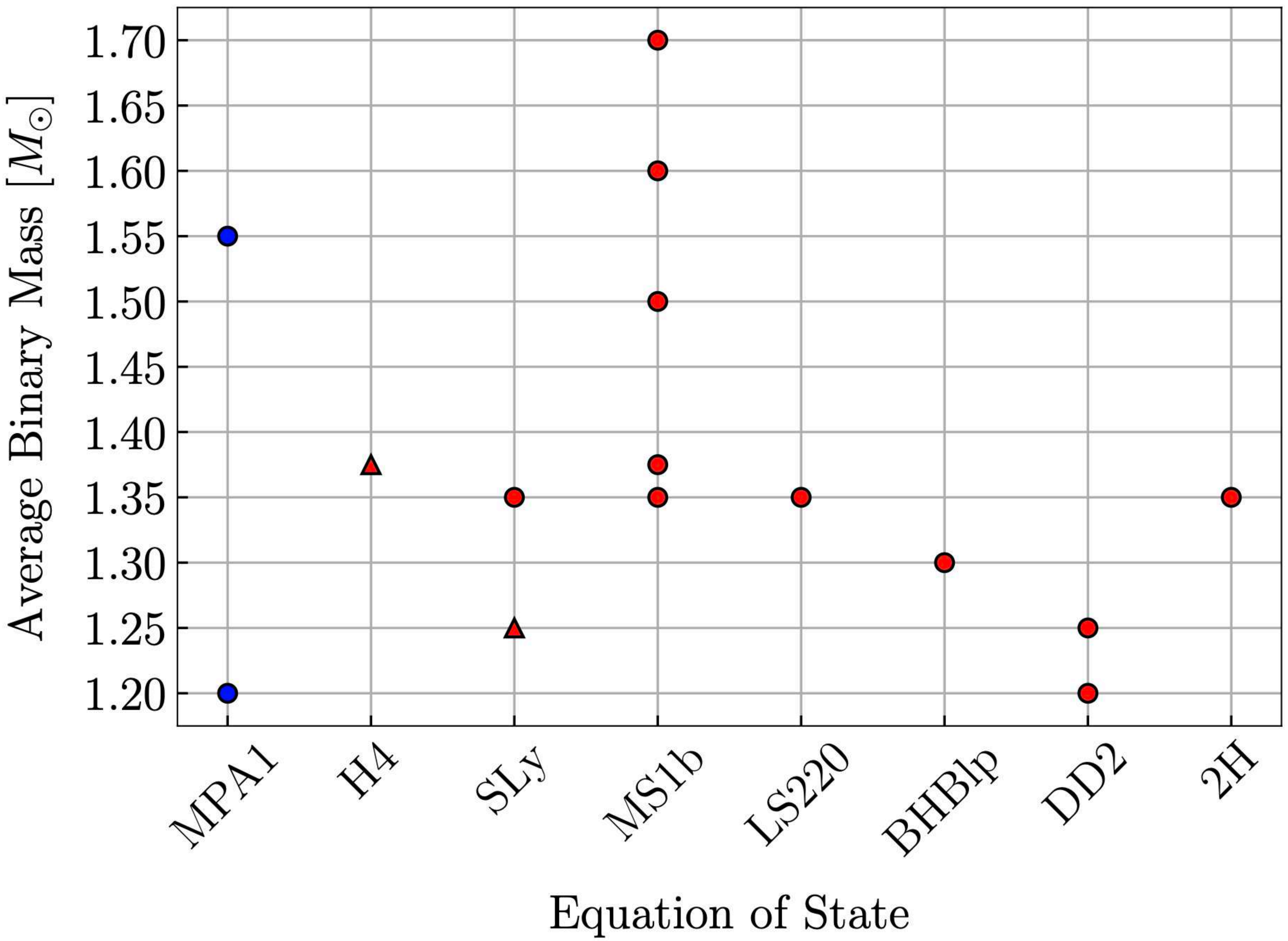
$$f_{\mathrm{spiral}}\, M_{\mathrm{chirp}} = 1.2 - 0.442 M_{\mathrm{chirp}} + 45.819 \tilde{\Lambda}^{-1/2}$$ Vretinaris et al. (2025)



13 numerical waveforms from the CoRe database (Gonzalez et al. 2022)

+2 numerical waveforms from Soultanis, Bauswein & Stergioulas (2022)

Label	EOS	q	(Average) Mass	References
THC:0036:R03	SLy	1.0	1.350	[47]
THC:0019:R05	LS220	1.0	1.350	[101, 102]
BAM:0088:R01	MS1b	1.0	1.500	[99, 100]
THC:0002:R01	BHBlp	1.0	1.300	[101, 102]
THC:0011:R01	DD2	1.0	1.250	[101, 102]
BAM:0070:R01	MS1b	1.0	1.375	[103]
BAM:0065:R03	MS1b	1.0	1.350	[104]
THC:0010:R01	DD2	1.0	1.200	[101, 102]
BAM:0002:R02	2H	1.0	1.350	[104]
BAM:0053:R01	H4	1.5	1.375	[105]
BAM:0124:R01	SLy	1.5	1.250	[103]
BAM:0090:R02	MS1b	1.0	1.600	[99, 100]
BAM:0092:R02	MS1b	1.0	1.700	[99, 100]
Soultanis et al.	MPA1	1.0	1.200	[67]
Soultanis et al.	MPA1	1.0	1.550	[67]



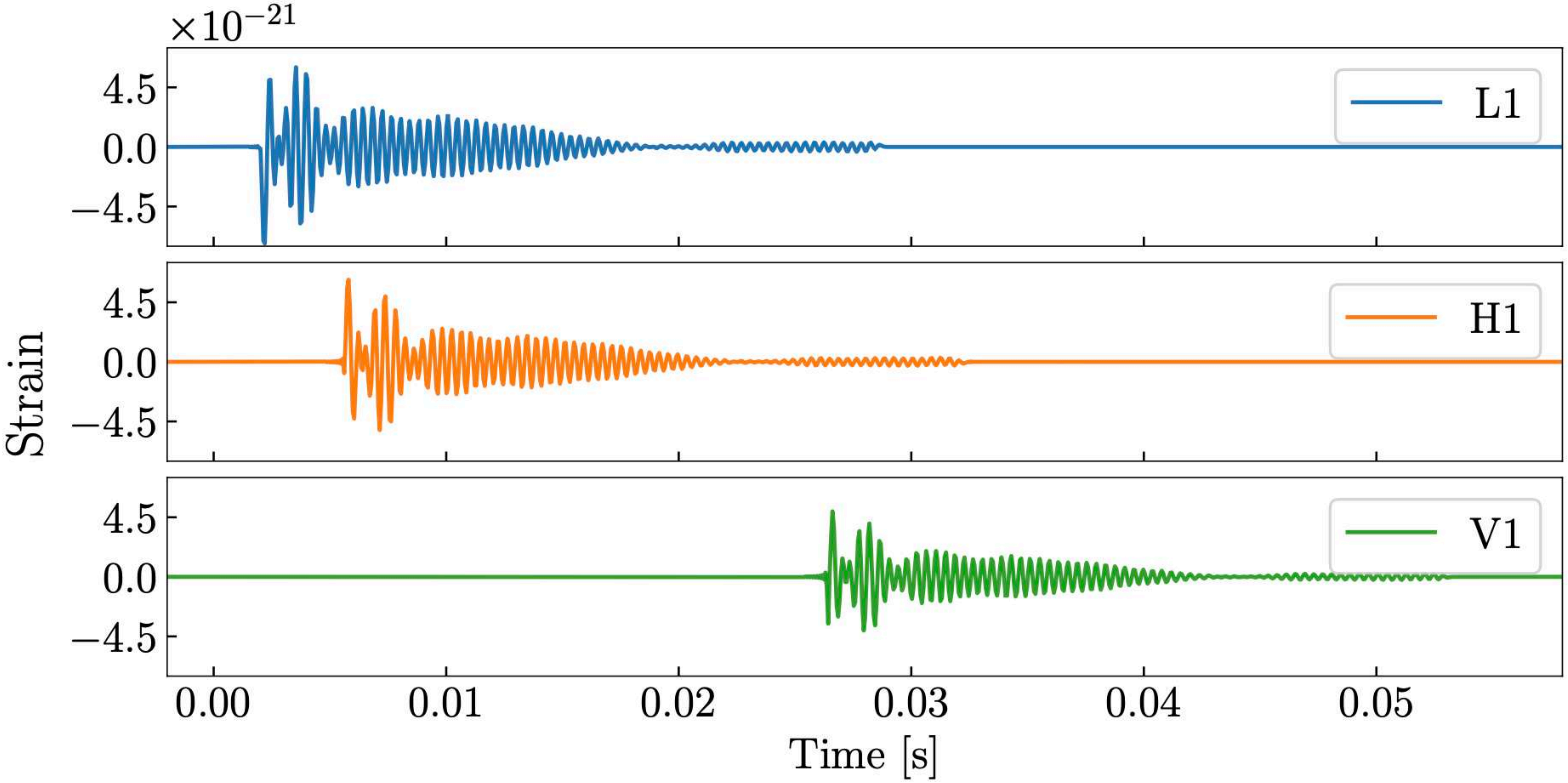
Extension of the set of 9 waveforms used in previous work by Easter et al. (2020)

Vretinaris S., Vretinaris G., Mermigkas, Karamanis, Stergioulas (2025)

INJECTIONS IN 3-DETECTOR NETWORK

Waveforms are injected in 3-detector HLV network at design sensitivity, using BILBY (Ashton et al. 2019)

We choose post-merger SNR of 8, 16 and 50, to simulate detection by 3G network (ET/CE) at distances as close as 200Mpc.



ANALYTIC BNS POST-MERGER WAVEFORM MODELS

Several analytic models exist in the time- and frequency-domains.

Here, we extend the analytic model of Easter et al. (2020) as follows:

$$
\begin{aligned}
h(\boldsymbol{\theta}, t) &= h_{+}(\boldsymbol{\theta}, t) - i h_{\times}(\boldsymbol{\theta}, t) \\
&= \sum_{j=1}^{4} \left[h_{j,+}(\boldsymbol{\theta}, t) - i h_{j,\times}(\boldsymbol{\theta}, t)\right] \\
h_{j,+}(\boldsymbol{\theta}, t) &= A_j \exp\left[-\frac{t}{T_j}\right] \cos\left[2\pi f_j t \left(1 + a_j t\right) + \psi_j\right]
\end{aligned}
$$

ANALYTIC BNS POST-MERGER WAVEFORM MODELS

Several analytic models exist in the time- and frequency-domains.

Here, we extend the analytic model of Easter et al. (2020) as follows:

$$
\begin{aligned}
h(\boldsymbol{\theta}, t) &= h_{+}(\boldsymbol{\theta}, t) - i h_{\times}(\boldsymbol{\theta}, t) \\
&= \sum_{j=1}^{4} \left[h_{j,+}(\boldsymbol{\theta}, t) - i h_{j,\times}(\boldsymbol{\theta}, t) \right] \\
h_{j,+}(\boldsymbol{\theta}, t) &= A_j \exp\left[-\frac{t}{T_j} \right] \cos\left[2\pi f_j t \left(1 + a_j t \right) + \psi_j \right]
\end{aligned}
$$

Intrinsic parameters $\boldsymbol{\theta} = \{A_j, T_j, f_j, a_j, \psi_j\} \quad \text{for} \quad j \in [1, 4]$

$h_{j,\times}(\boldsymbol{\theta},t)$ is obtained by applying a π/2 phase shift to $h_{j,+}(\boldsymbol{\theta},t)$

- We have added a 4th oscillator at high frequencies > f_{peak} (e.g. f_{2+0})
- Amplitudes A_j are free parameters

INFORMED PRIORS

In Easter et al. (2020) flat priors in a wide frequency range of 1-5 kHz for every oscillator were used.

- Here, we take advantage of the empirical relations to set Gaussian priors in a narrow frequency range around each expected frequency.

- I addition, the priors differ, according to the type of the post-merger waveform:

  - Type I: Gaussian priors, $\mathcal{N}(f_{2-0}, \sigma^2)$ for f_{2-0} and uniform priors $\mathcal{U}(1,5)$[kHz] for $f_{\rm spiral}$.
  - Type II: Gaussian priors, $\mathcal{N}(f_{2-0}, \sigma^2)$ for f_{2-0} and $\mathcal{N}(f_{\rm spiral}, \sigma^2)$ for $f_{\rm spiral}$.
  - Type III: Gaussian priors, $\mathcal{N}(f_{\rm spiral}, \sigma^2)$ for $f_{\rm spiral}$ and uniform priors $\mathcal{U}(1,5)$[kHz] for f_{2-0}.

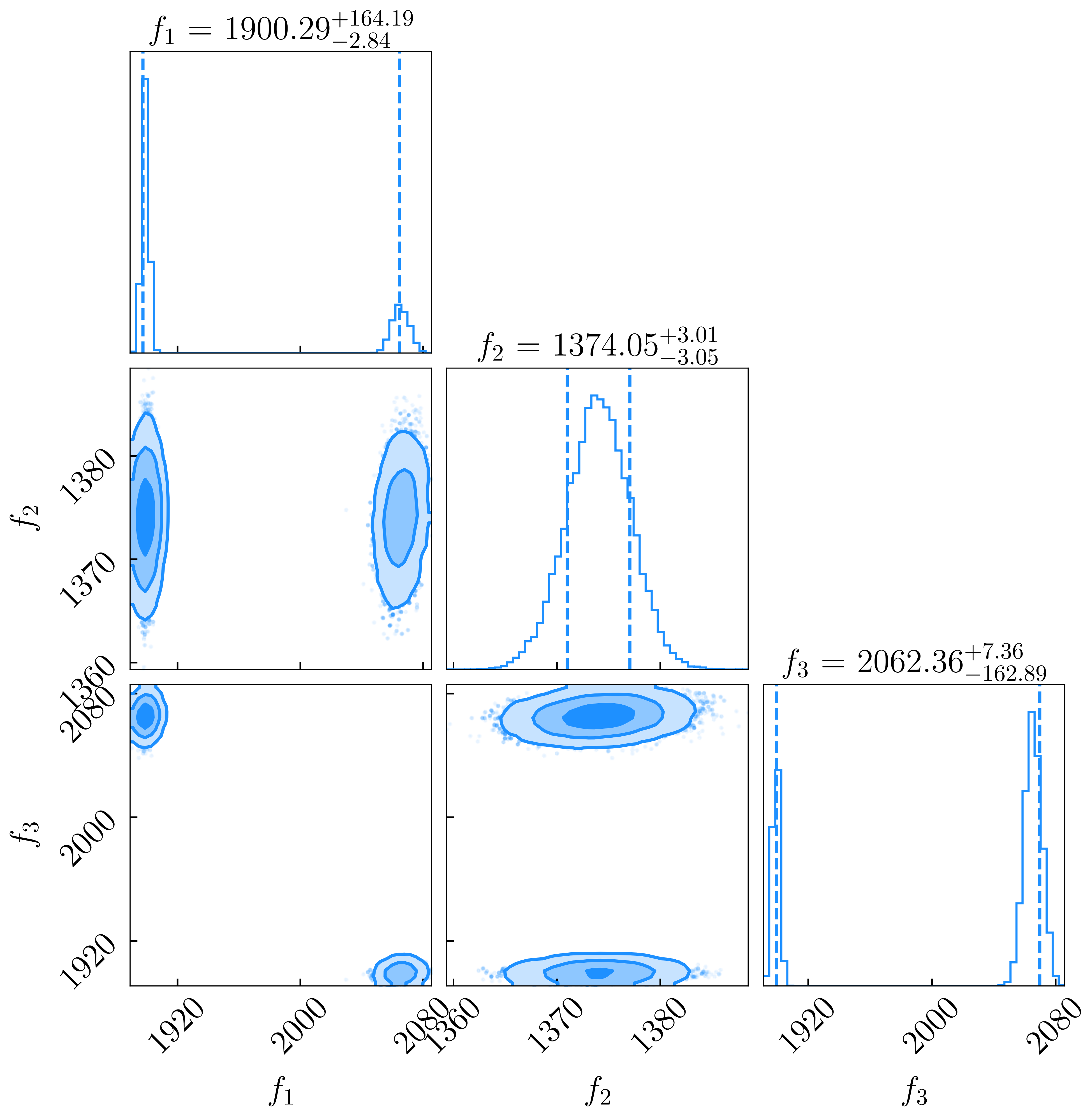
where σ = 3 x max error of empirical relations. For $f_{\rm peak}$: Gaussian ; for f_4: flat in [$f_{\rm peak}$+0.3kHz, 5 kHz].

priors for Aj: uniform in [-24, -19]
priors for aj: uniform in [-6.4, 6.4]
priors for remaining parameters: as in Easter et al. (2020)

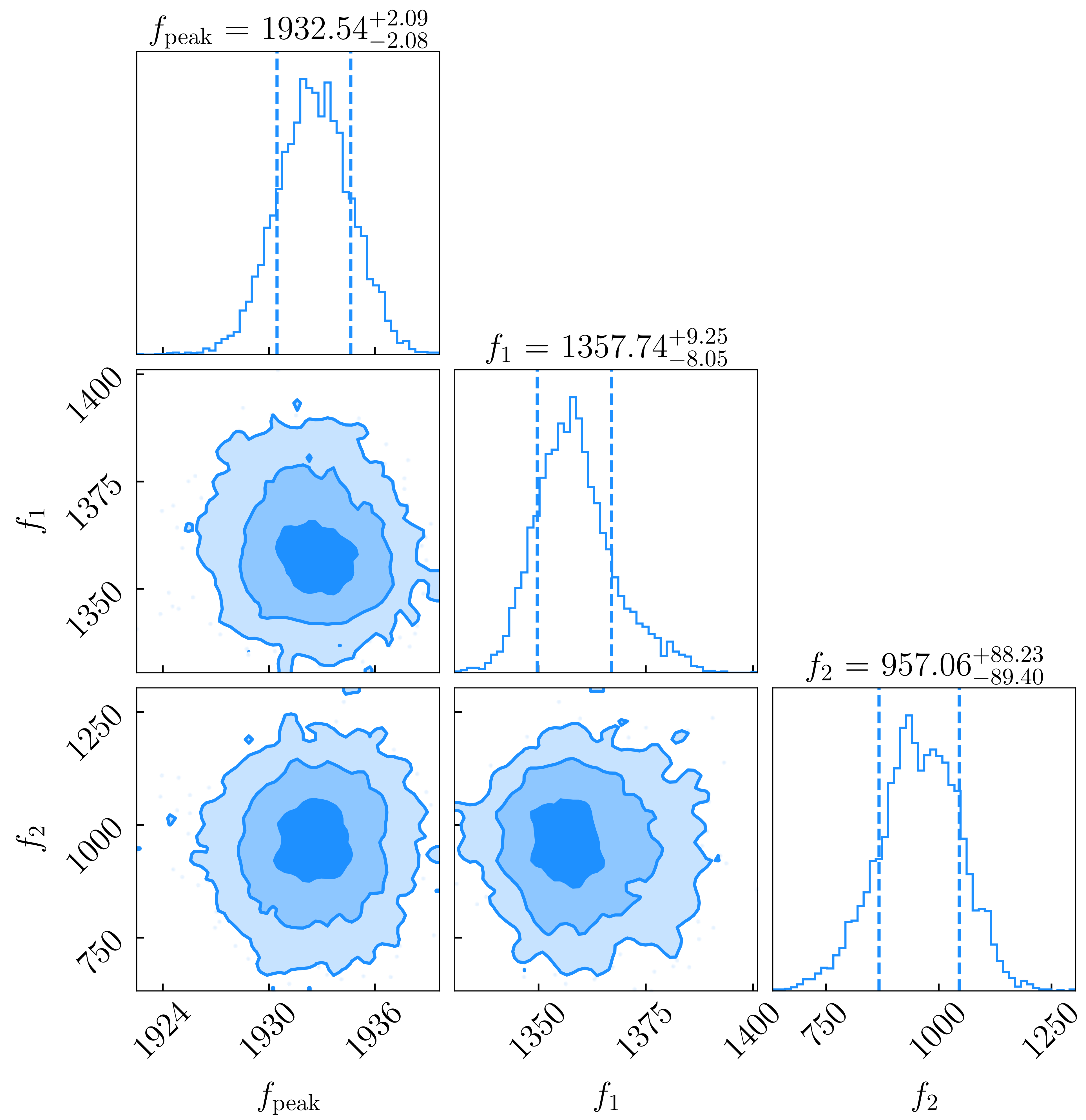
POSTERIOR DISTRIBUTIONS

EOS 2H 1.35+1.35, SNR = 50

using flat priors



using informed priors (through empirical relations)

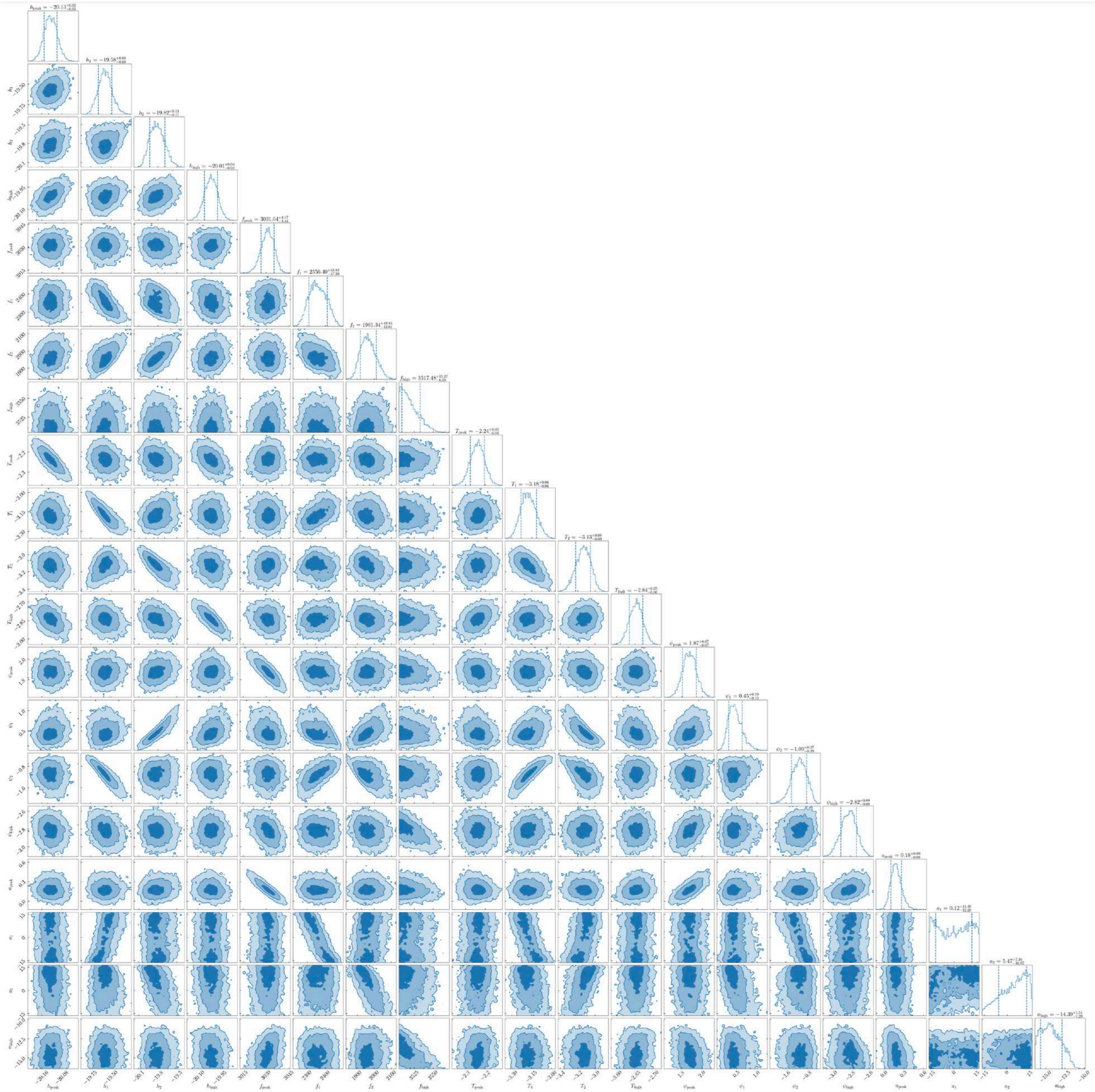


POSTERIOR DISTRIBUTIONS

EOS MPA 1.55+1.55, SNR = 50

using informed priors

all parameters



POCOMC: PRECONDITIONED MONTE CARLO SAMPLER

Preconditioned Monte Carlo Sampling (Karamanis et al. 2022)

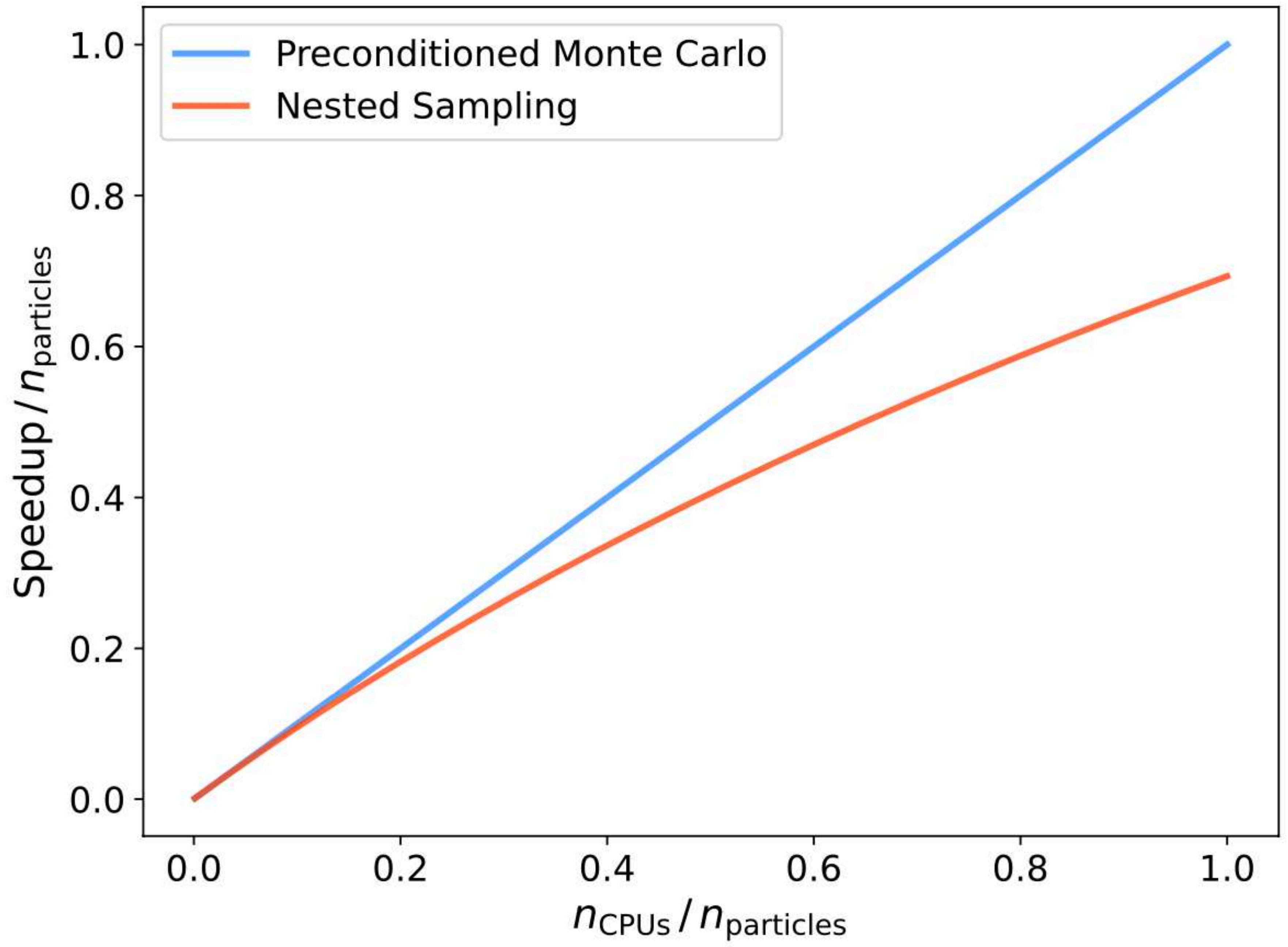
PMC targets an **annealed version of the posterior**, with density given by

$$p_t(\boldsymbol{\theta} \mid d) \propto \mathcal{L}^{\beta_t}(\boldsymbol{\theta} \mid d)\pi(\boldsymbol{\theta})$$

where β_t a parameter (inverse temperature).

We use 2000 particles that transition from the prior ($\beta_0 = 0$) to the posterior distribution ($\beta_T = 1$), through a sequence of reweighing, resampling and mutation steps.

https://github.com/minaskar/pocomc



PocoMC is highly parallelizable

POCOMC: PRECONDITIONED MONTE CARLO SAMPLER

Preconditioned Monte Carlo Sampling (Karamanis et al. 2022)

PMC targets an **annealed version of the posterior**, with density given by

$$p_t(\boldsymbol{\theta} \mid d) \propto \mathcal{L}^{\beta_t}(\boldsymbol{\theta} \mid d)\pi(\boldsymbol{\theta})$$

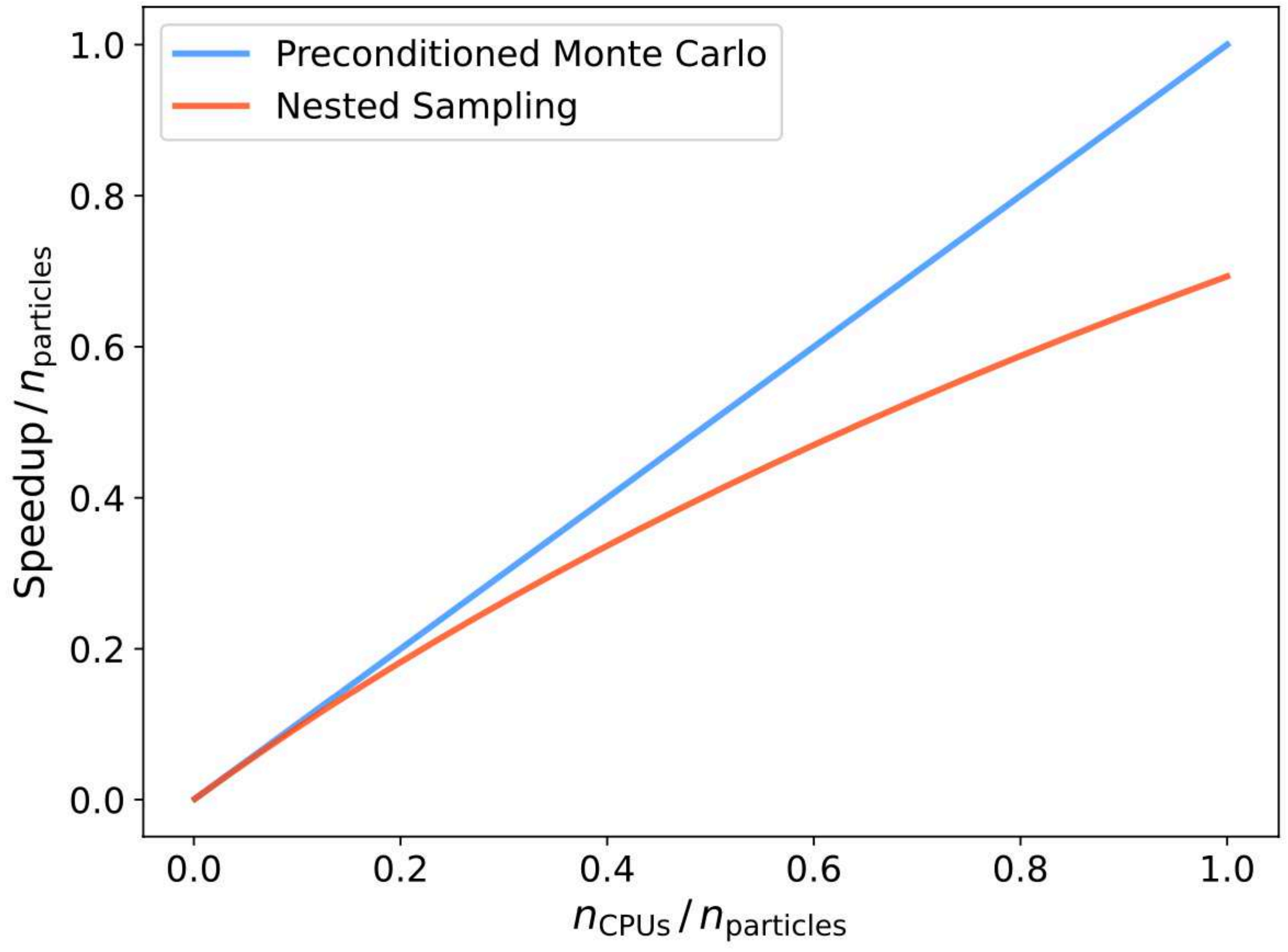
where β_t a parameter (inverse temperature).

We use 2000 particles that transition from the prior ($\beta_0 = 0$) to the posterior distribution ($\beta_T = 1$), through a sequence of reweighing, resampling and mutation steps.

After each iteration, a **normalizing flow** transforms the distribution to a simpler one (almost Gaussian), decorrelating the parameters. This allows for a much faster sampling.

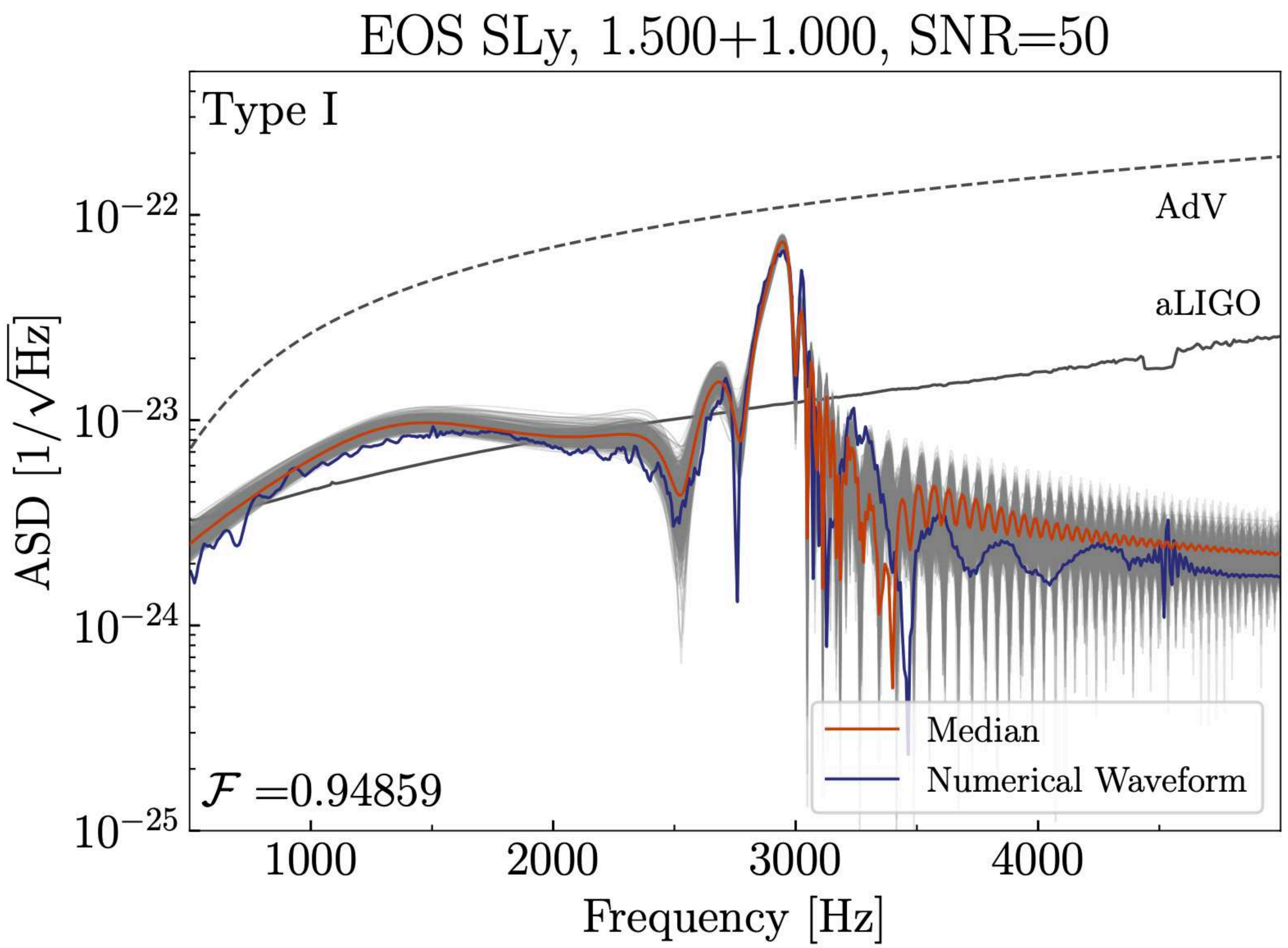
On same number of CPUs: pocoMC is ~**10 times faster** than dynesty.

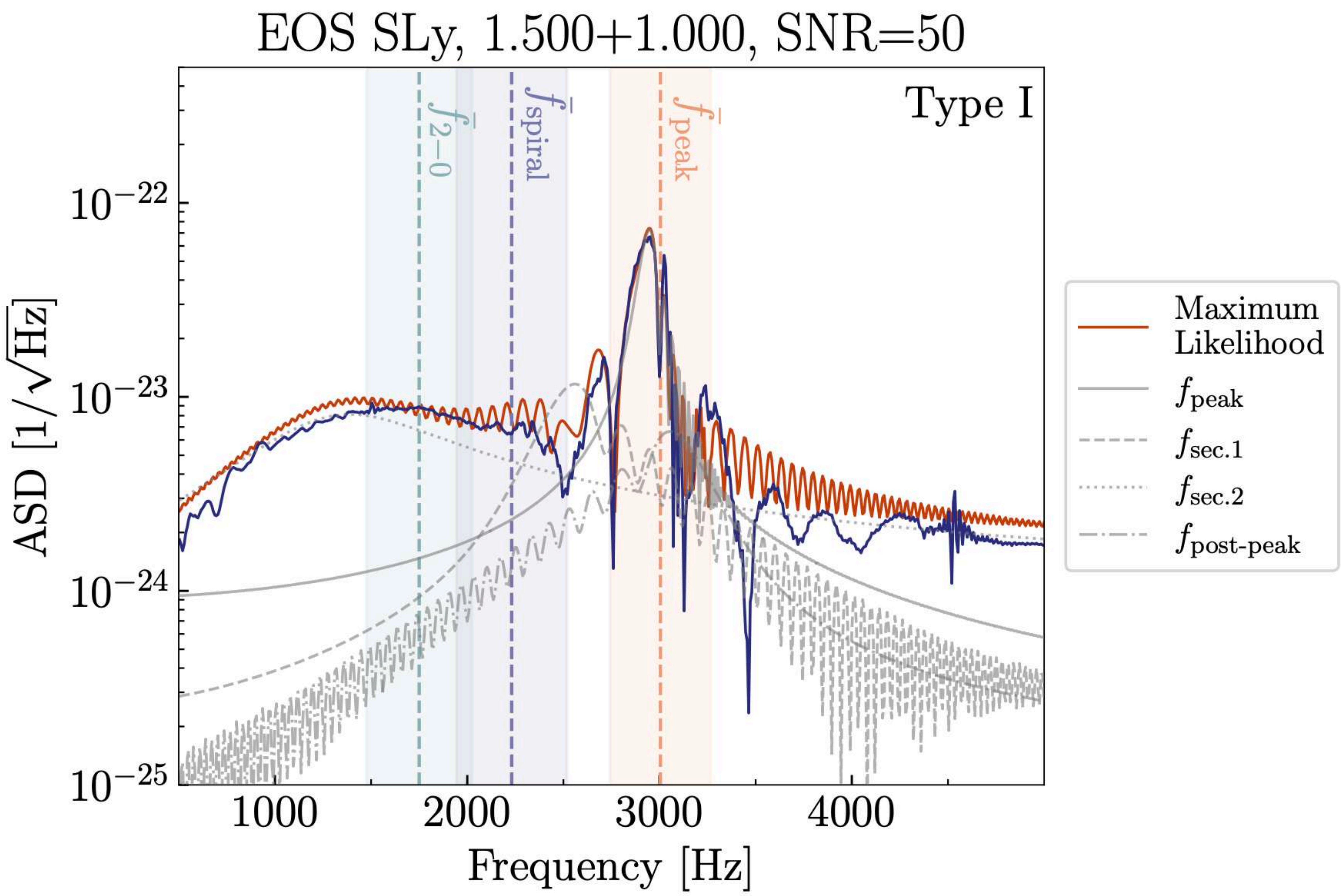
https://github.com/minaskar/pocomc



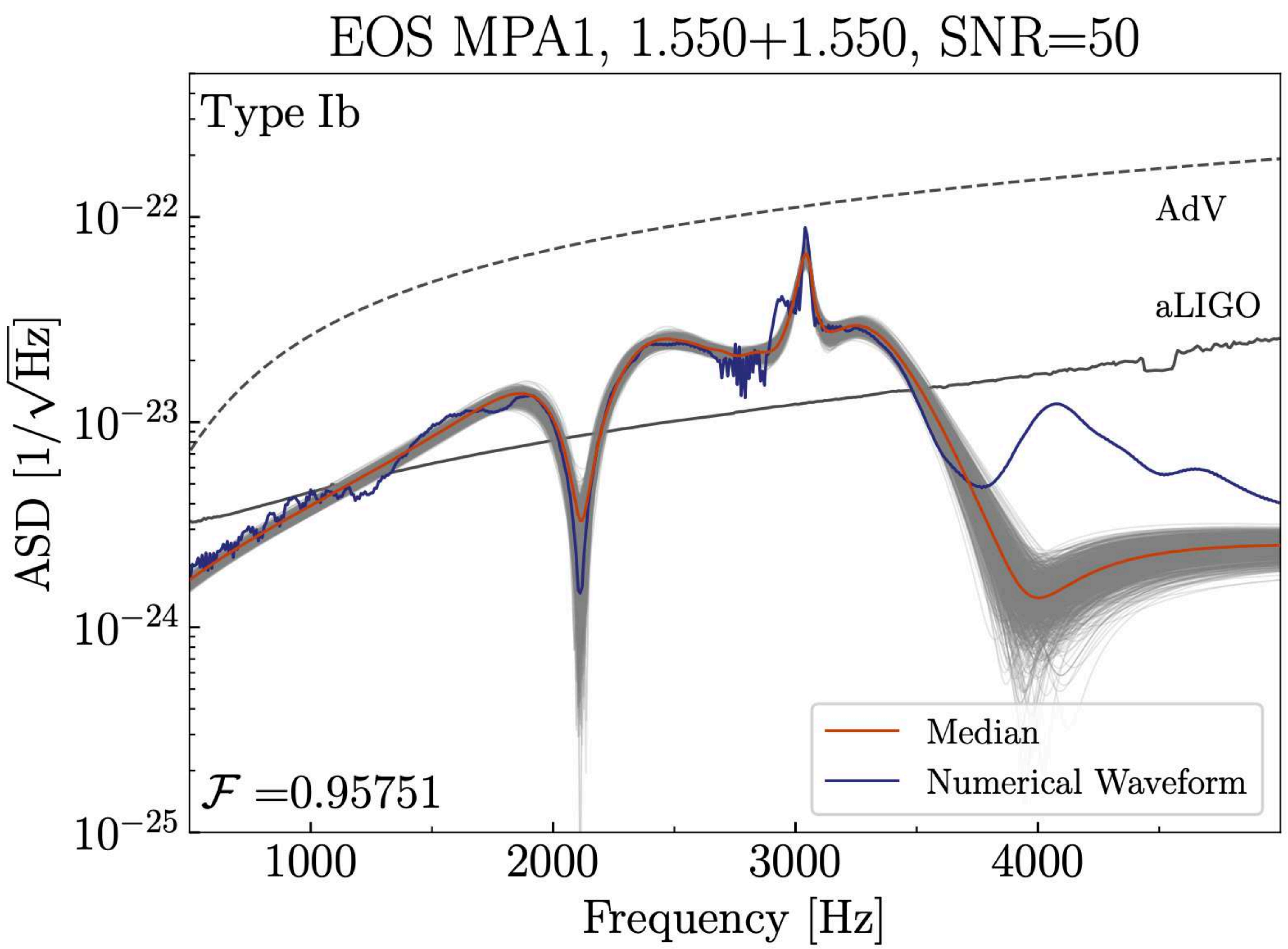
PocoMC is highly parallelizable

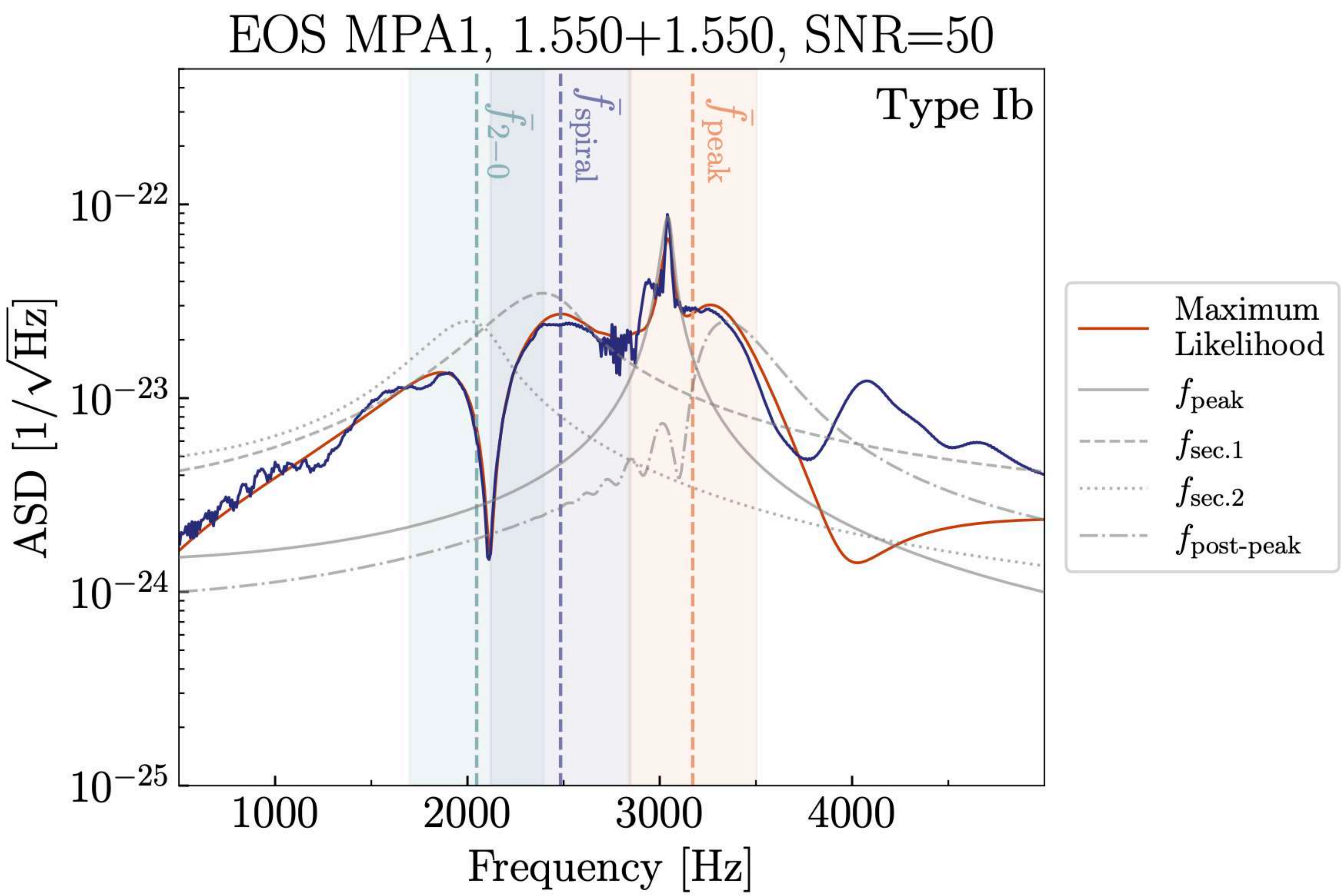
RECONSTRUCTION IN THE FREQUENCY DOMAIN



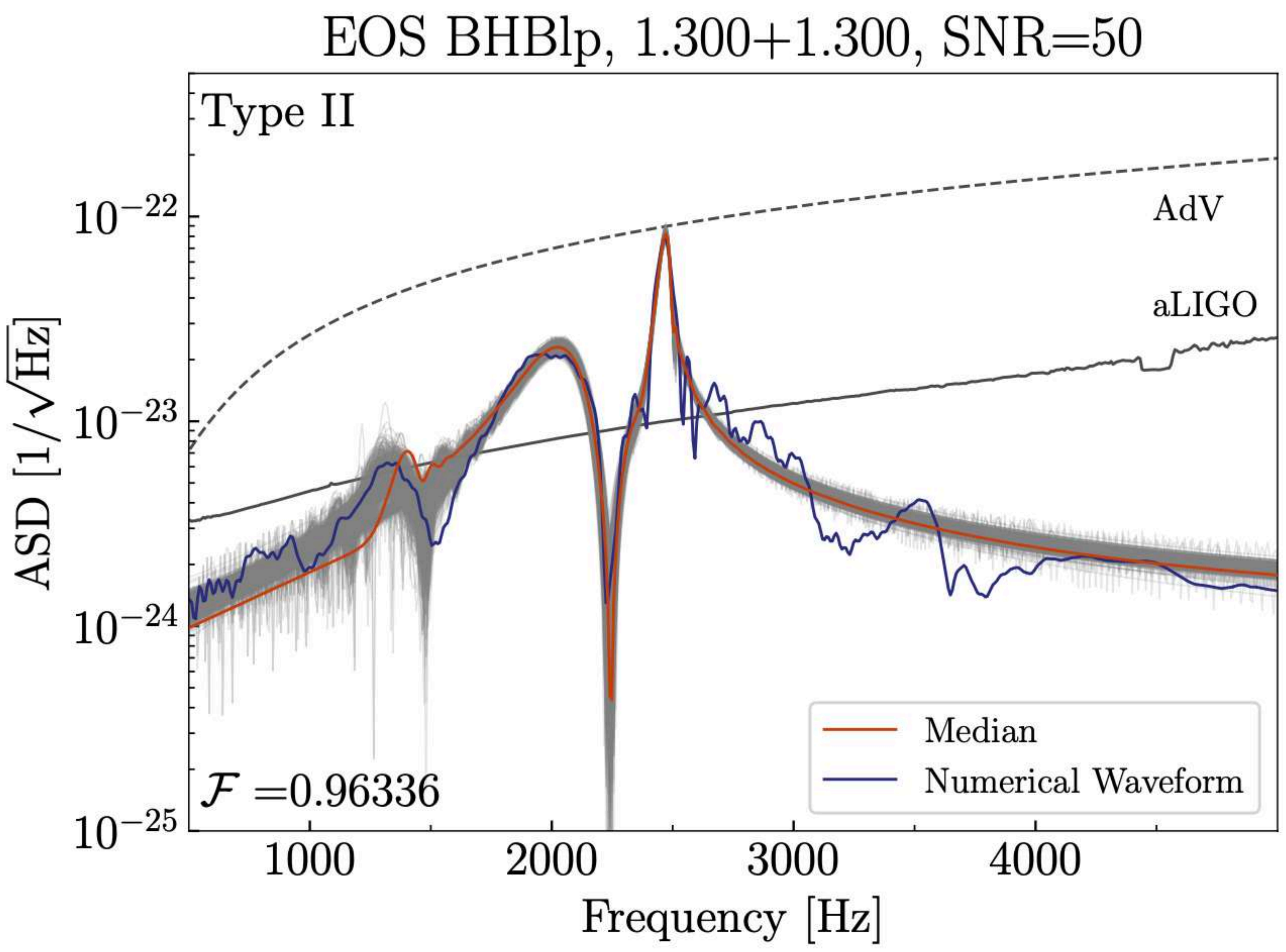


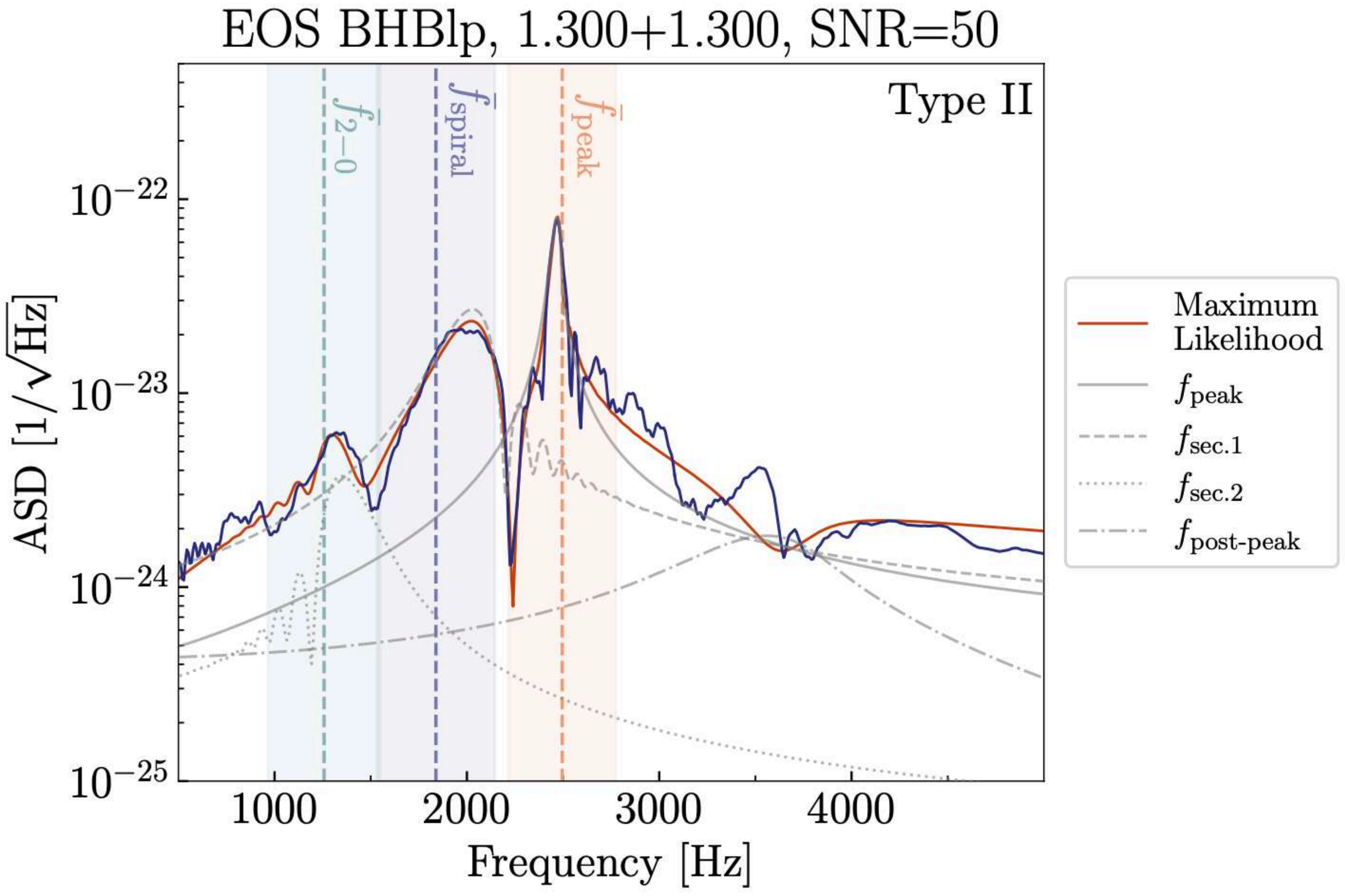
RECONSTRUCTION IN THE FREQUENCY DOMAIN



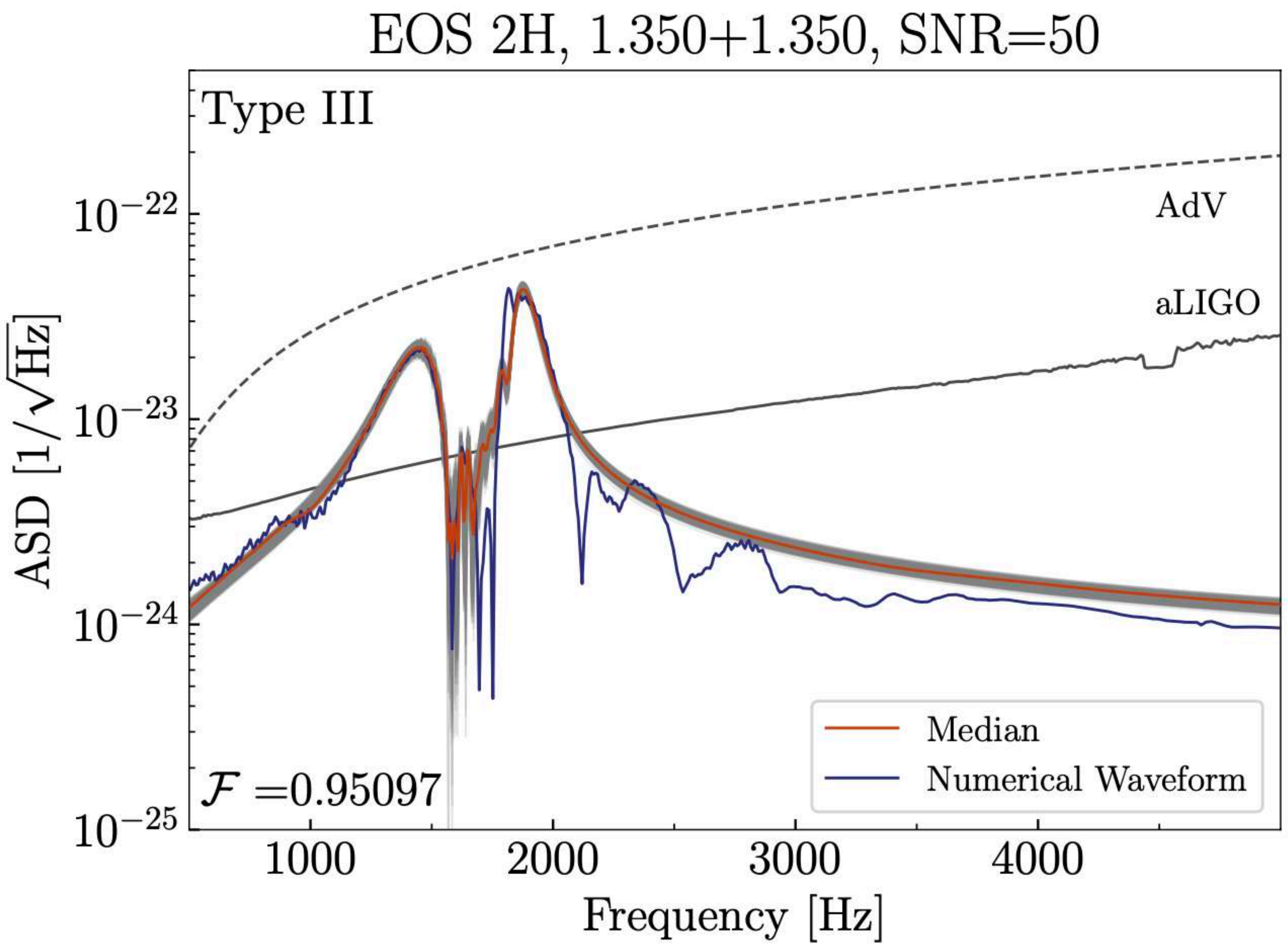


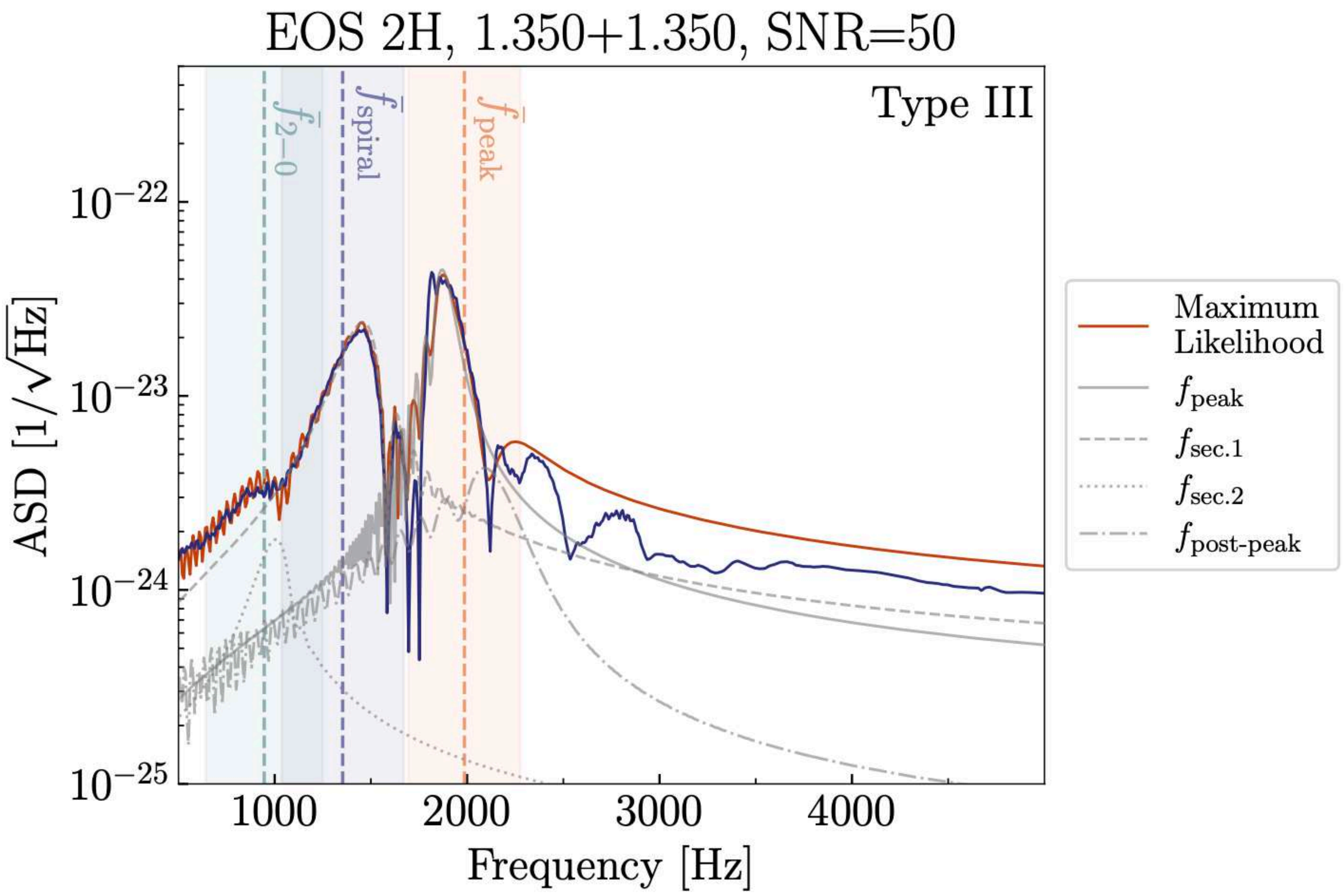
RECONSTRUCTION IN THE FREQUENCY DOMAIN





RECONSTRUCTION IN THE FREQUENCY DOMAIN



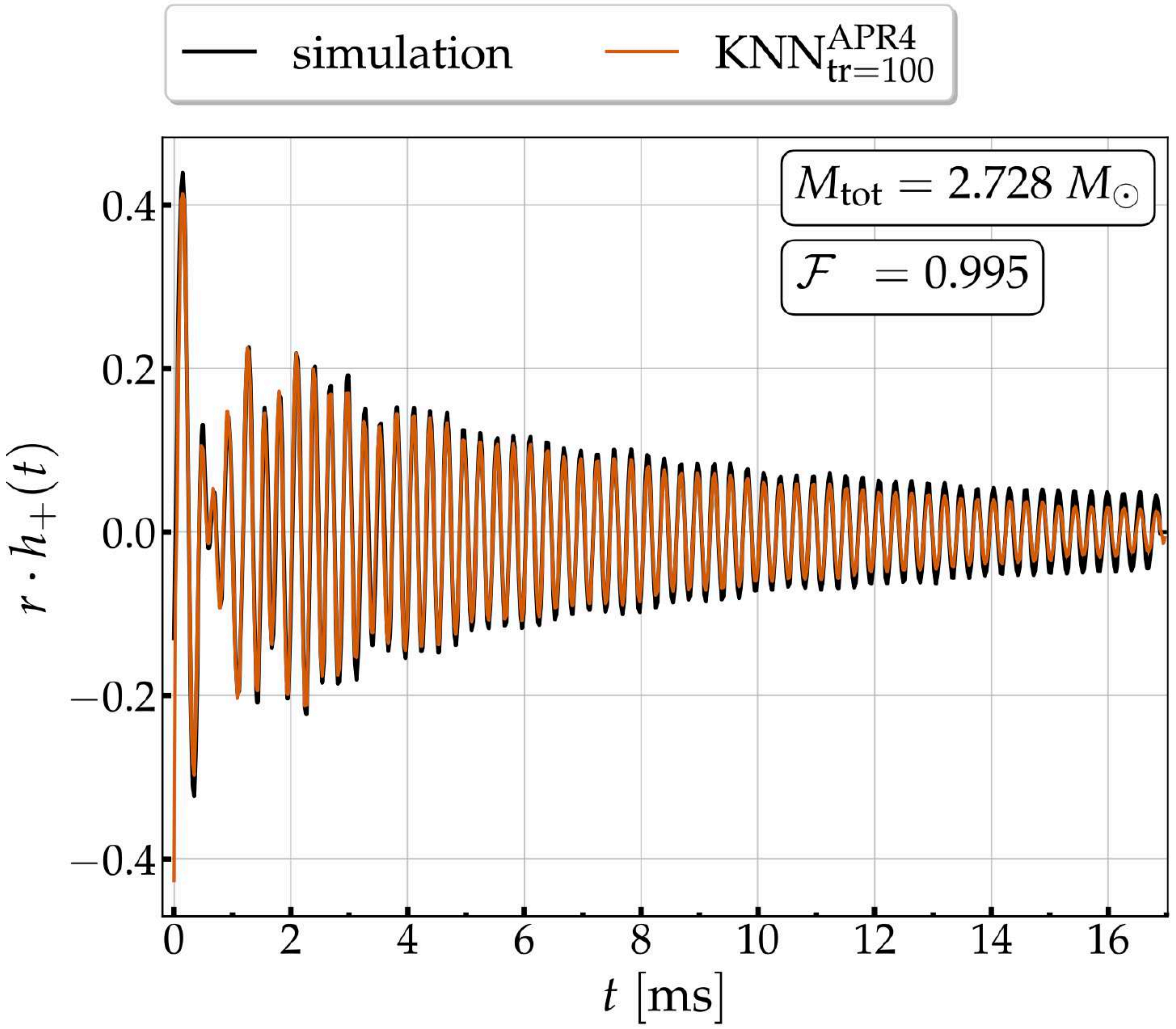


APPLICATION OF MACHINE LEARNING TO THE POST-MERGER PHASE

Problem: Only O(200) substantially different numerical BNS simulations are currently available.

Solution: Construct surrogate model of post-merger GWs as function of e.g. M, q, EOS(Λ)

Time domain surrogate model using
K-Nearest Neighbor (KNN) regression:



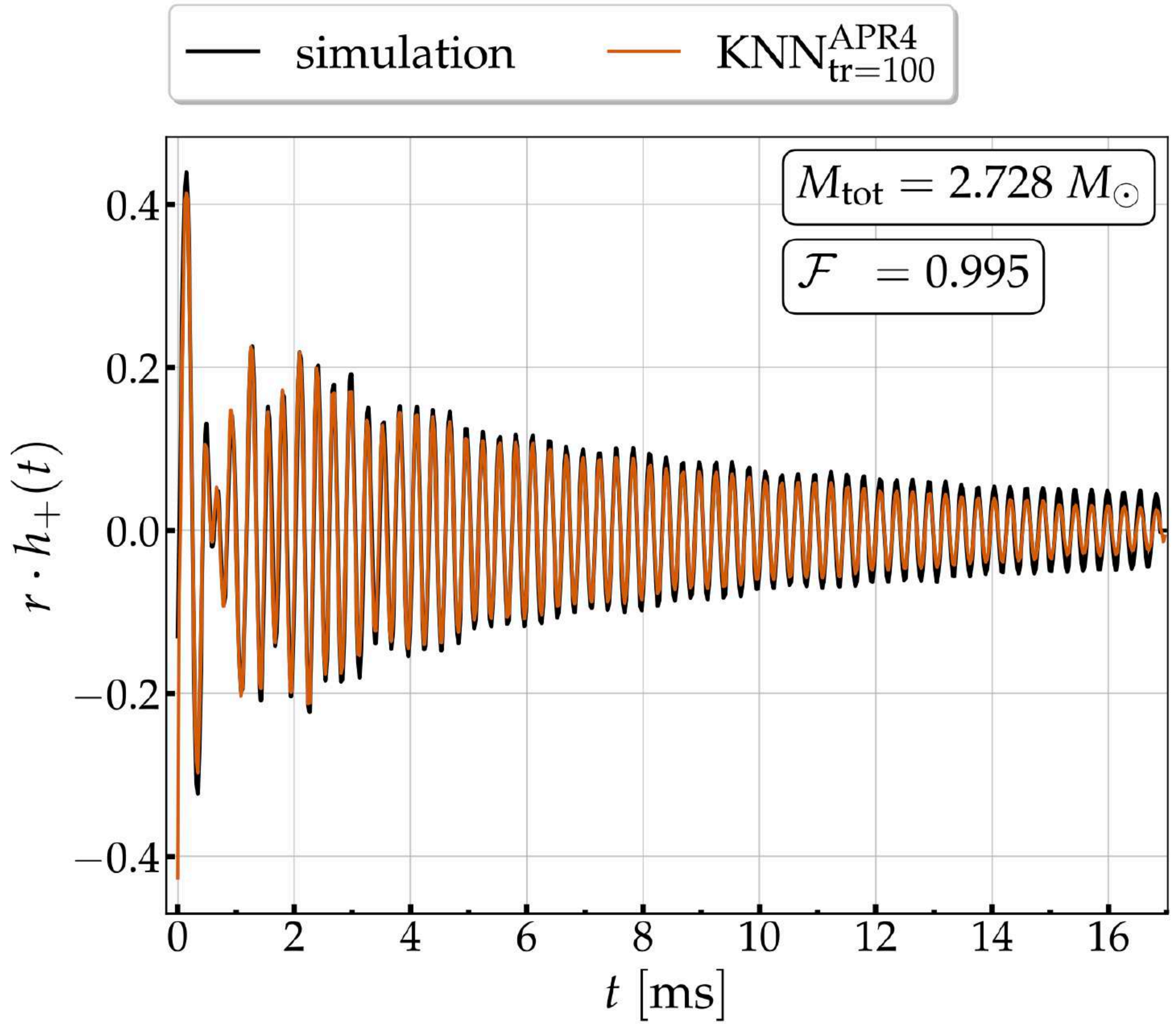
Soultanis et al. (2025)

APPLICATION OF MACHINE LEARNING TO THE POST-MERGER PHASE

Problem: Only O(200) substantially different numerical BNS simulations are currently available.

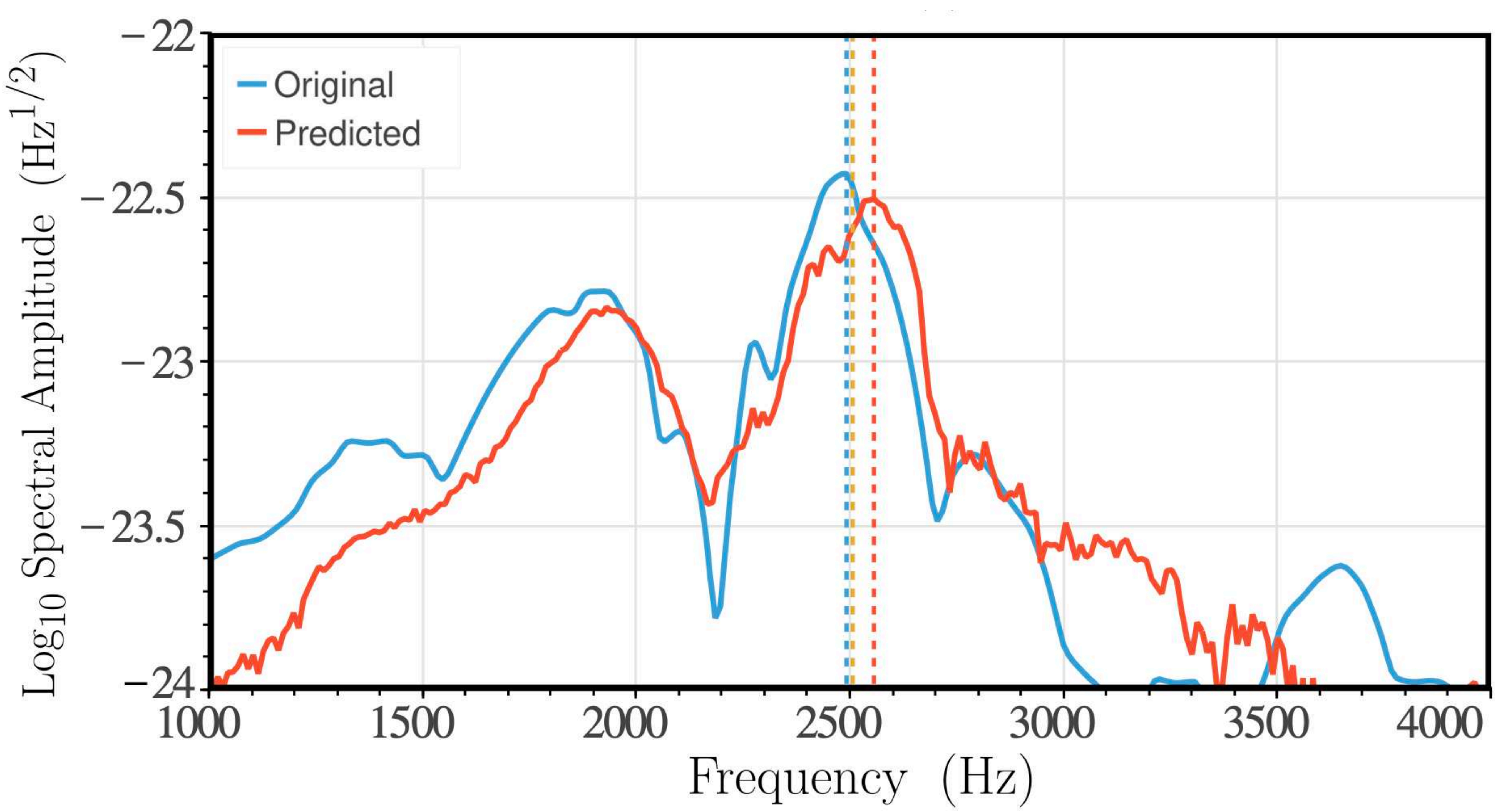
Solution: Construct surrogate model of post-merger GWs as function of e.g. M, q, EOS(Λ)

Time domain surrogate model using
K-Nearest Neighbor (KNN) regression:



Soultanis et al. (2025)

Frequency domain surrogate model using
Artificial Neural Networks (ANN) regression:



Pesios et al. (2024)

Training set: 20-100 different M_{tot} (q=1) between 2.4 and 2.8 M_{sun}.

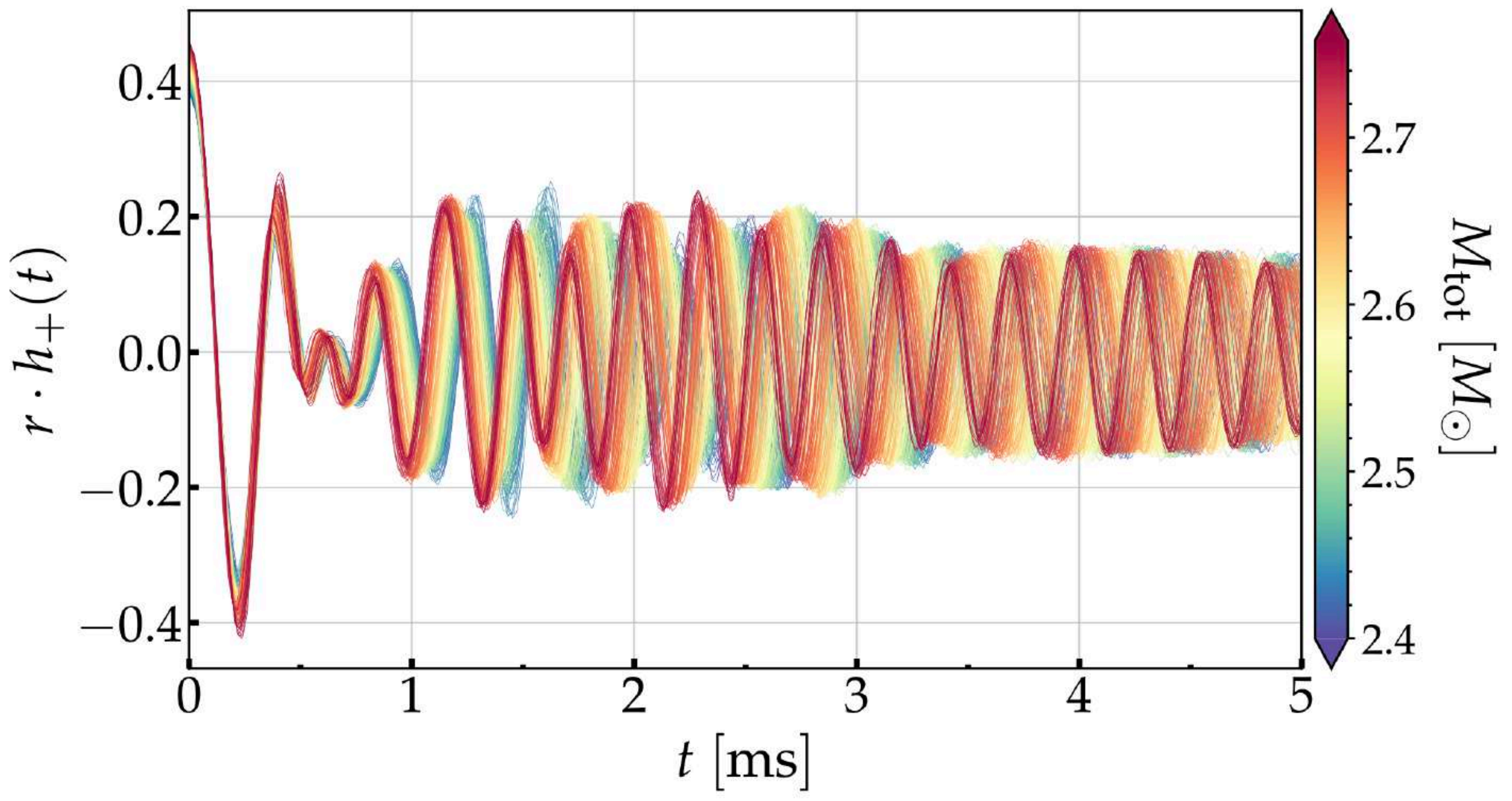
Choose specific EOS (APR4; SFHX).

Complex GW strain:

$$h(t) = h_+(t) + i h_\times(t)$$

$$= |h(t)| \cdot e^{+i\phi(t)},$$

Signals are aligned at merger time:



K-NEAREST NEIGHBOR REGRESSION IN THE TIME DOMAIN

Training set: 20-100 different M_{tot} (q=1) between 2.4 and 2.8 M_{sun}.

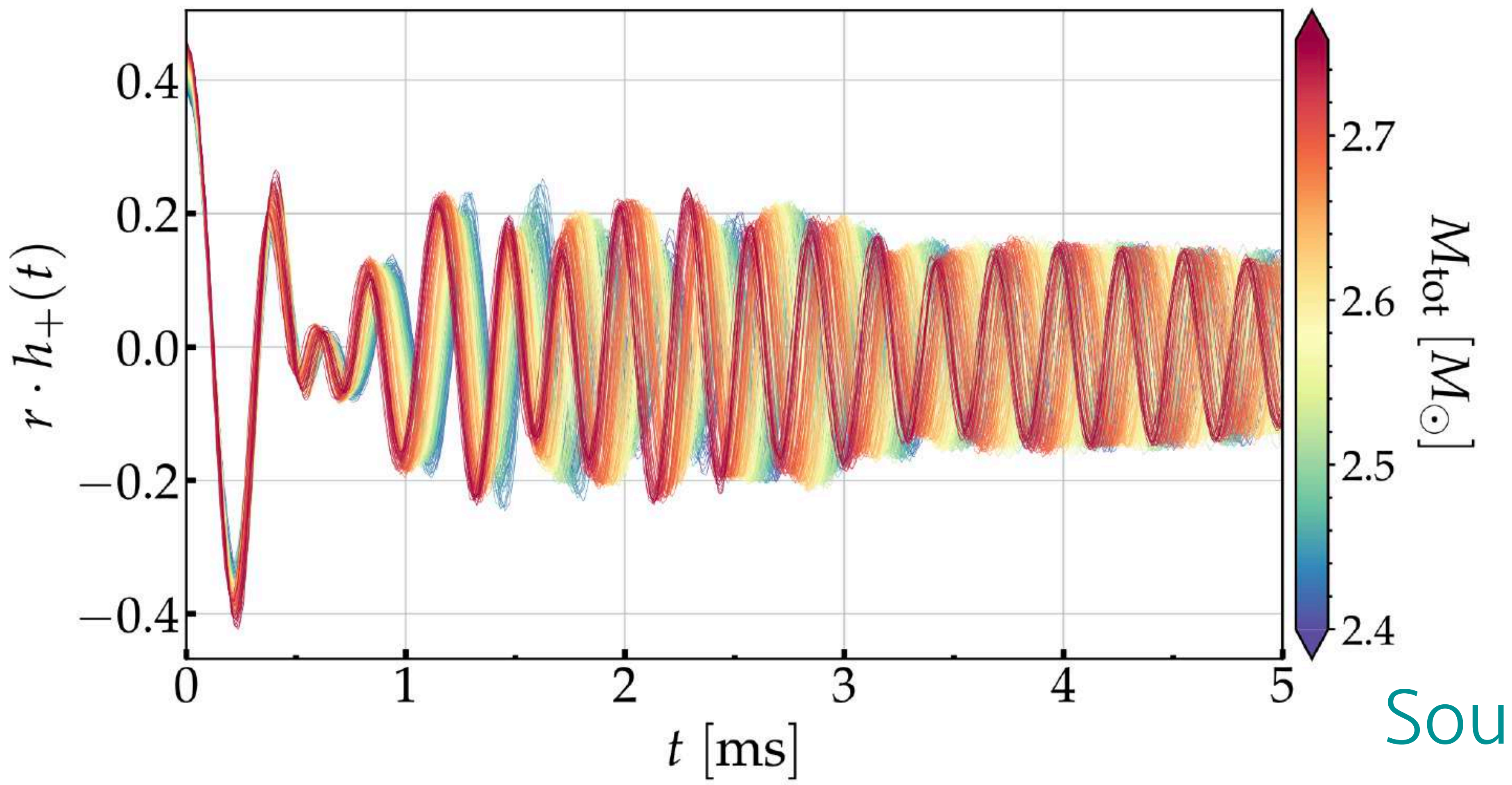
Choose specific EOS (APR4; SFHX).

Complex GW strain:

$$h(t) = h_+(t) + ih_\times(t)$$

$$= |h(t)| \cdot e^{+i\phi(t)},$$

Signals are aligned at merger time:



Input features: $\vec{X}_i = \{M_{\rm tot}, t_j\}$

Predictions: $\vec{Y}_i = \text{strain data}$

KNN algorithm: For given input $\vec{X}_0$, find set N_0 of K nearest neighbors

The prediction is a weighted average

$$\vec{Y}_0 = \frac{1}{K} \sum_{\vec{X}_i \in \mathcal{N}_0} \frac{w_i \cdot \vec{Y}_i}{W}, \quad \text{where } W = \sum_{\vec{X}_i \in \mathcal{N}_0} w_i$$

and $w_i = 1/d_i$ where d_i is the distance between $\vec{X}_0$ and $\vec{X}_i$.

Hyperparameters: K and w_i
(tuned to optimal choices using a validation set)

Soultanis, Maltsev, Bauswein, Chatziioannou, Roepke, Stergioulas (2025)

K-NEAREST NEIGHBOR REGRESSION IN THE TIME DOMAIN

Noise-weighted inner product for two signals:

$$\langle h_1(t), h_2(t) \rangle \equiv 4\,\mathrm{Re} \int_0^\infty df \, \frac{\tilde{h}_1(f) \cdot \tilde{h}_2^*(f)}{S_h(f)}$$

Overlap:

$$\mathcal{O} \equiv \frac{\langle h_1(t), h_2(t) \rangle}{\sqrt{\langle h_1(t), h_1(t) \rangle \, \langle h_2(t), h_2(t) \rangle}}$$

Faithfulness (maximized overlap)

$$\mathcal{F} \equiv \max_{\phi_0, t_0} \frac{\langle h_1(t), h_2(t) \rangle}{\sqrt{\langle h_1(t), h_1(t) \rangle \, \langle h_2(t), h_2(t) \rangle}}$$

K-NEAREST NEIGHBOR REGRESSION IN THE TIME DOMAIN

Noise-weighted inner product for two signals:

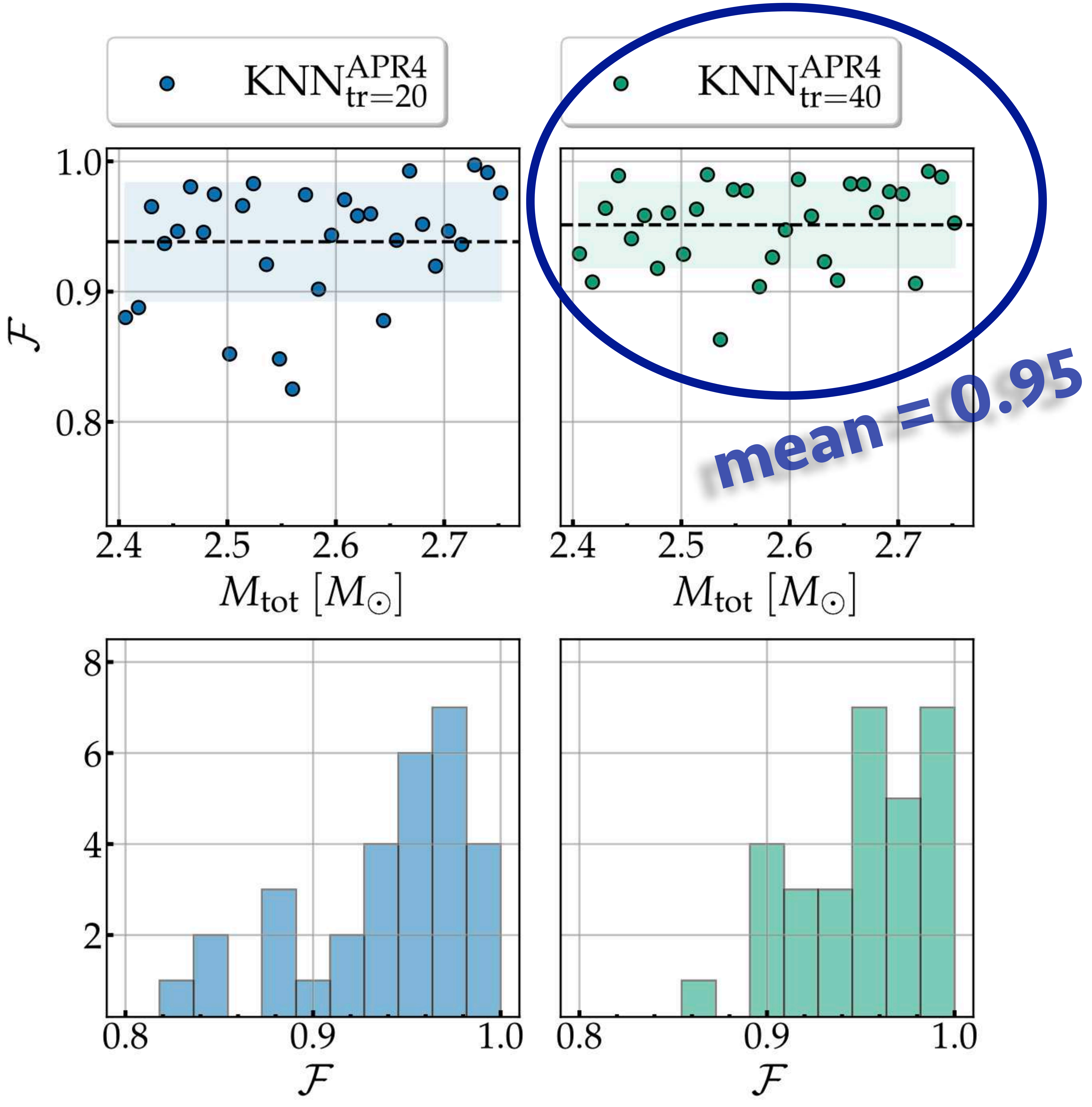
$$\langle h_1(t), h_2(t)\rangle \equiv 4\,\mathrm{Re}\int_0^\infty df\,\frac{\tilde{h}_1(f)\cdot\tilde{h}_2^*(f)}{S_h(f)}$$

Overlap:

$$\mathcal{O} \equiv \frac{\langle h_1(t), h_2(t)\rangle}{\sqrt{\langle h_1(t), h_1(t)\rangle\,\langle h_2(t), h_2(t)\rangle}}$$

Faithfulness (maximized overlap)

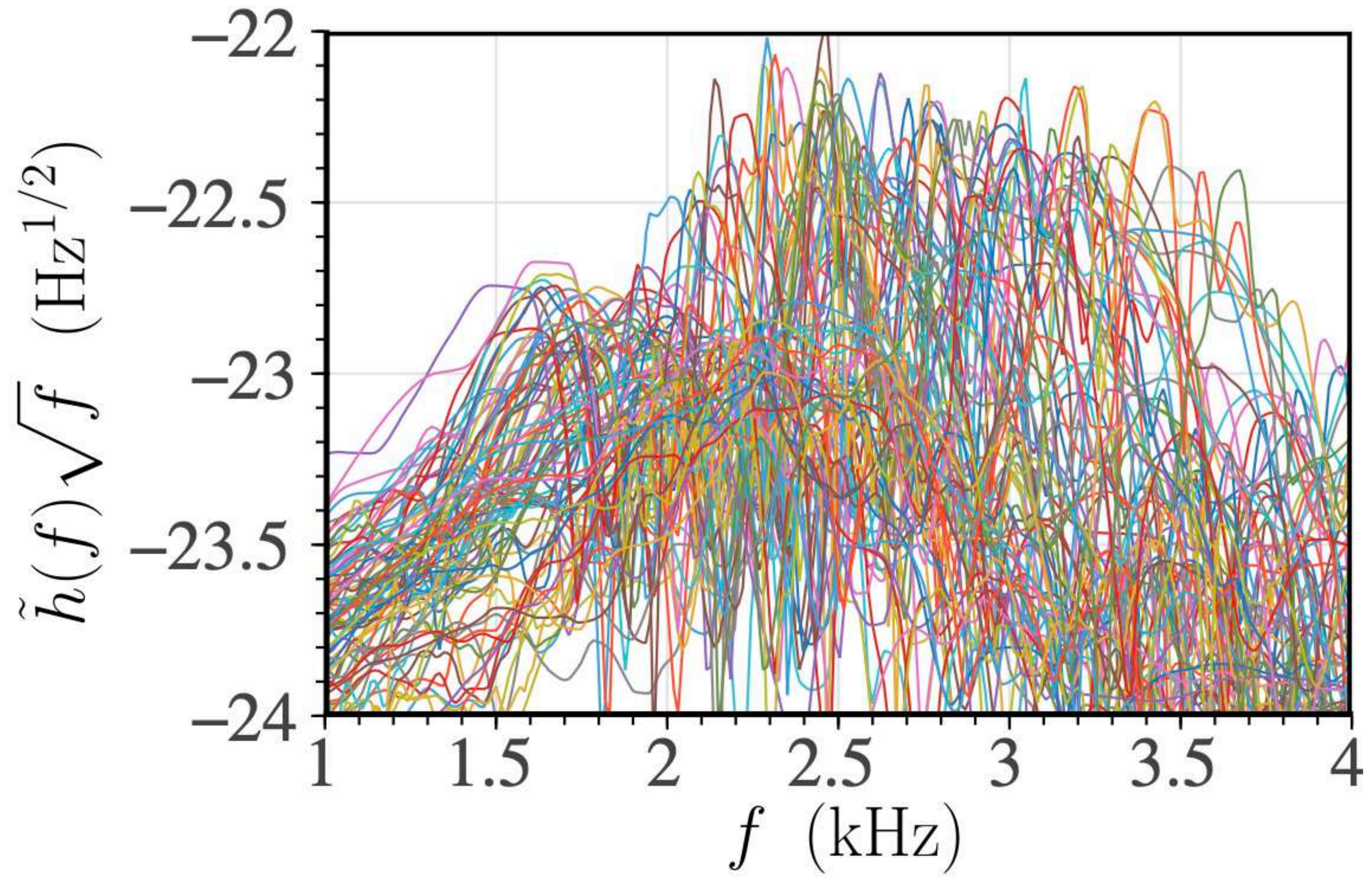
$$\mathcal{F} \equiv \max_{\phi_0, t_0} \frac{\langle h_1(t), h_2(t)\rangle}{\sqrt{\langle h_1(t), h_1(t)\rangle\,\langle h_2(t), h_2(t)\rangle}}$$



Soultanis et al. (2025)

Expanded training set: 87 equal-mass models using 14 different EOS

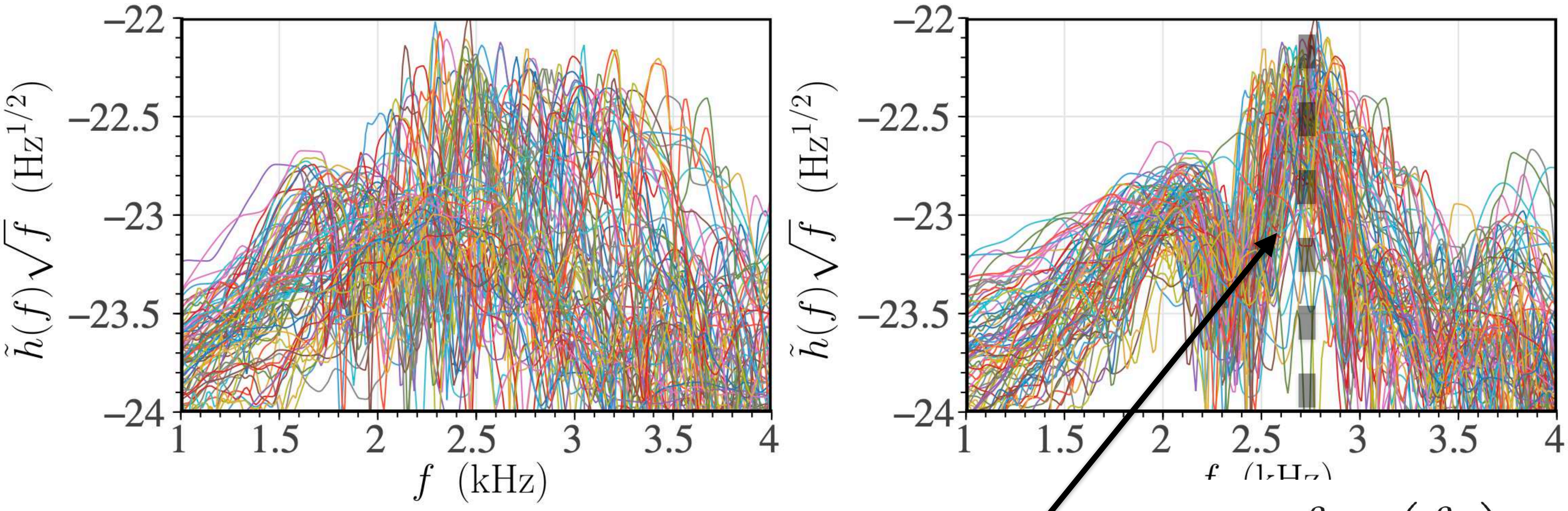
Surrogate model now depends on both mass and tidal deformability.



ANN REGRESSION IN THE FREQUENCY DOMAIN

Expanded training set: 87 equal-mass models using 14 different EOS

Surrogate model now depends on both mass and tidal deformability.



Partial alignment of spectra using empirical relation: $$f_{\text{peak}}(\kappa_2^\tau, M) = 4\frac{\beta_1}{M}\ln\left(\frac{\beta_0}{8\kappa_2^\tau}\right)$$

Pesios, Koutalios, Kugiumtzis, Stergioulas (2024)

ARTIFICIAL NEURON

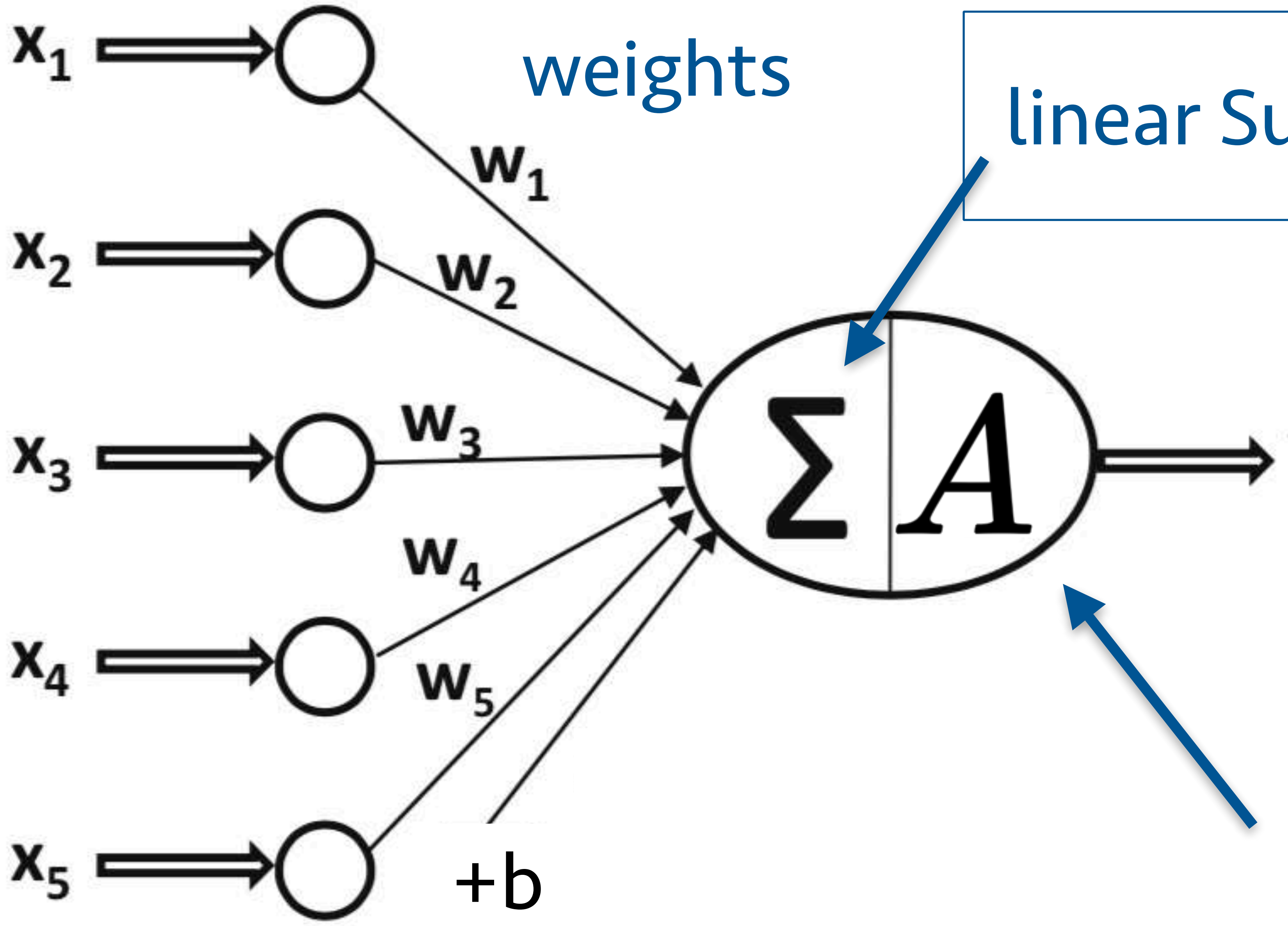
- Each neuron in the network maps several input values $x_1, .., x_n$ to an output value y

weights

linear Sum $\Sigma = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + b$

bias

input



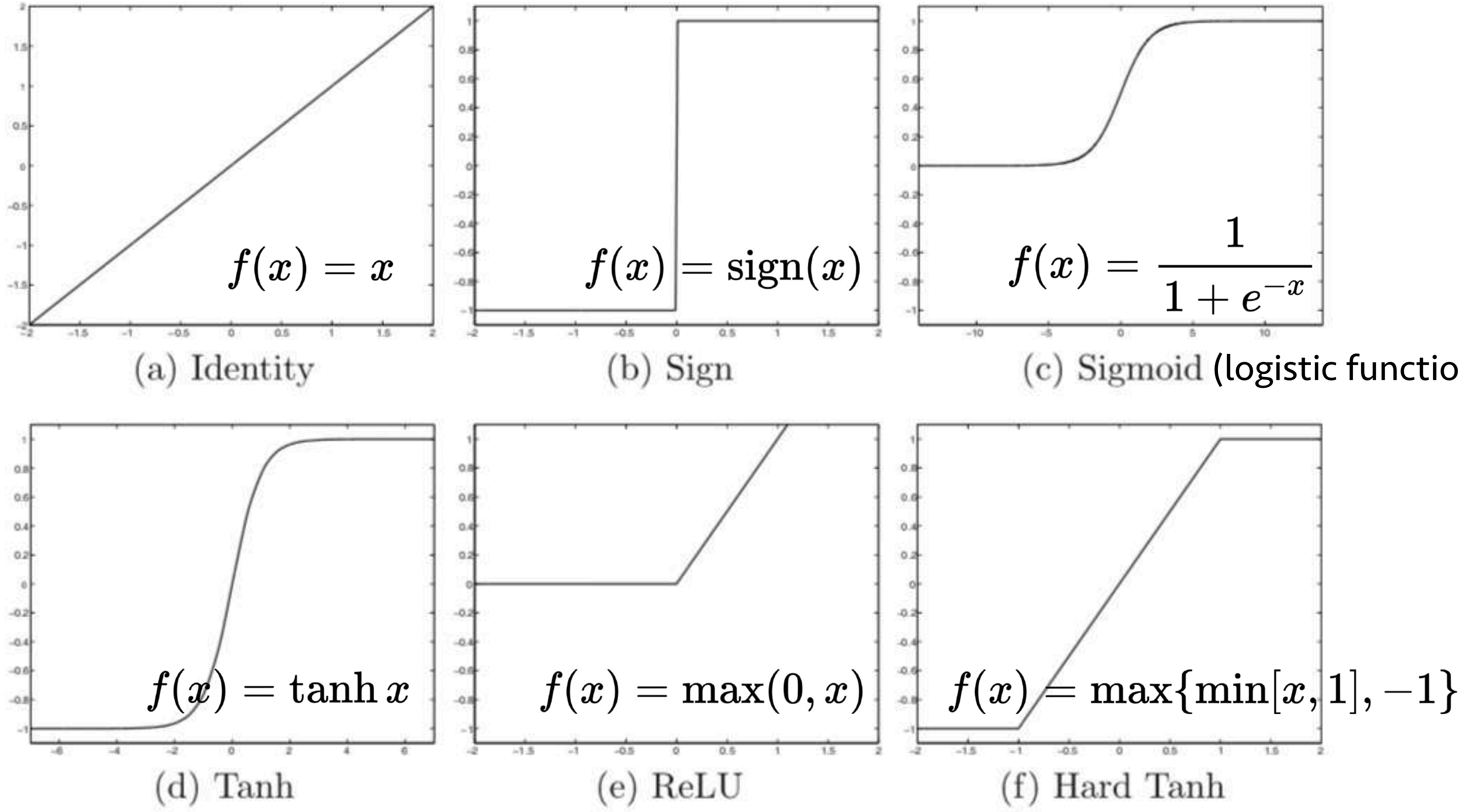
$y = A(\Sigma)$

output

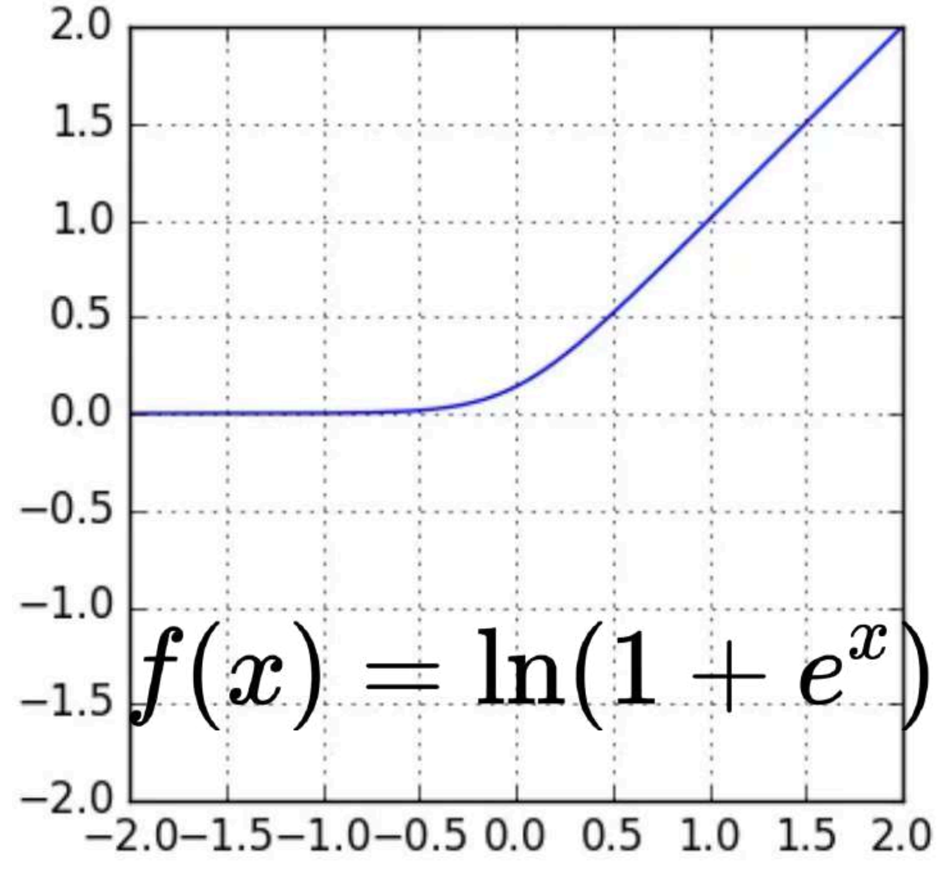
(nonlinear) Activation function

ACTIVATION FUNCTIONS

- Different common activation functions:



(Rectified Linear Unit)



Softplus

ANN SURROGATE IN THE FREQUENCY DOMAIN

Input features:

1) Mass, 2) tidal coupling constant, 3) dR/dM

Prediction:

Magnitude of GW spectrum (1-4 kHz).

Pesios et al. (2024)

ANN SURROGATE IN THE FREQUENCY DOMAIN

Input features:

1) Mass, 2) tidal coupling constant, 3) dR/dM

Prediction:

Magnitude of GW spectrum (1-4 kHz).

4-layer feed-forward ANN

Added Gaussian noise and dropout layers

Adam optimizer

Layer	Type	Shape	Activation	Params
1	Gaussian noise (0.1)	(None, 3)	...	0
2	Dense	(None, 200)	Linear	800
3	Gaussian noise (0.05)	(None, 200)	...	0
4	Dropout (0.15)	(None, 200)	...	0
5	Dense	(None, 400)	Sigmoid	80400
6	Gaussian noise (0.1)	(None, 400)	...	0
7	Dropout (0.15)	(None, 400)	...	0
8	Dense	(None, 400)	Sigmoid	160400
9	Gaussian noise (0.1)	(None, 400)	...	0
10	Dropout (0.05)	(None, 400)	...	0
11	Dense	(None, 370)	Linear	148370

Pesios et al. (2024)

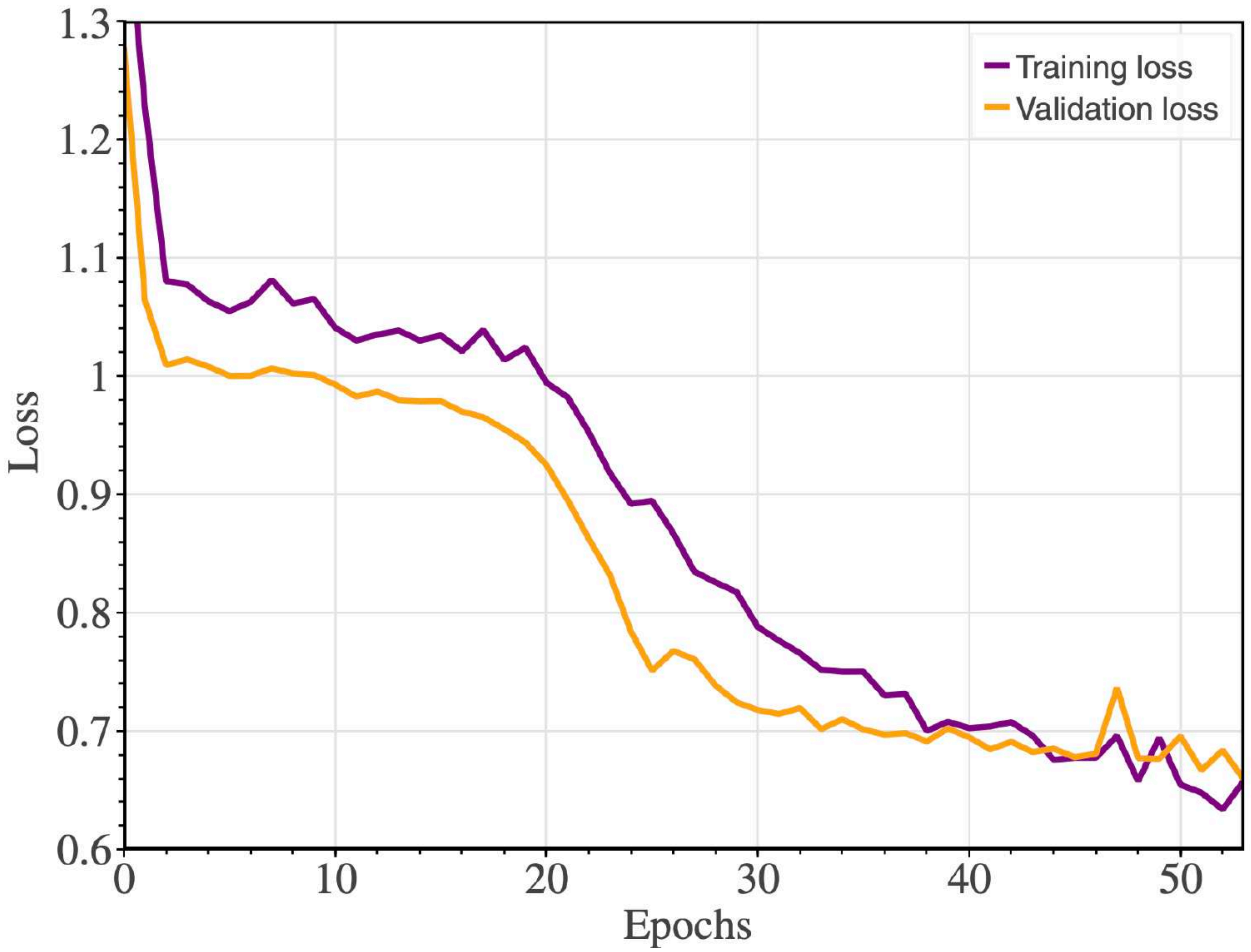
ANN SURROGATE IN THE FREQUENCY DOMAIN

Input features:

1) Mass, 2) tidal coupling constant, 3) dR/dM

Prediction:

Magnitude of GW spectrum (1-4 kHz).



4-layer feed-forward ANN

Added Gaussian noise and dropout layers

Adam optimizer

Layer	Type	Shape	Activation	Params
1	Gaussian noise (0.1)	(None, 3)	···	0
2	Dense	(None, 200)	Linear	800
3	Gaussian noise (0.05)	(None, 200)	···	0
4	Dropout (0.15)	(None, 200)	···	0
5	Dense	(None, 400)	Sigmoid	80400
6	Gaussian noise (0.1)	(None, 400)	···	0
7	Dropout (0.15)	(None, 400)	···	0
8	Dense	(None, 400)	Sigmoid	160400
9	Gaussian noise (0.1)	(None, 400)	···	0
10	Dropout (0.05)	(None, 400)	···	0
11	Dense	(None, 370)	Linear	148370

Pesios et al. (2024)

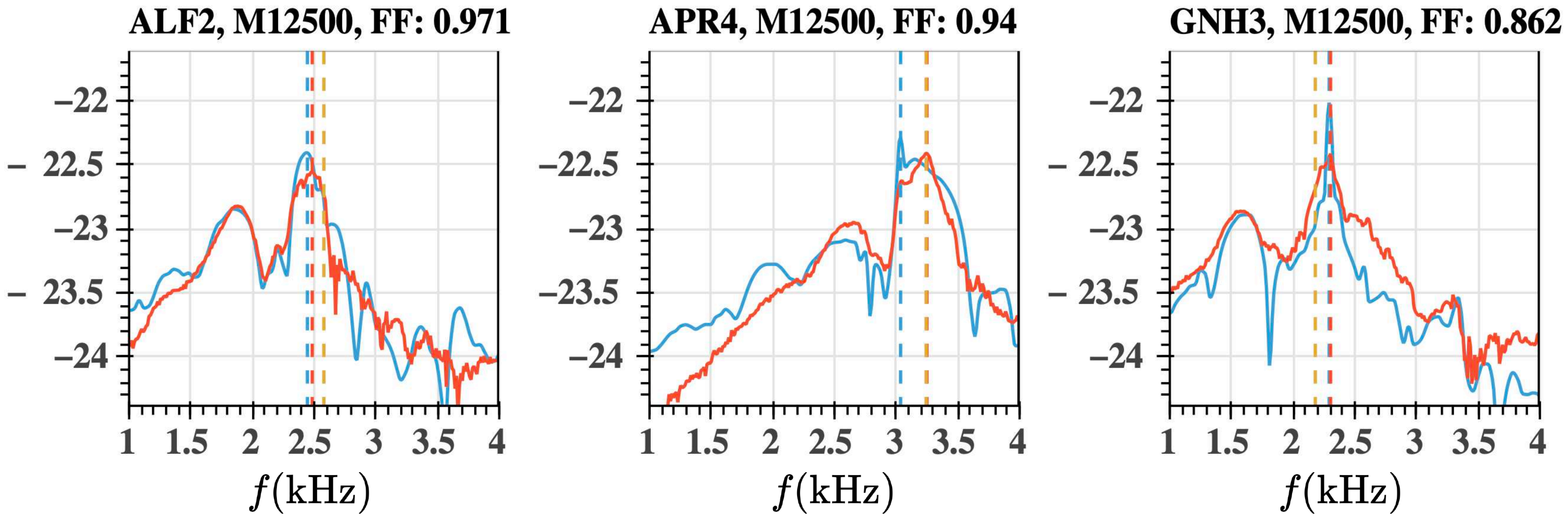
ANN SURROGATE IN THE FREQUENCY DOMAIN

Typical examples of predicted magnitudes of GW spectra

FF = O = overlap

Impact of uncertainty in empirical relation offset by re-calibration of spectra.

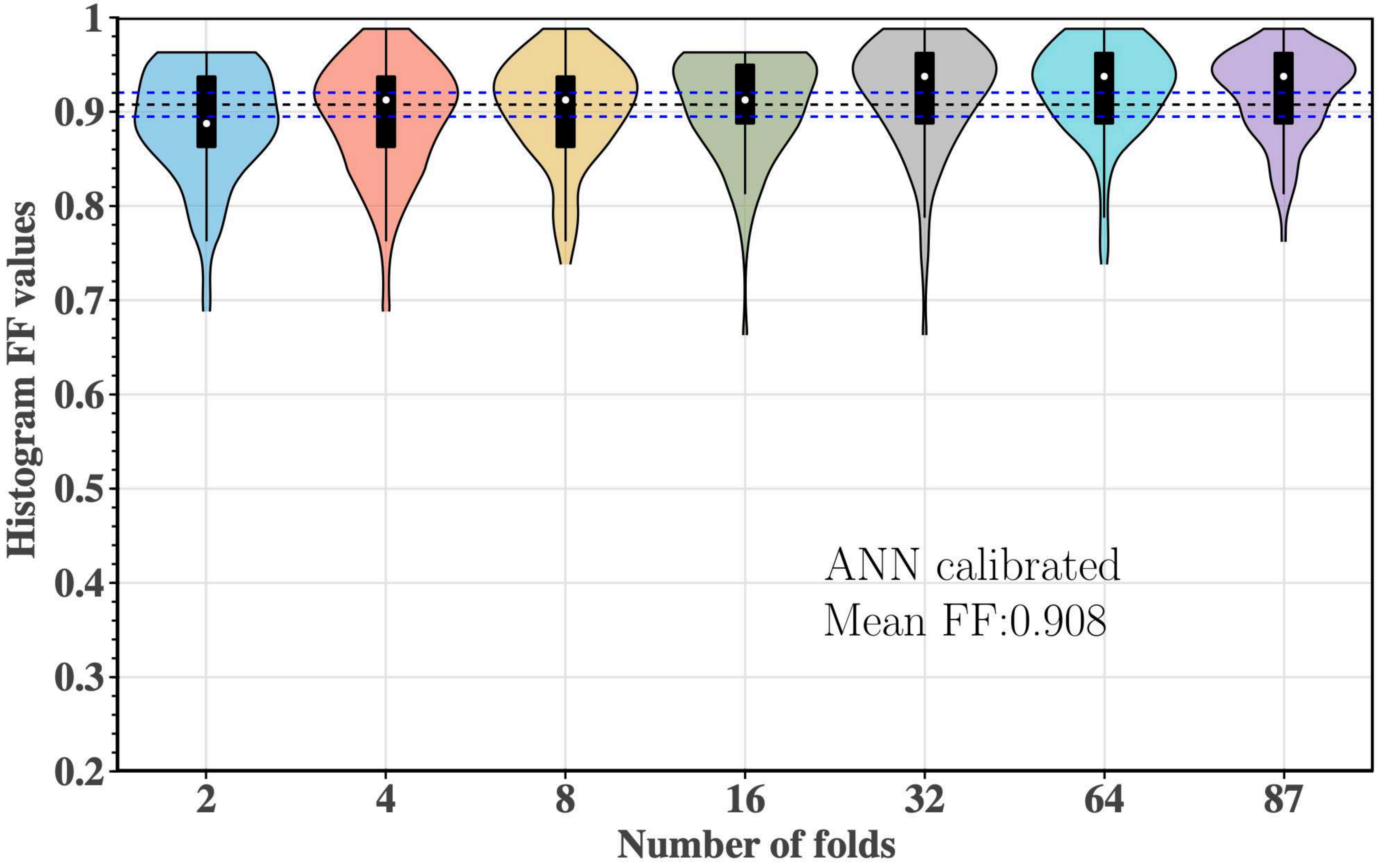
ANN outperforms multivariate linear regression.



Pesios et al. (2024)

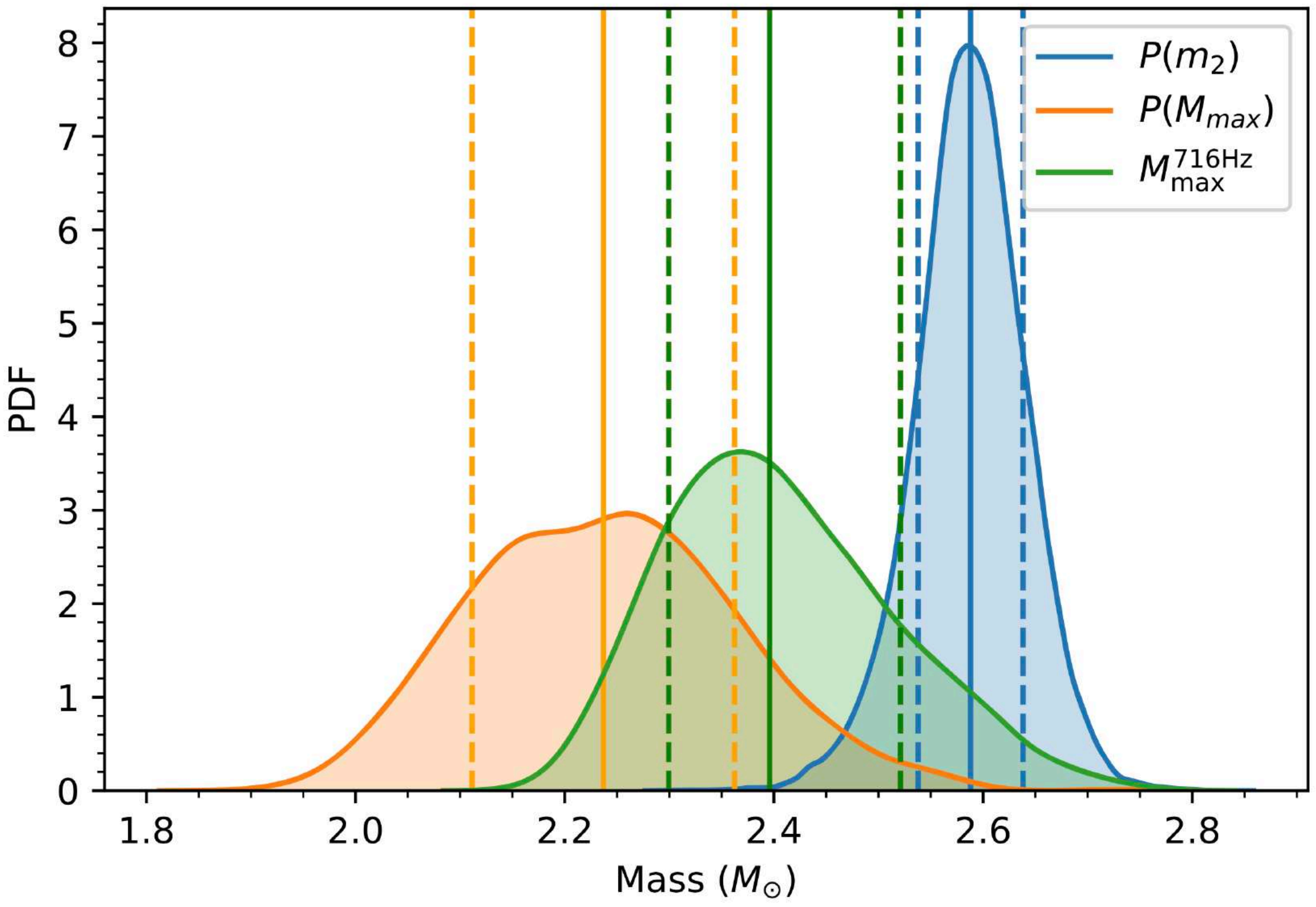
ANN SURROGATE IN THE FREQUENCY DOMAIN

Cross-validation study of fitting factors distribution:

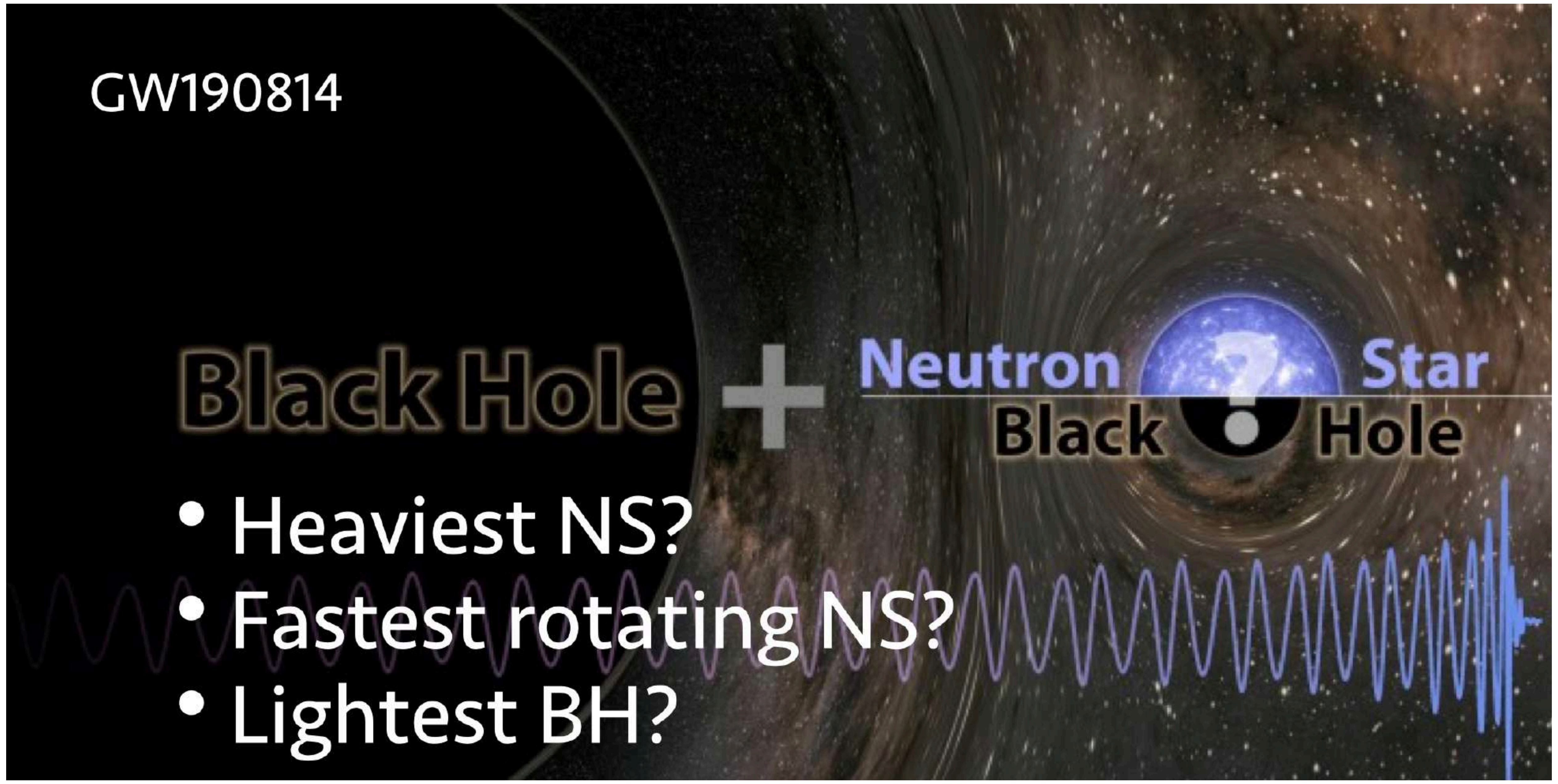


PART C.
NEUTRON STAR MODELS IN ALTERNATIVE THEORIES OF GRAVITY

What was the nature of the lighter component in GW190814?

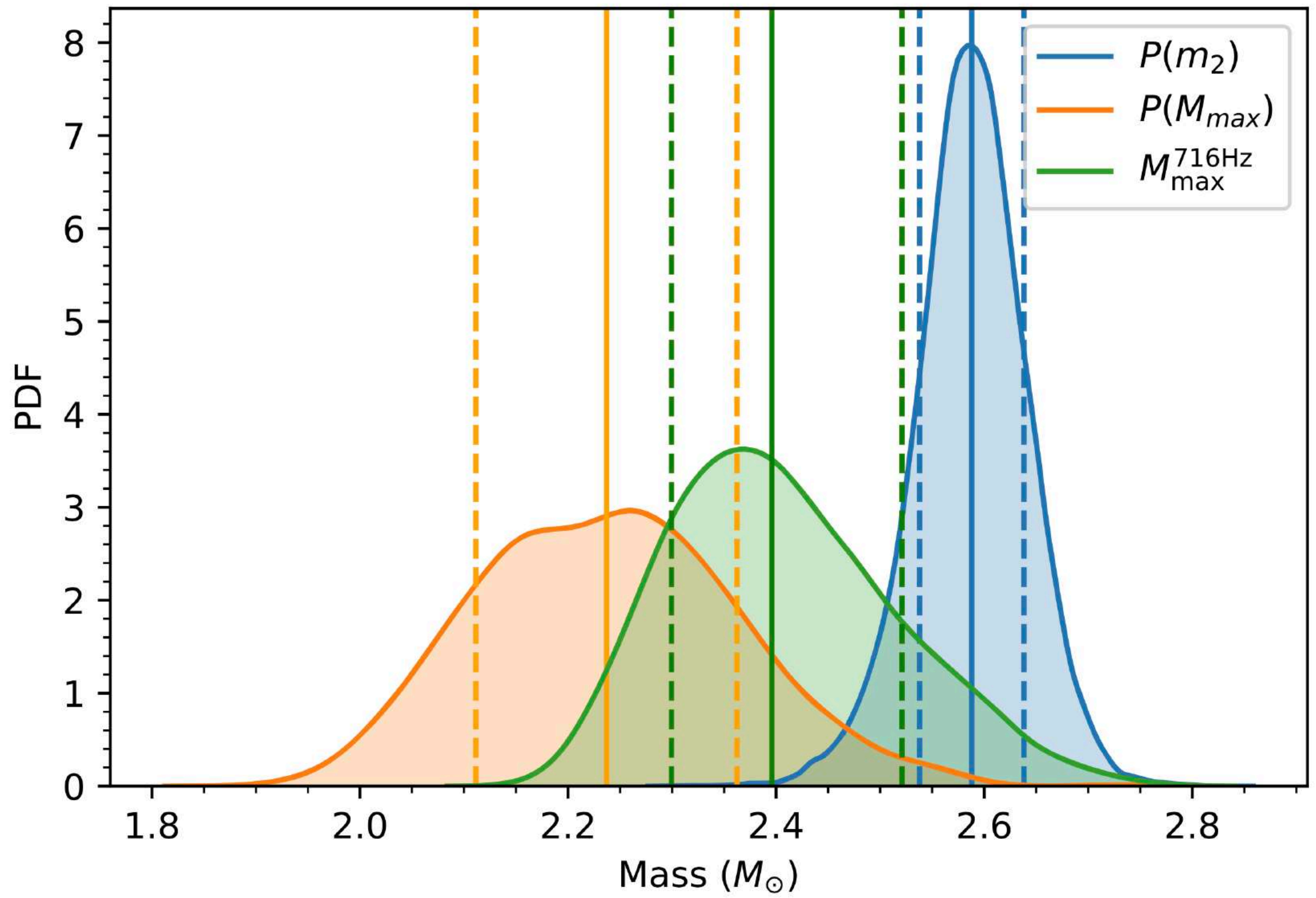


Biswas, Nandi, Char, Bose, Stergioulas (2021)

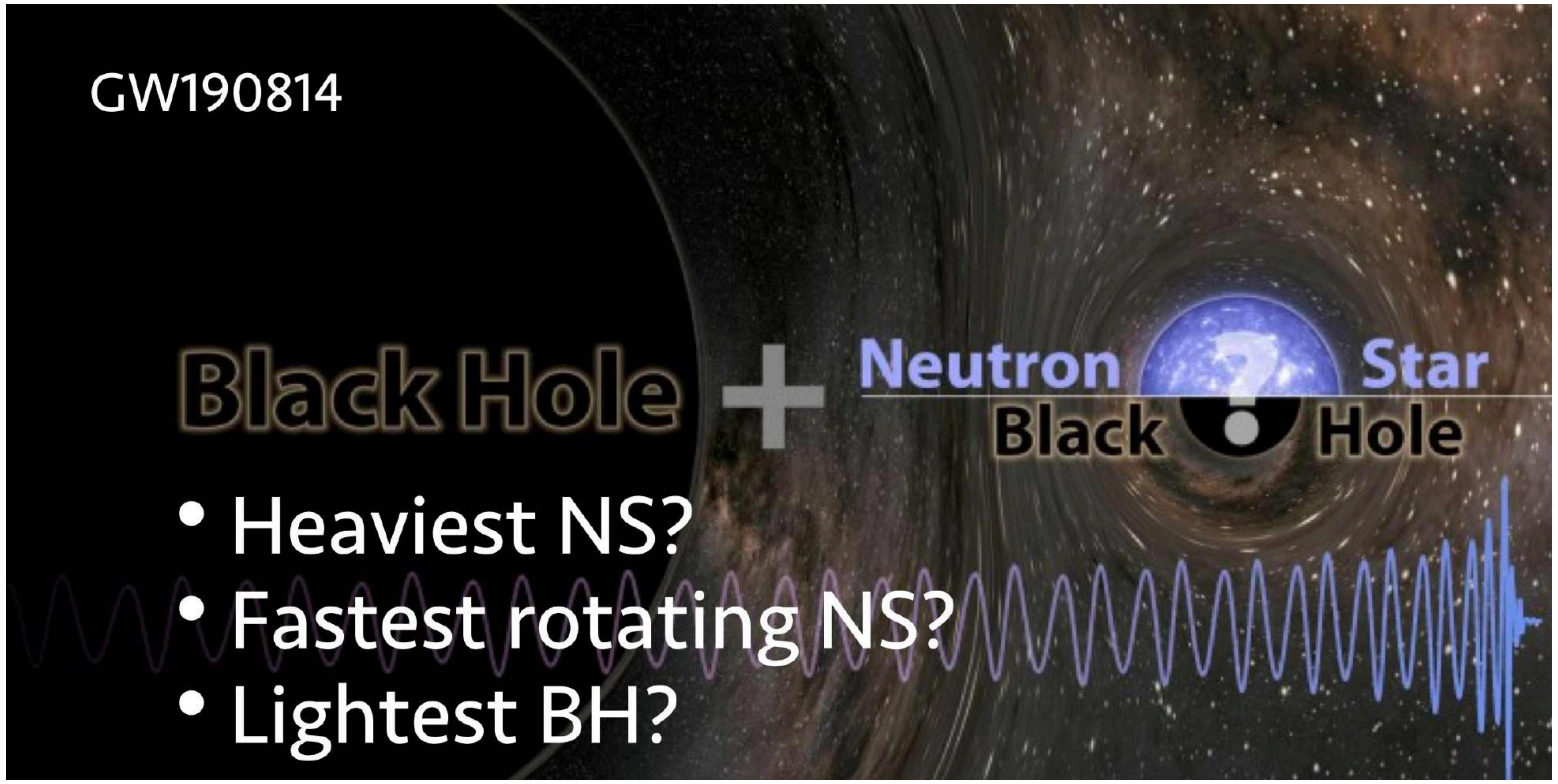


HIGH MASS NEUTRON STARS?

What was the nature of the lighter component in GW190814?



Biswas, Nandi, Char, Bose, Stergioulas (2021)



<u>Another possibility</u>: if the correct theory of gravity is not GR, then heavier NS might exist!

How can this degeneracy be resolved?

NS IN 4D EINSTEIN-GAUSS-BONNET GRAVITY

We consider the action

$$S=\frac{1}{2\kappa}\int \mathrm{d}^4x\sqrt{-g}\{R+\alpha[\phi\mathcal{G}+4G_{\mu\nu}\nabla^\mu\phi\nabla^\nu\phi-4(\nabla\phi)^2\Box\phi+2(\nabla\phi)^4]\}+S_\mathrm{m}$$

where $\kappa=8\pi G/c^4, \mathcal{G}=R^2-4R_{\mu\nu}R^{\mu\nu}+R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$ is the Gauss-Bonnet scalar and S_m is the matter Lagrangian.

This theory possesses an exact vacuum solution describing nonrotating, static black holes with line element

$$\mathrm{d}s^2=-h(r)\mathrm{d}t^2+\frac{\mathrm{d}r^2}{f(r)}+r^2\,\mathrm{d}\Omega^2$$

where

$$h(r)=f(r)=1+\tfrac{r^2}{2\alpha}\left(1-\sqrt{1+\tfrac{8\alpha M}{r^3}}\right)$$

(M ->ADM mass)

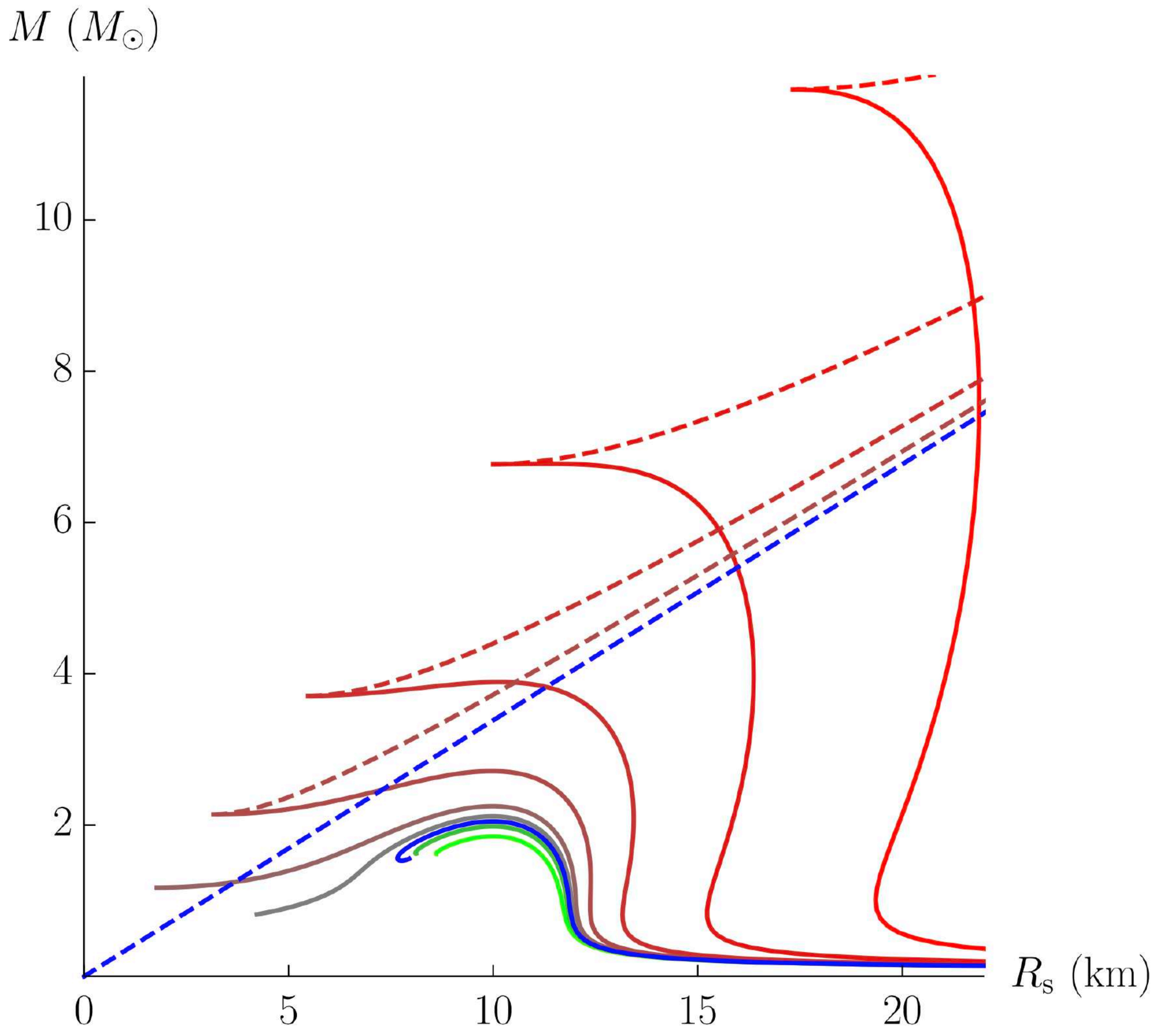
and the shift-symmetric scalar field is

$$\phi(r)=\int\mathrm{d}r\frac{\sqrt{f}-1}{r\sqrt{f}}$$

Charmousis, Lehebel, Smyrniotis, Stergioulas (2022)

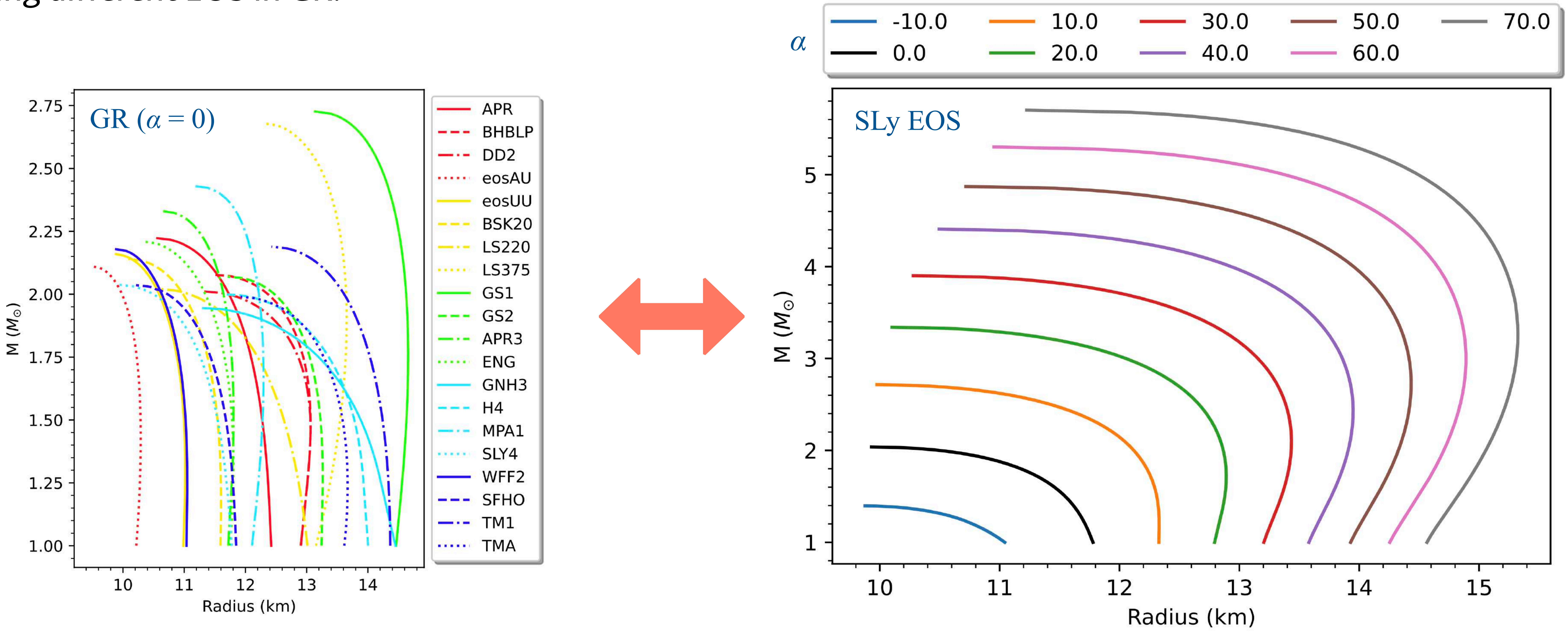
SEQUENCES OF EQUILIBRIUM MODELS

For the **SLy EOS**, we show representative cases of sequences of NS equilibrium models (solid lines) and BHs (dashed lines). For α>0, the NS solutions merge with the minimum mass BH solution.



EOS-GRAVITY DEGENERACY

In this theory, constructing equilibrium sequences for the same EOS, but different α, mimics the effect of using different EOS in GR.



Need to use a large number of observations to break the degeneracy!

Liodis, Smyrniotis, Stergioulas (2024)

ANN SURROGATE MODELS FOR NUMERICAL SOLUTIONS

EOS collection

Name	Number #
APR	1
BHBLP	2
DD2	3
eosAU	4
eosUU	5
BSk20	6
LS220	7
LS375	8
GS1	9
GS2	10
APR3 (PP)	11
ENG (PP)	12
GNH3 (PP)	13
H4 (PP)	14
MPA1 (PP)	15
SLy4 (PP)	16
WFF2 (PP)	17
SFHo	18
TM1	19
TMA	20

Surrogate models

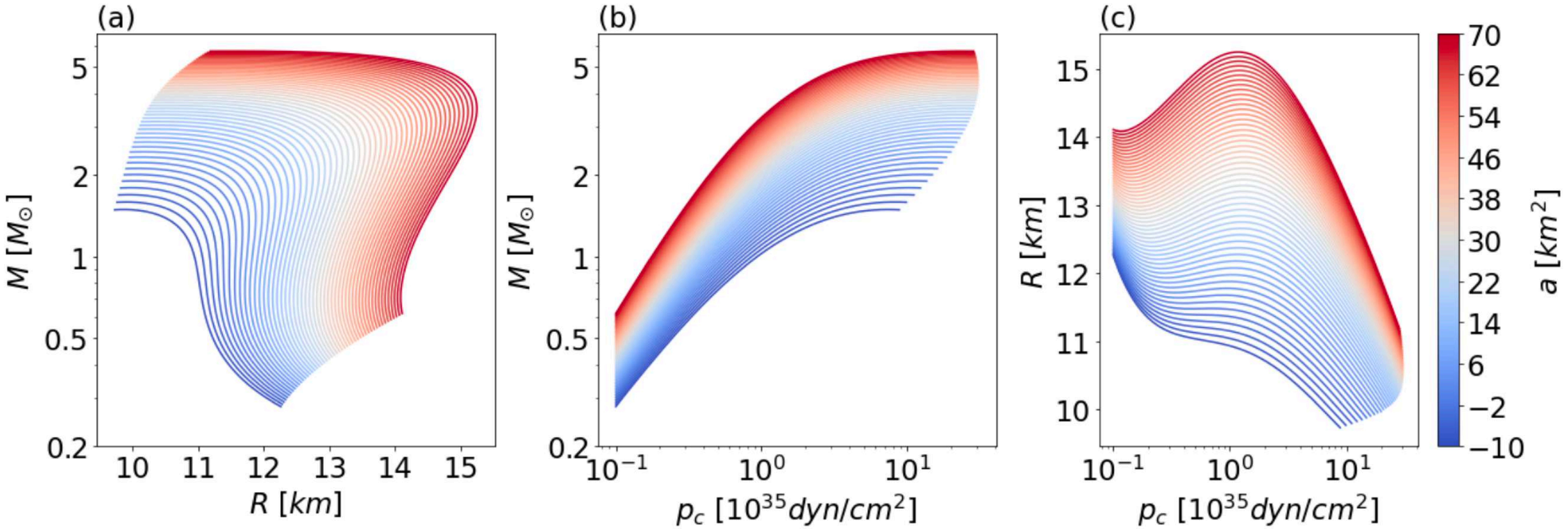
$$f_1(\mathrm{EoS}; \alpha, \mathrm{p_c}) \to (\mathrm{M}, \mathrm{R})$$

$$f_2(\mathrm{EoS}; \alpha, \mathrm{M}) \to \mathrm{R}$$

Network architecture for each EOS

Layer	Type f_1	Type f_2
Input layer	(α, p_c)	(α, M)
Hidden layer 1	25-tanh	25-tanh
Hidden layer 2	35-relu	35-relu
Hidden layer 3	25-tanh	25-tanh
Output layer	(M, R)	R

Training set for each EOS: 200 (p_c) x 51 (α) = 10200 equilibrium models



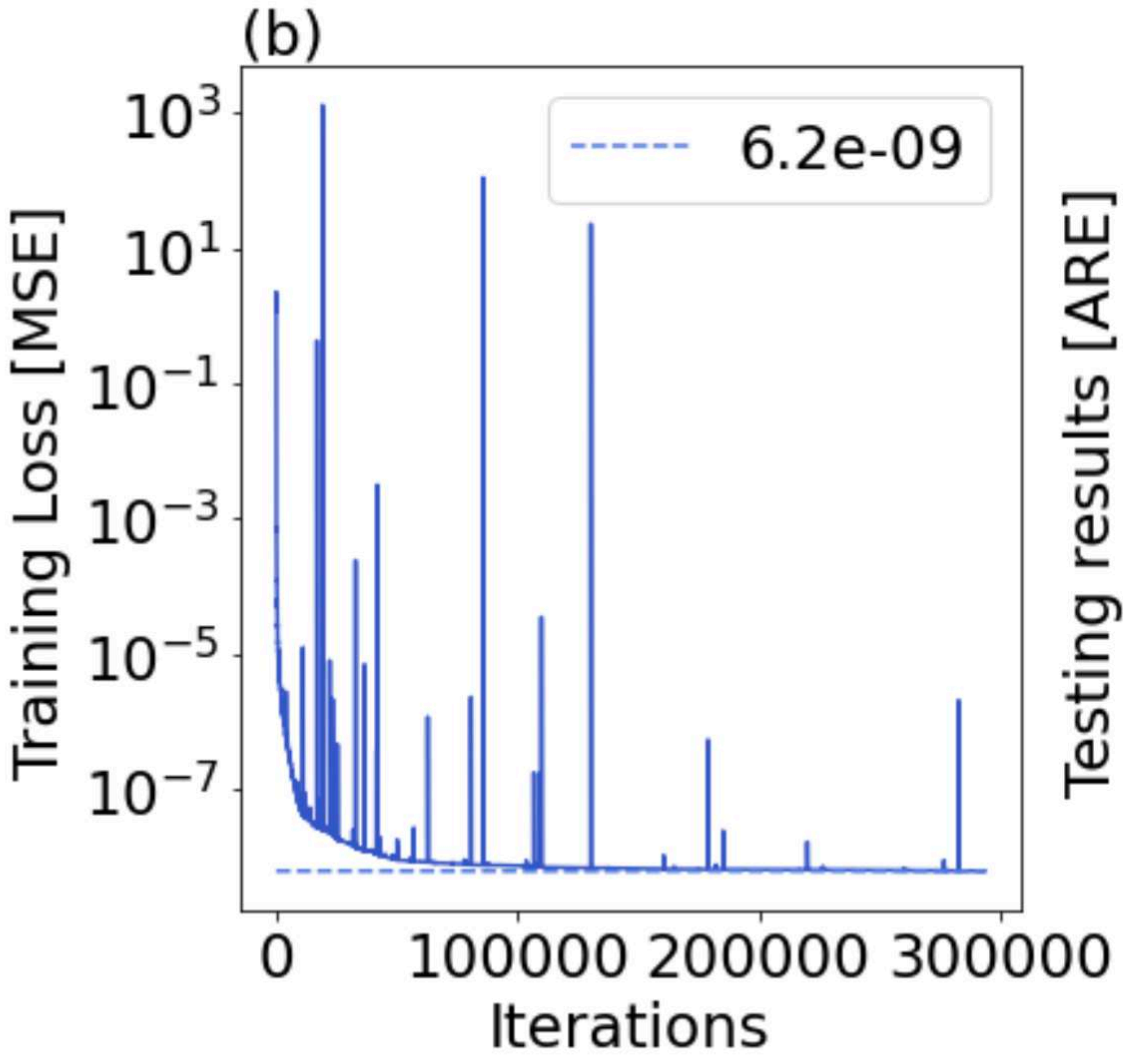
Liodis, Smyrniotis, Stergioulas (2024)

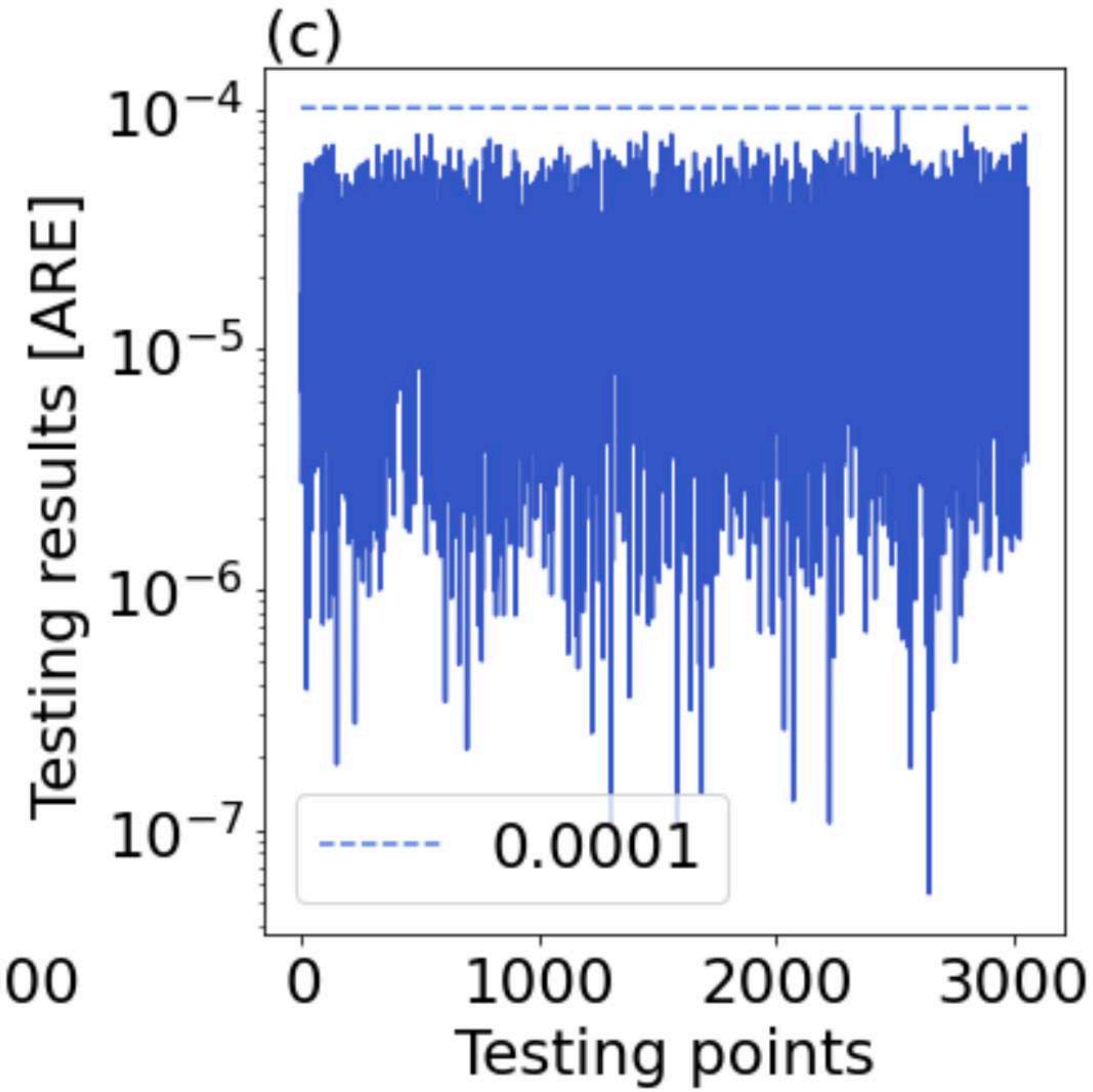
ACCURACY OF f_1 ANN SURROGATE MODEL FOR EOS BSk20

Loss function:
Mean Square Error (MSE)

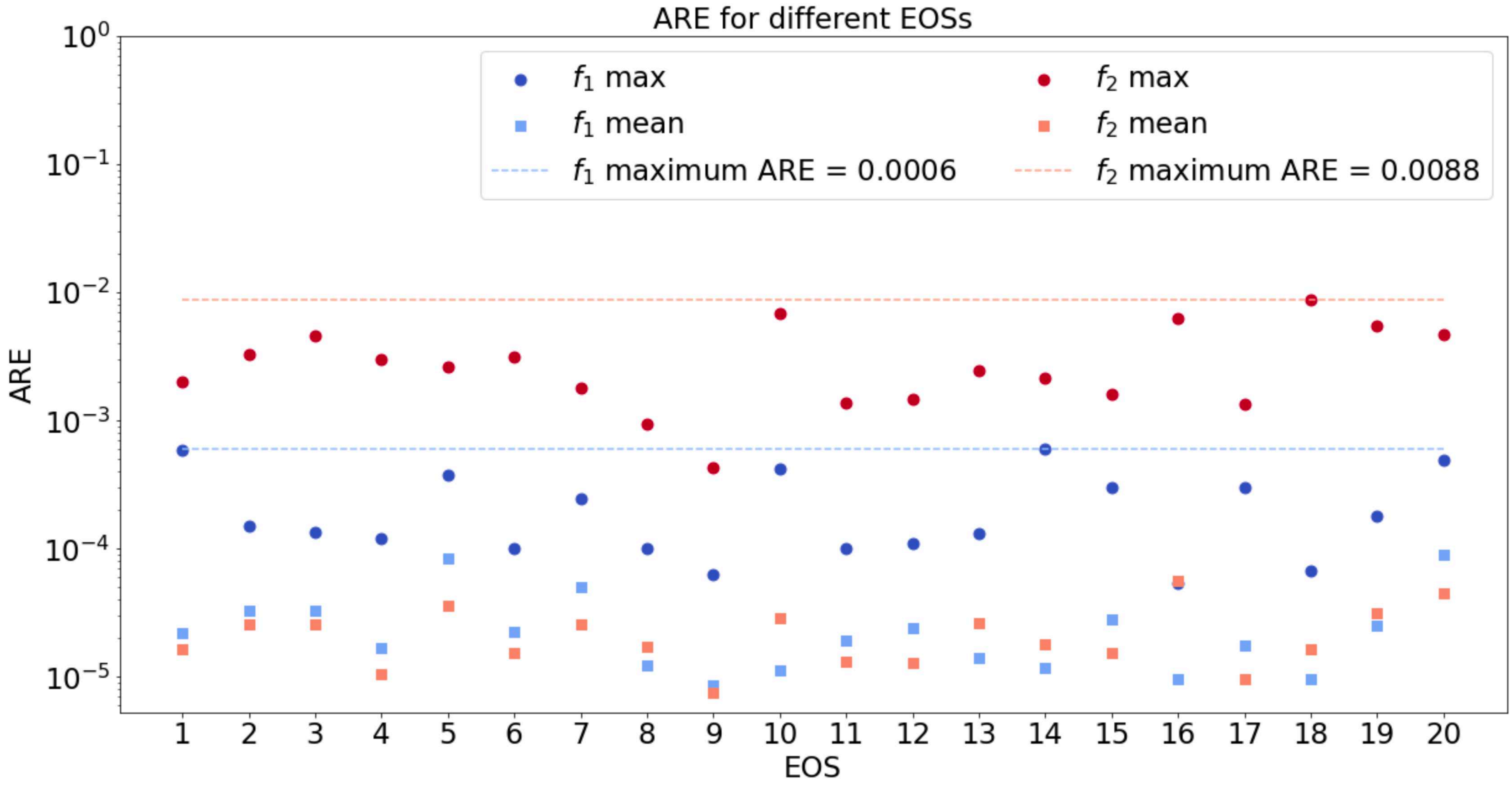
Accuracy test:
Absolute Relative Error

$$\mathrm{ARE}_i = \frac{1}{m}\sum_{j=1}^{m}\left|\frac{Y_i^j - Y_{i,\,\mathrm{true}}^j}{Y_{i,\,\mathrm{true}}^j}\right|$$





ACCURACY OF ANN SURROGATE MODELS FOR ALL EOS

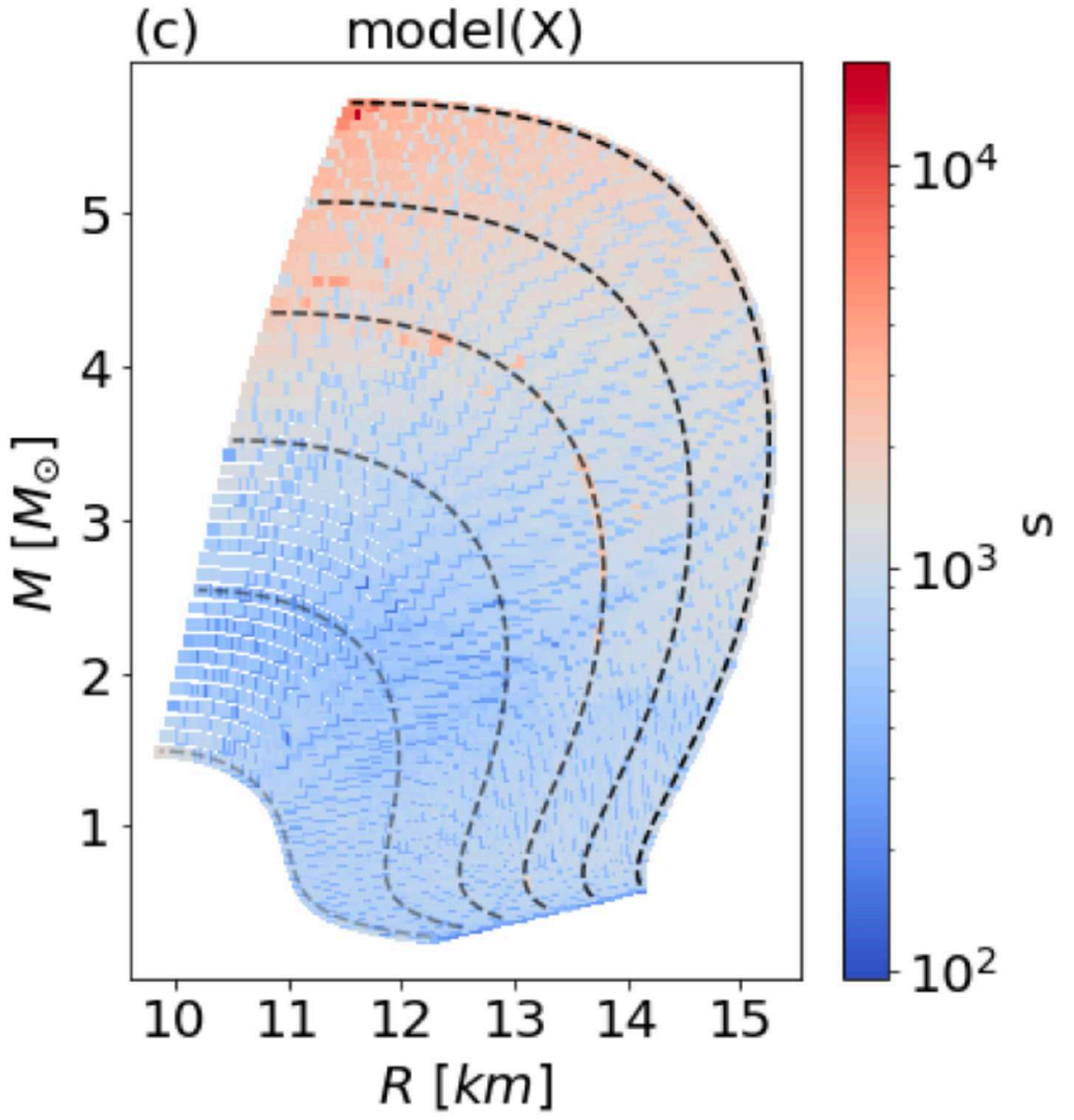


SPEED-UP ACHIEVED WITH THE ANN SURROGATE MODELS

Speed-up of ANN surrogate model compared to original numerical code

$$s = \frac{\Delta t_{\mathrm{ANN}}}{\Delta t_{\mathrm{num}}}$$

Numerical code run time:	(Mean) 1003.5 ms	(Min) 147.96 ms	(Max) 18122.4 ms
Output method	Speed Up (Mean)	Speed Up (Minimum)	Speed Up (Maximum)
model.predict(X)	25.12	0.97	464.56
model(X)	921.9	95.6	18102.1
model.predict($\mathbf{X}$)	31295.5	4614.5	565157.2



The achieved speed-up allows for millions of MCMC calls in a Bayesian inference computation with minimal overhead.

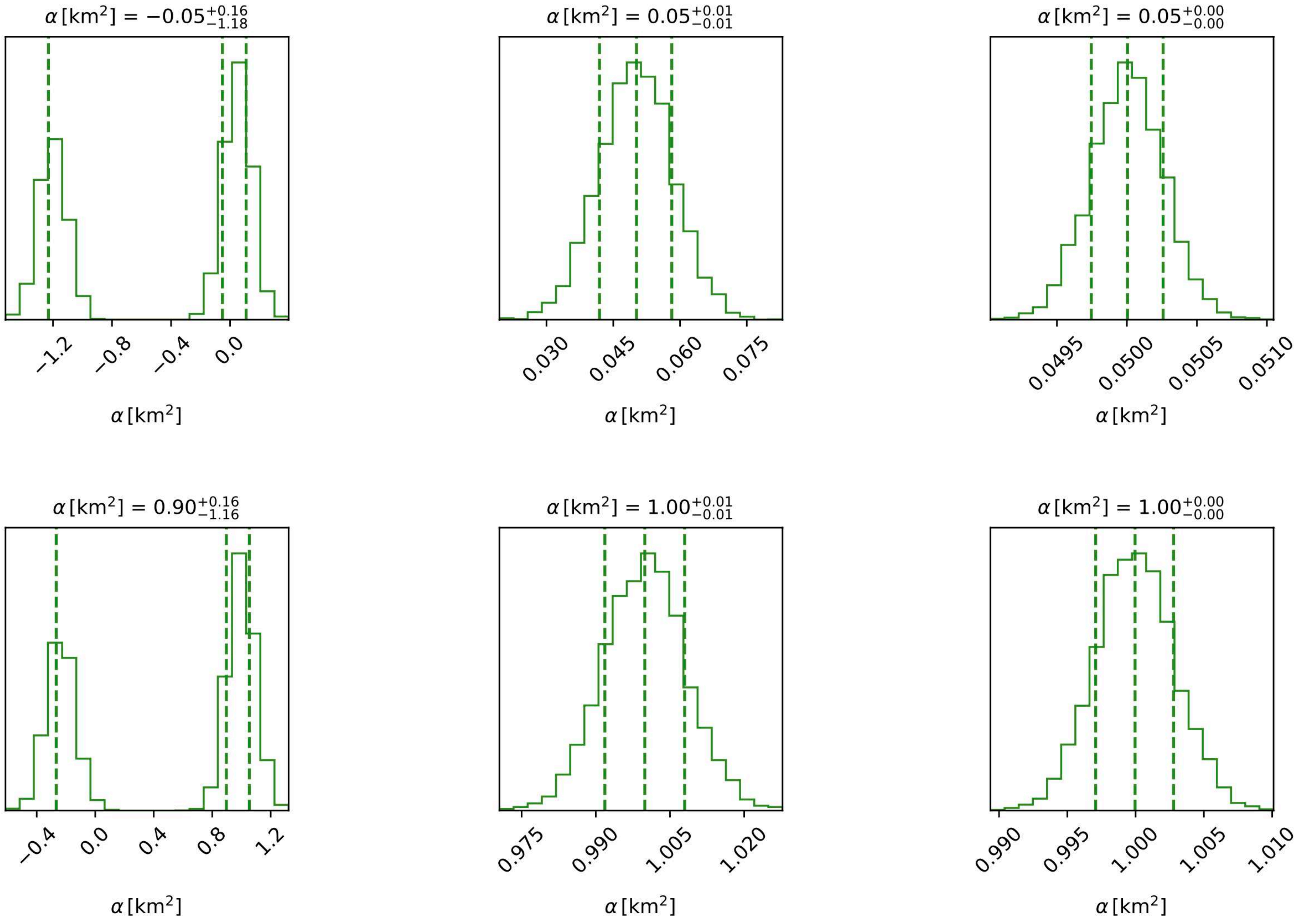
RECOVERY OF INJECTED α VALUES

Hierarchical Bayesian Inference using astrophysical constraints for a set of 20 EOS:

A+ | Voyager | 3G

$$\alpha = 0.05\,\mathrm{km}^2$$

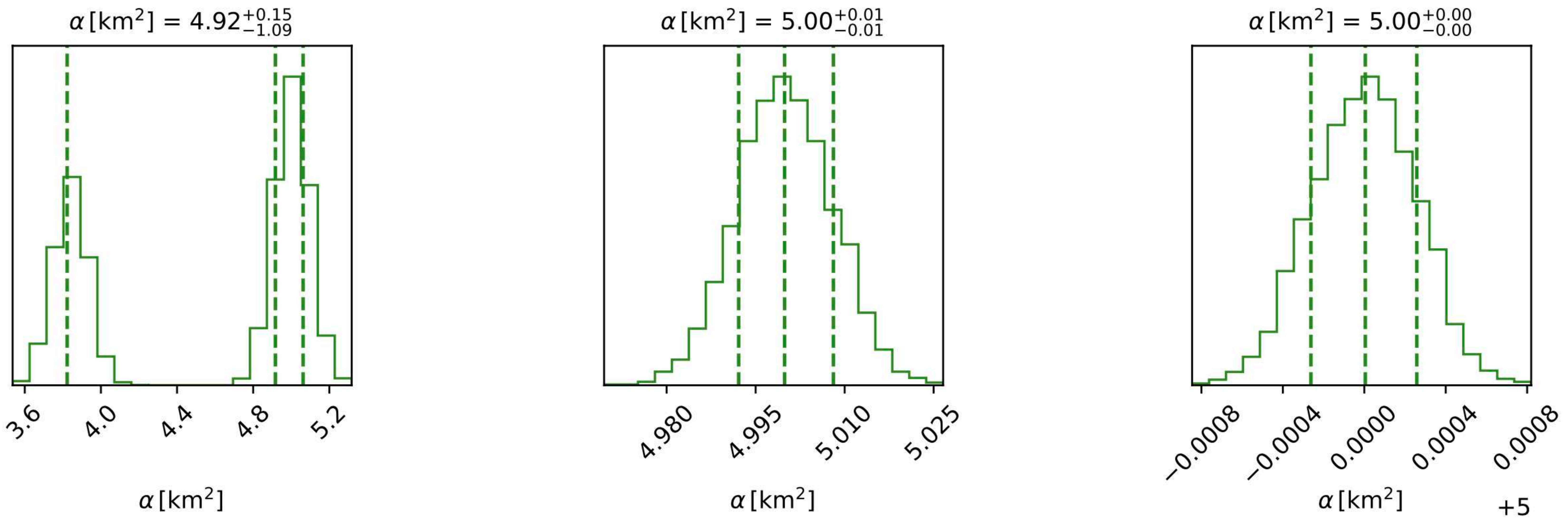
$$\alpha = 1.0\,\mathrm{km}^2$$



RECOVERY OF INJECTED α VALUES

A+ Voyager 3G

$$\alpha = 5.0\,\mathrm{km}^2$$



A+ : for an injection of $\alpha = 5$ km^2, the GR value of $\alpha = 0$ is excluded at the 3σ level.

Voyager/3G : even departures from GR at the $\alpha = 0.05$ km^2 level are accurately recovered.

Biswas, Smyrniotis, Liodis, Stergioulas (2024)

THANK YOU FOR YOUR ATTENTION