LEVERAGING MACHINE LEARNING FOR GRAVITATIONAL WAVE DISCOVERIES

From Binary Black Hole Mergers to Probing High-Density Nuclear Matter

NIKOLAOS STERGIOULAS

DEPARTMENT OF PHYSICS

ARISTOTLE UNIVERSITY OF THESSALONIKI

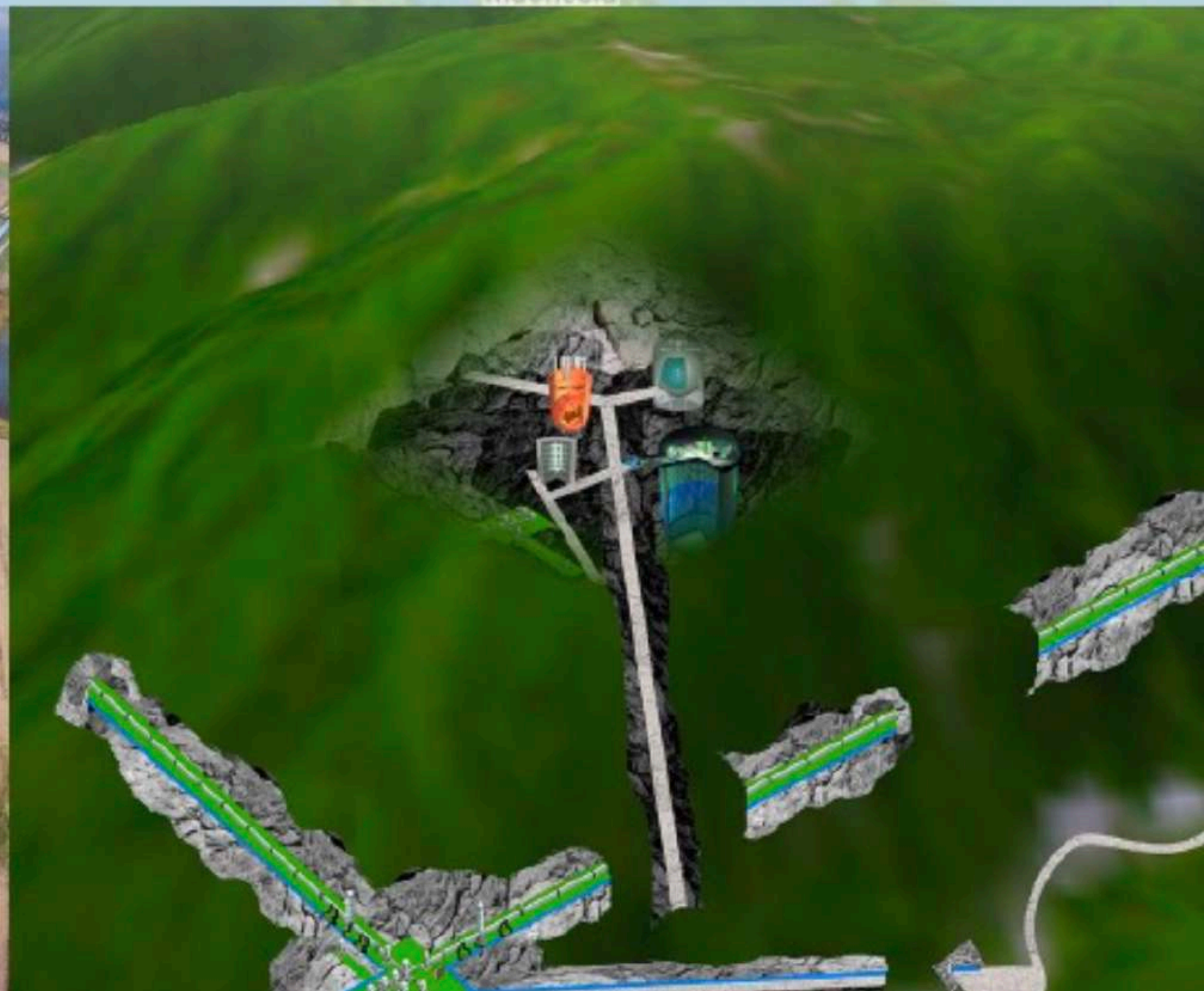
Ministry of Science and Higher Education
Republic of Poland

11th Conference of the Polish Society on Relativity, Wroclaw, July 21, 2025

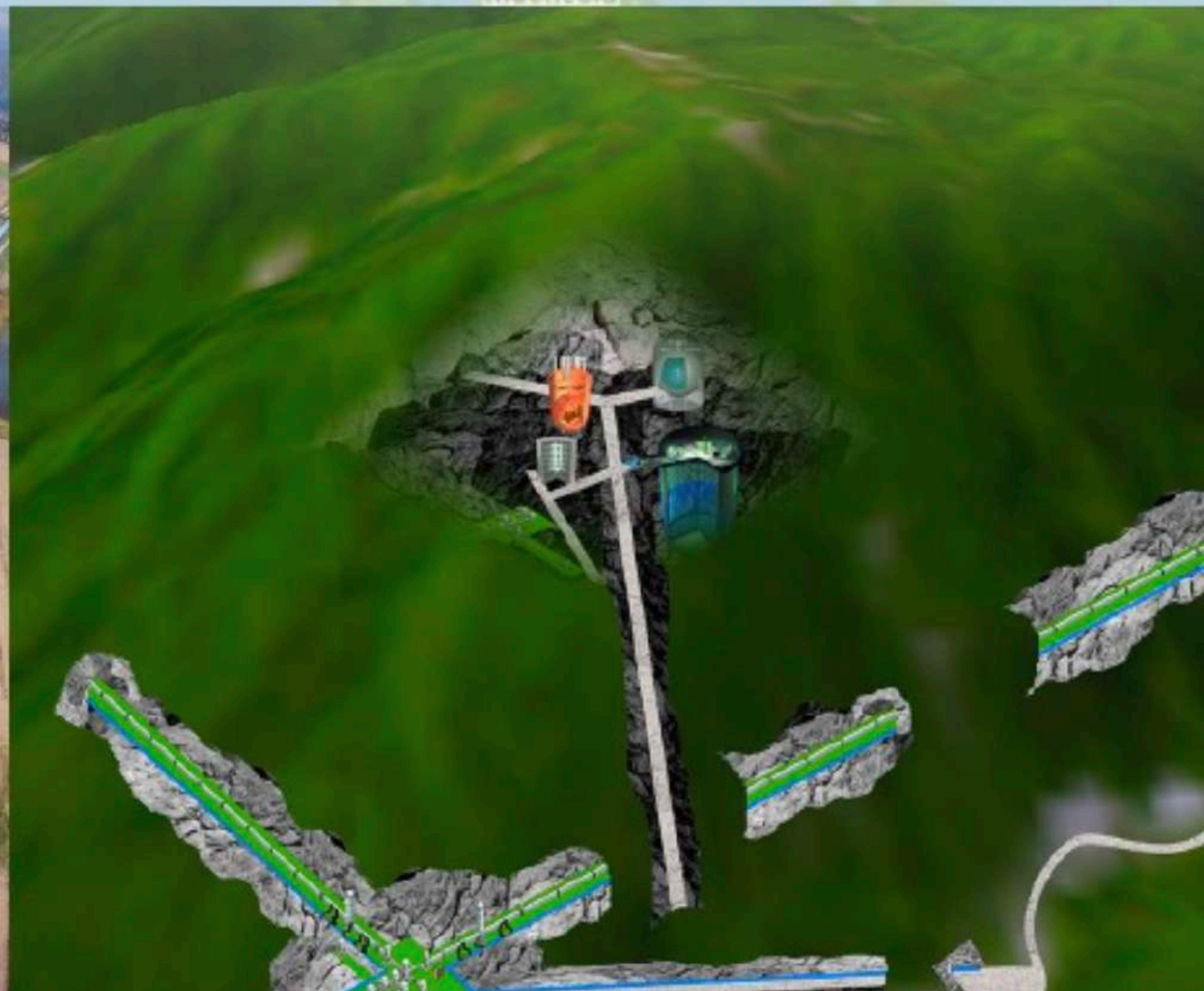
PART A. BINARY BLACK HOLE MERGERS

INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)

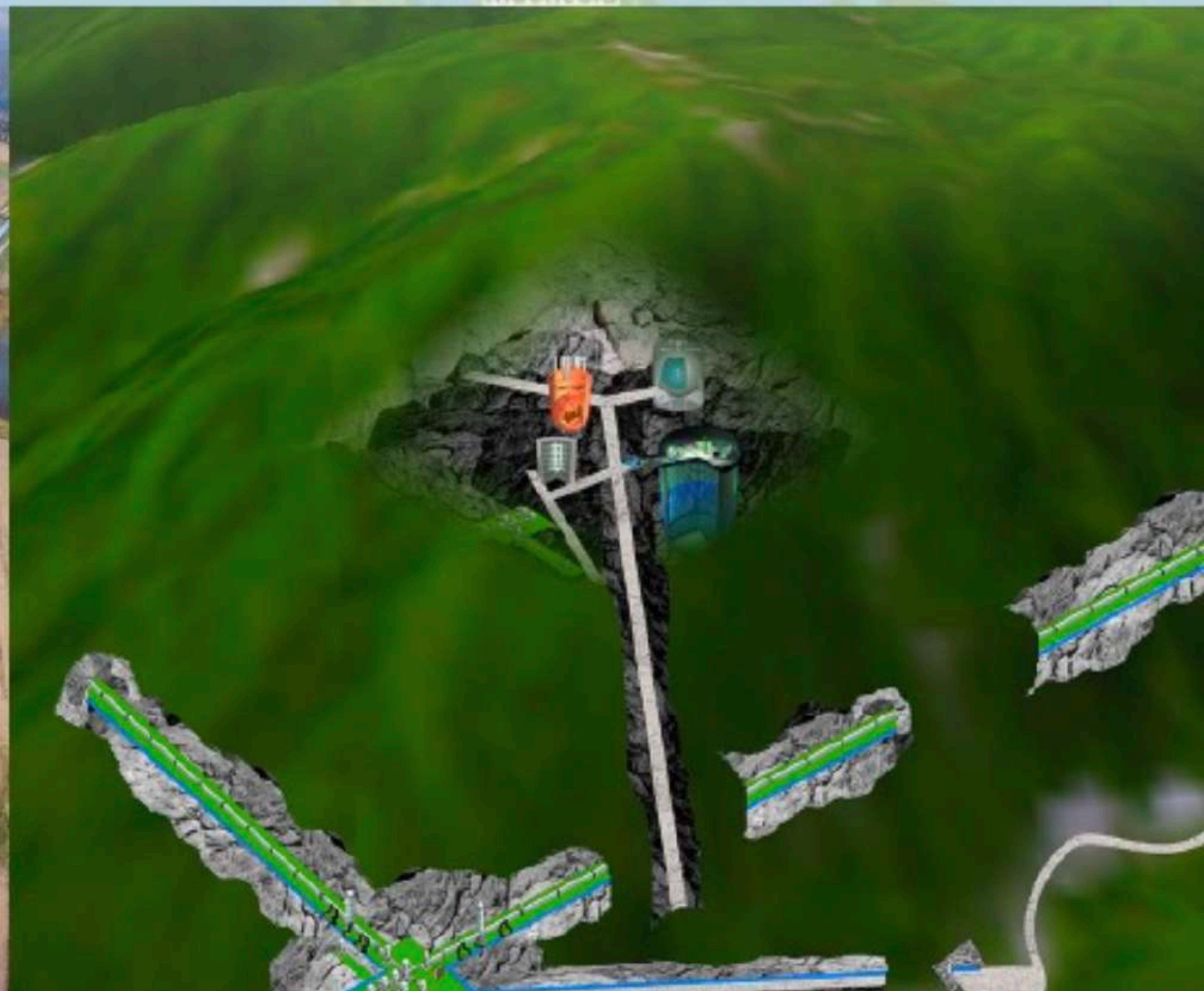
INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



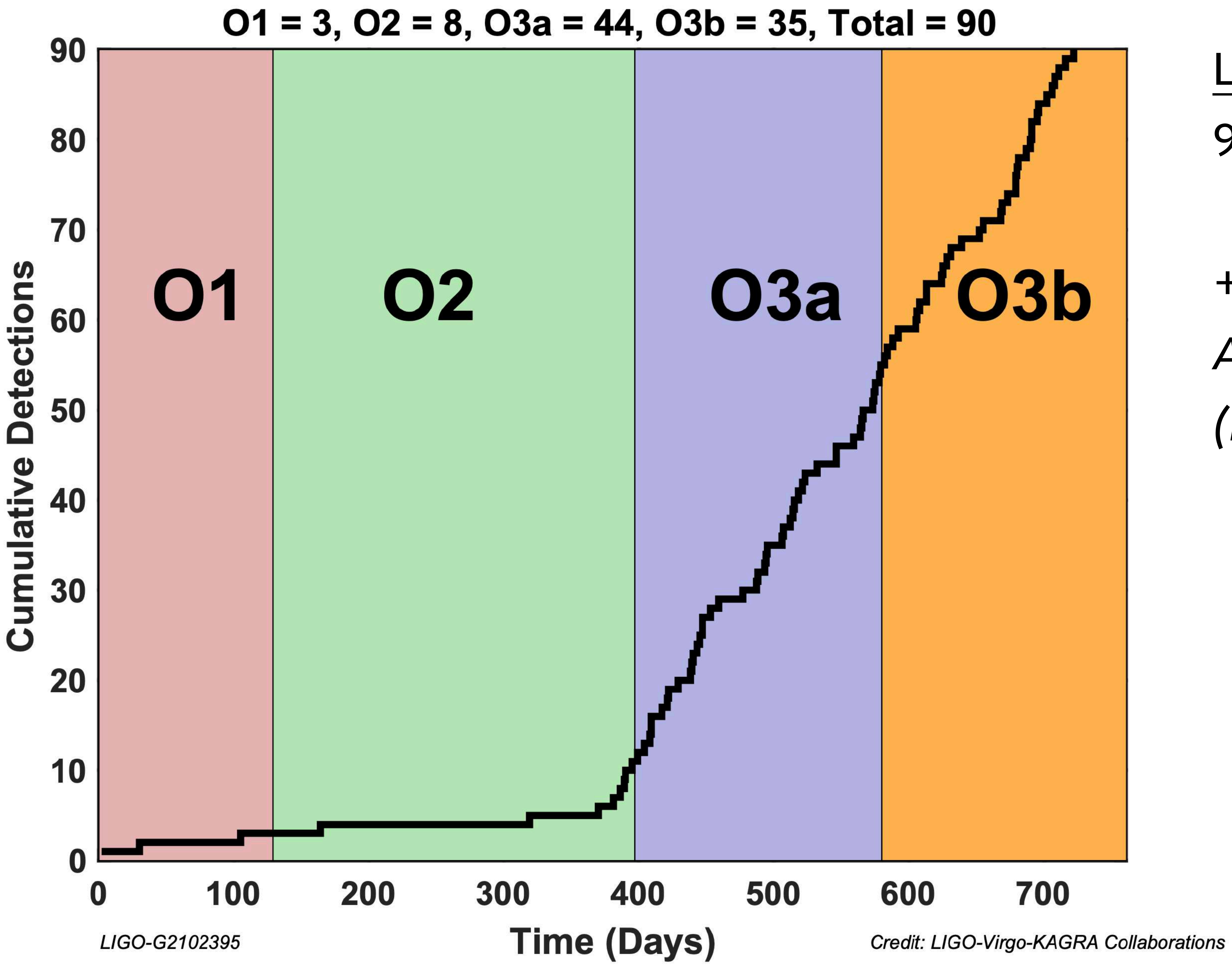
INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



TOTAL NUMBER OF DETECTIONS BY LVK COLLABORATION IN 2015-2020

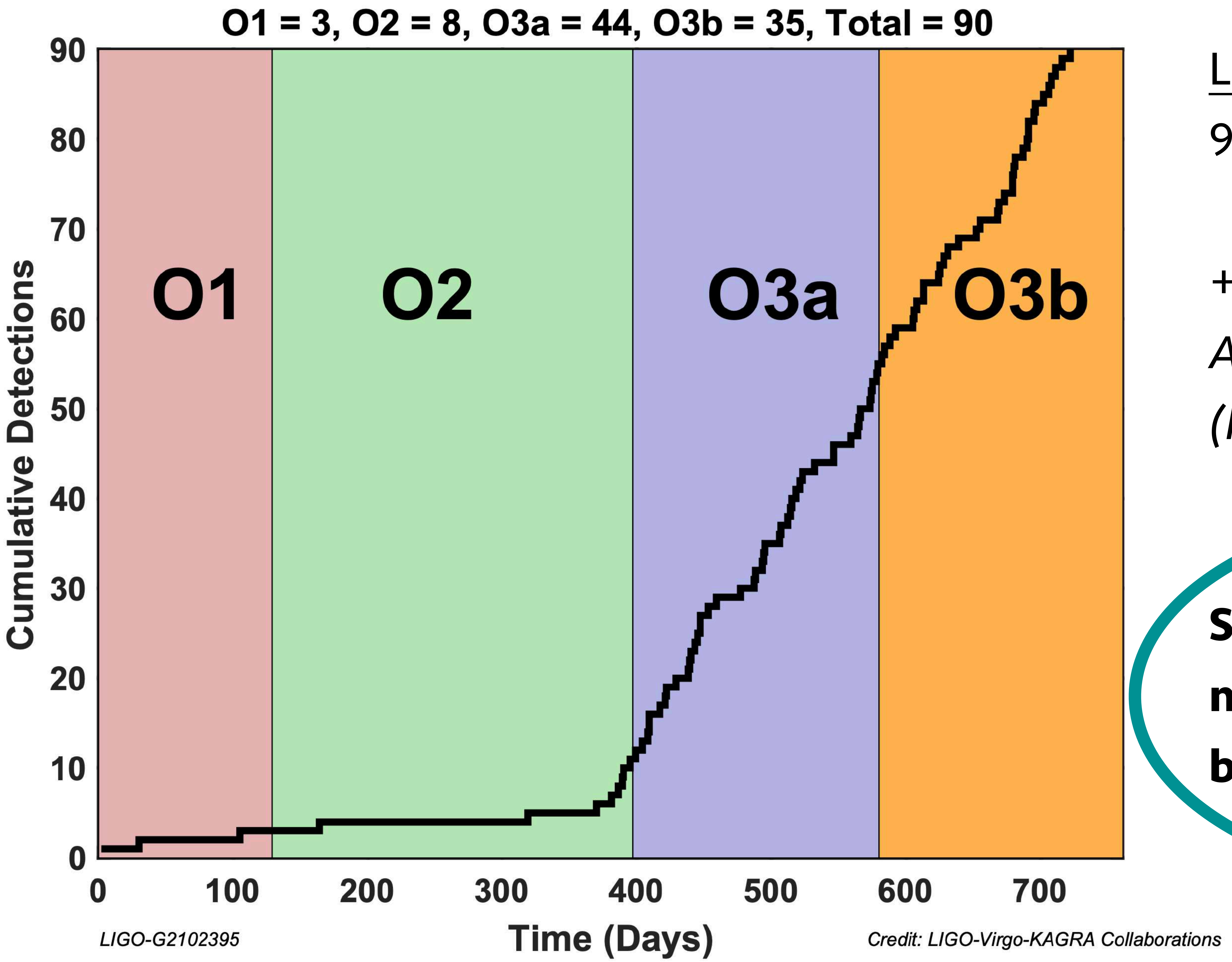


<u>LIGO-Virgo-KAGRA (LVK)</u>:

90 published detections

+ additional detections by AEI Hannover and IAS (Princeton) in O3 data

TOTAL NUMBER OF DETECTIONS BY LVK COLLABORATION IN 2015-2020



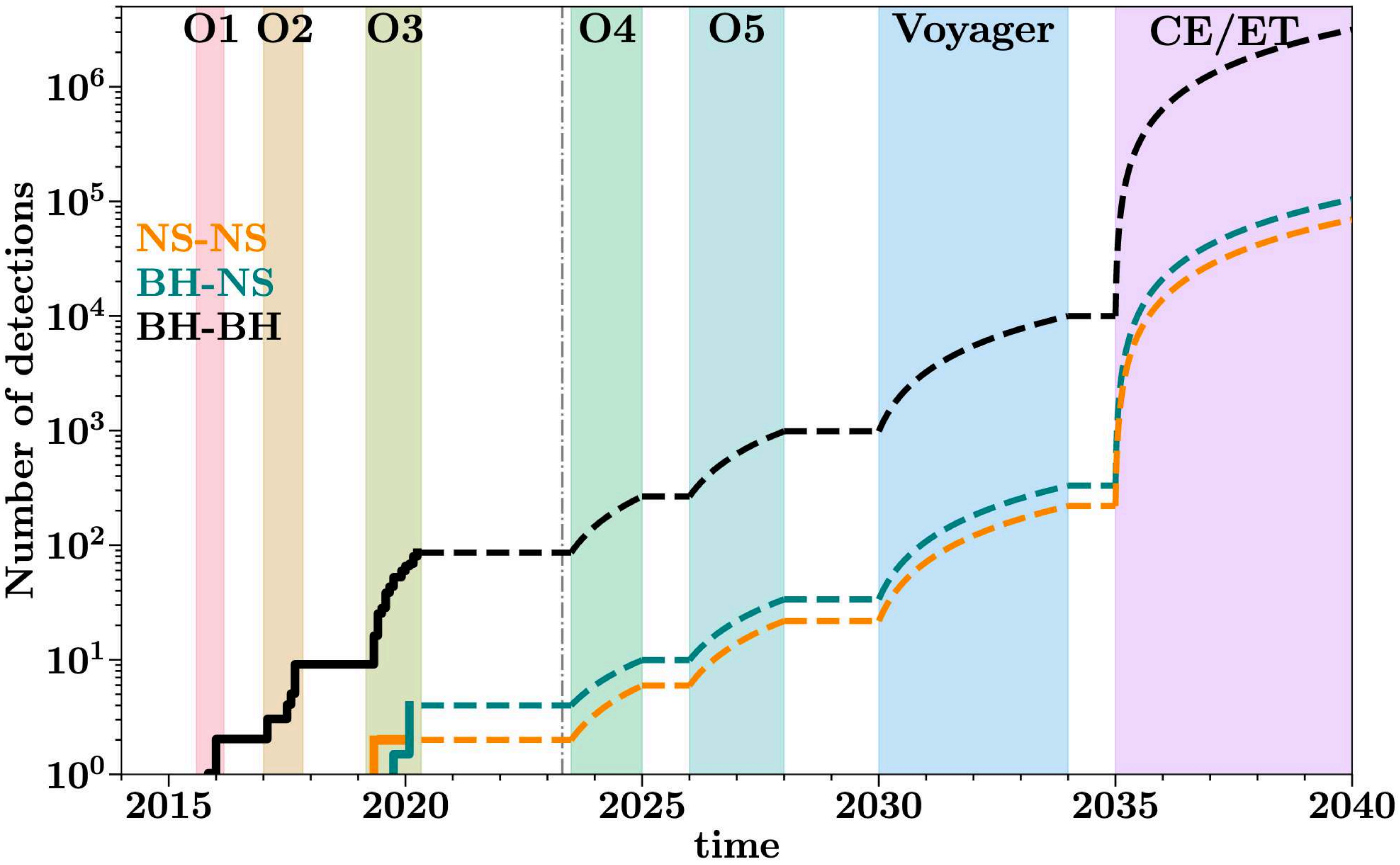
LIGO-Virgo-KAGRA (LVK):

90 published detections

+ additional detections by AEI Hannover and IAS (Princeton) in O3 data

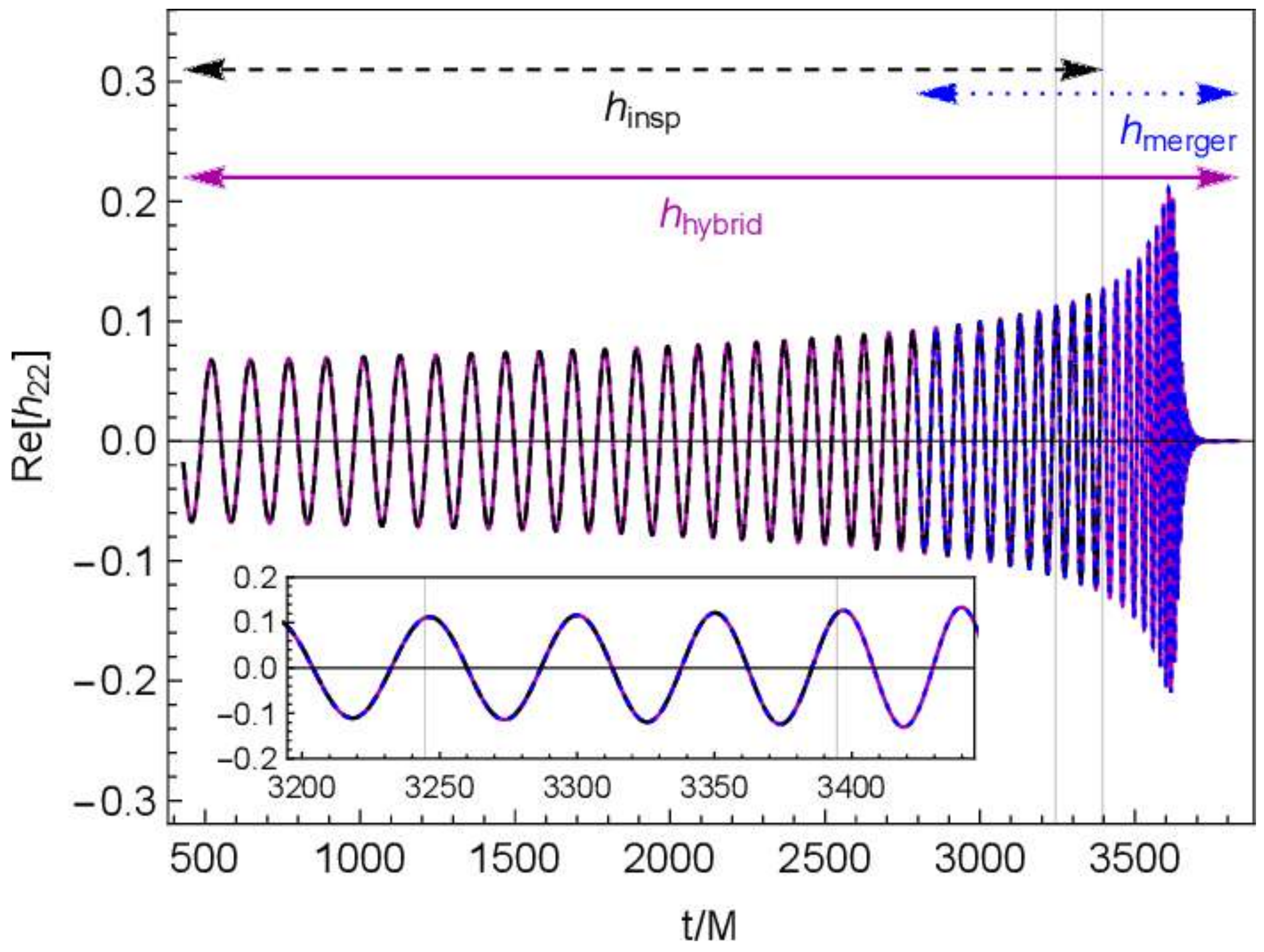
Stay tuned for upcoming new catalog of O4 detections by the LVK Collaboration!

EXPECTED NUMBER OF DETECTIONS

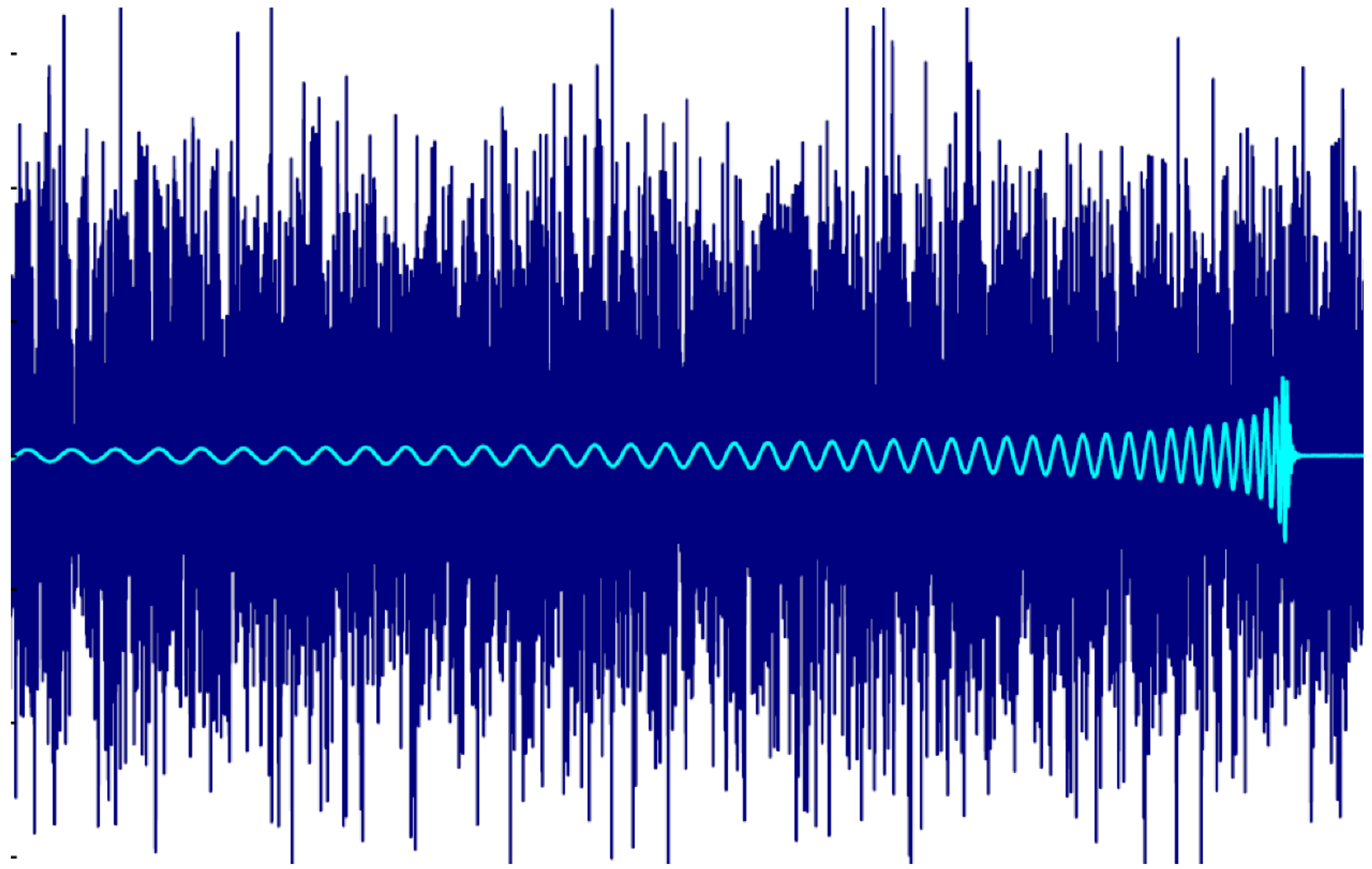


MACHINE LEARNING IN GRAVITATIONAL WAVE ASTRONOMY

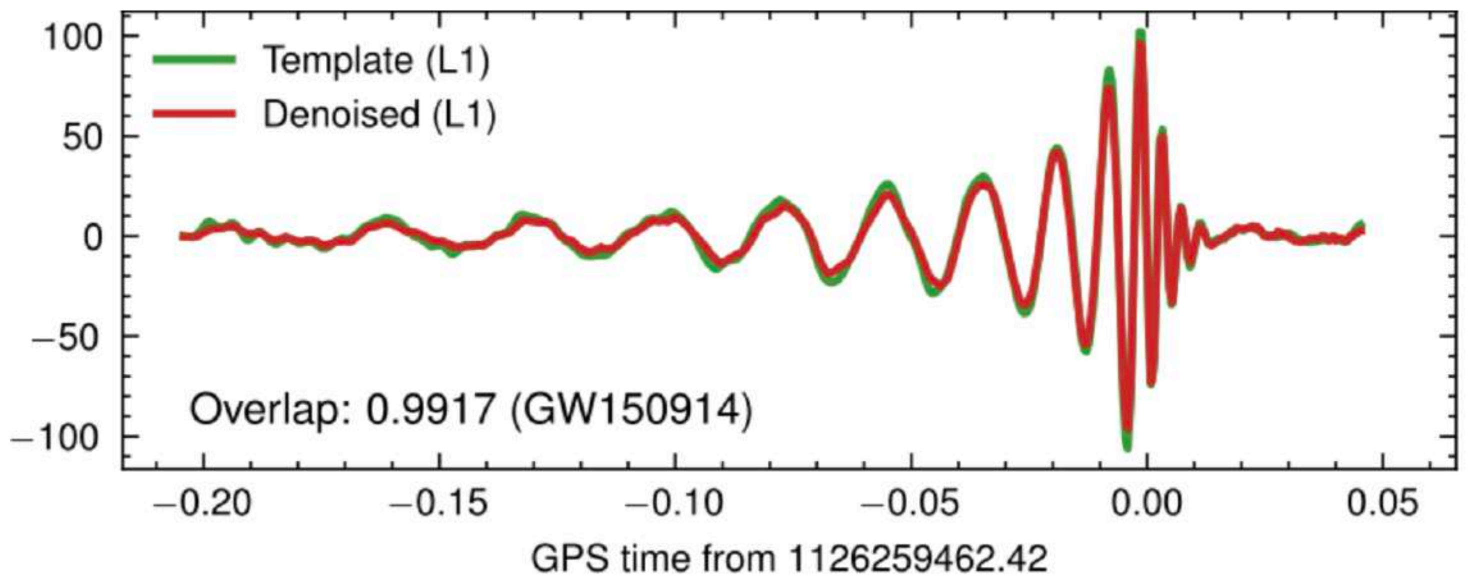
Waveform Modeling



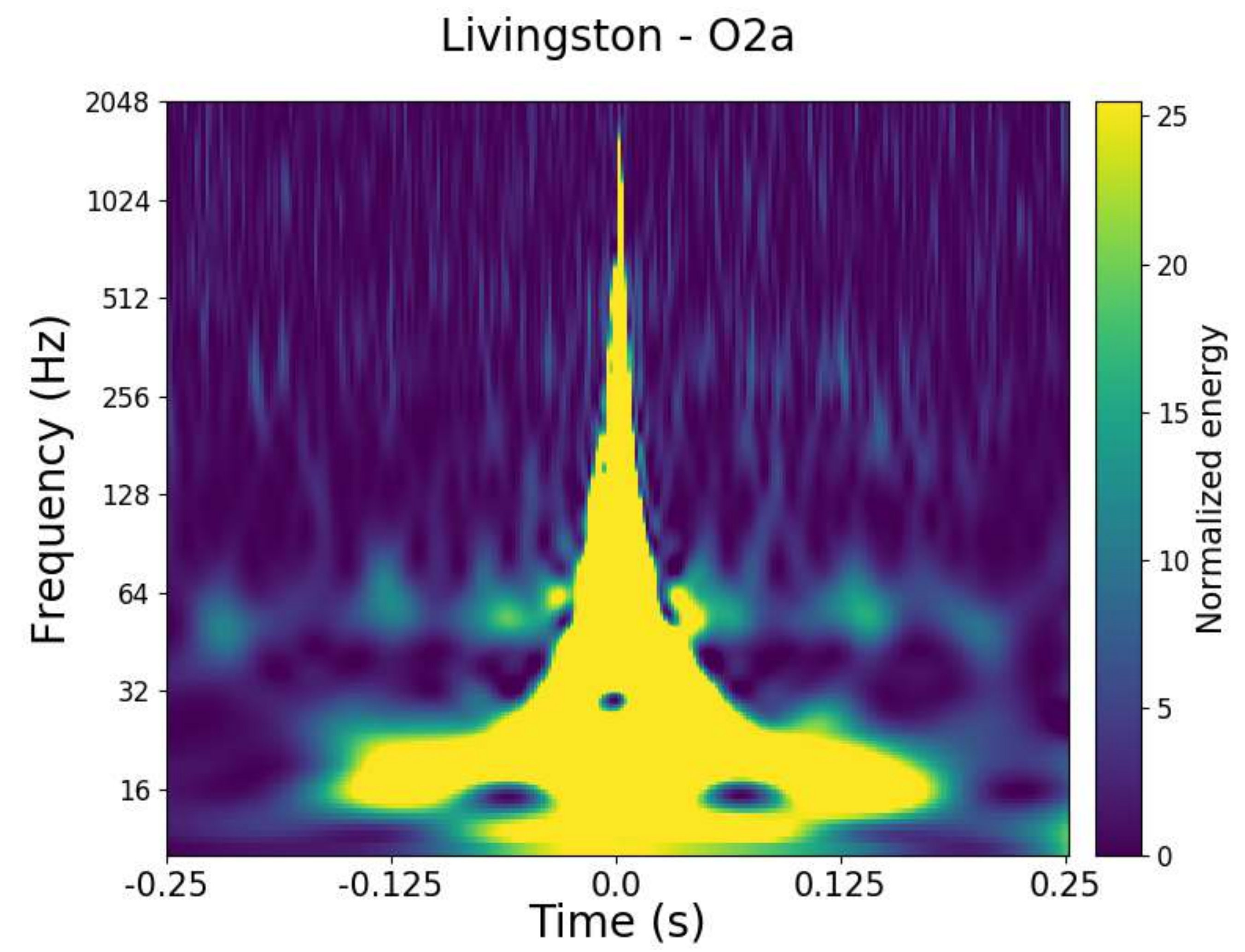
Detection



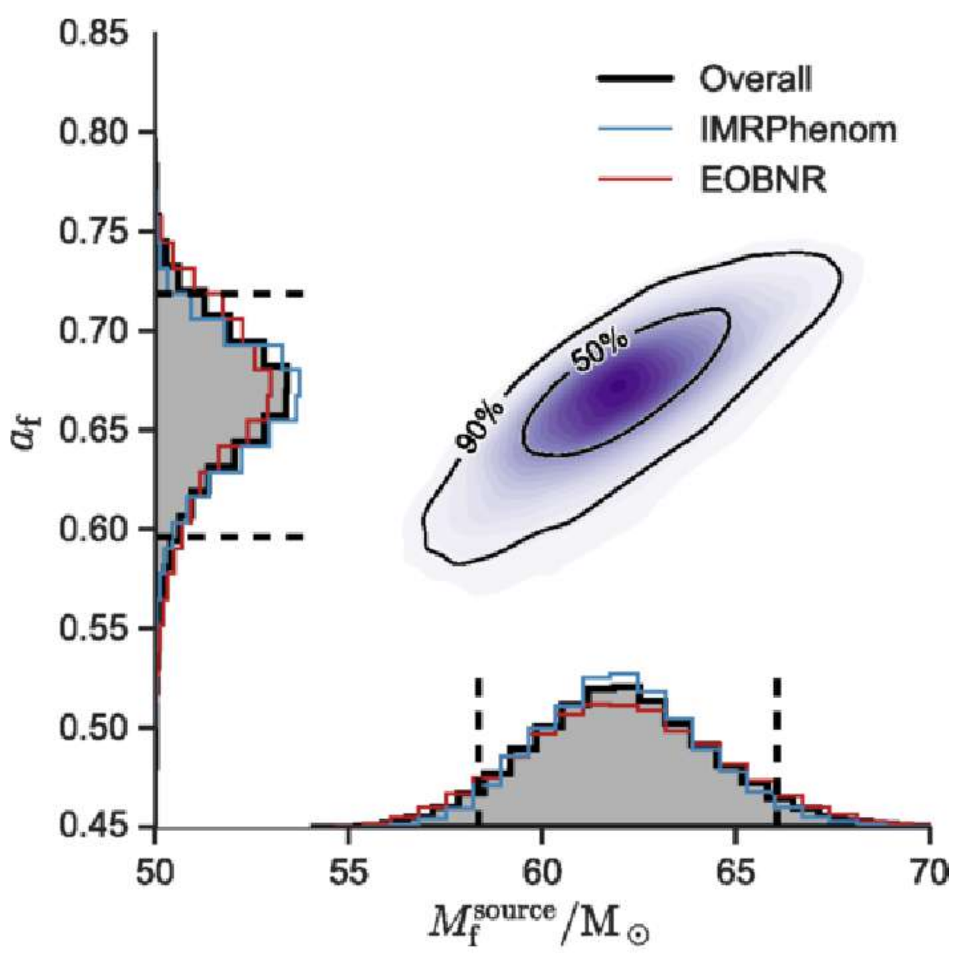
Denoising



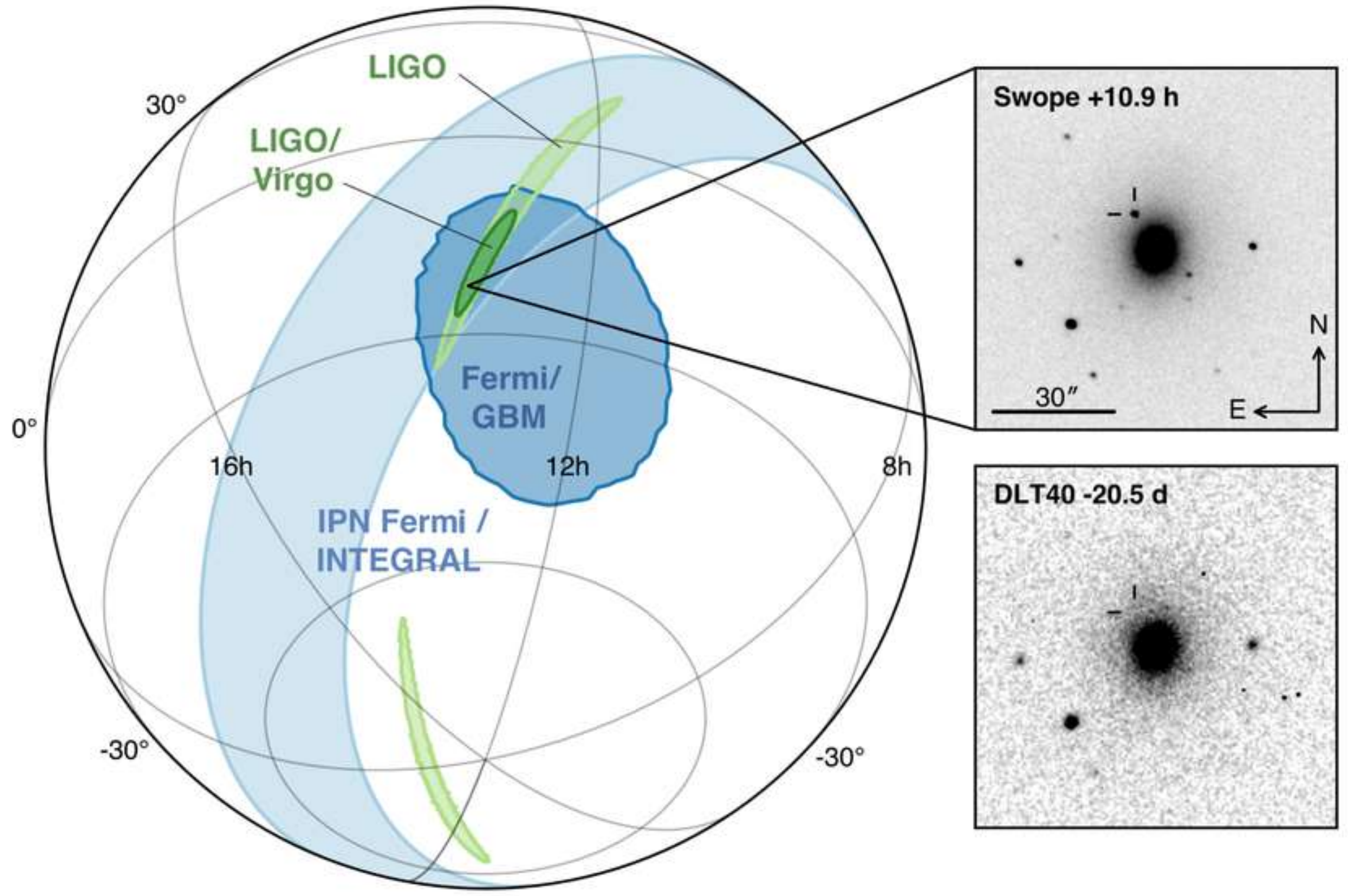
Glitch Classification



Parameter Estimation



Sky Localization



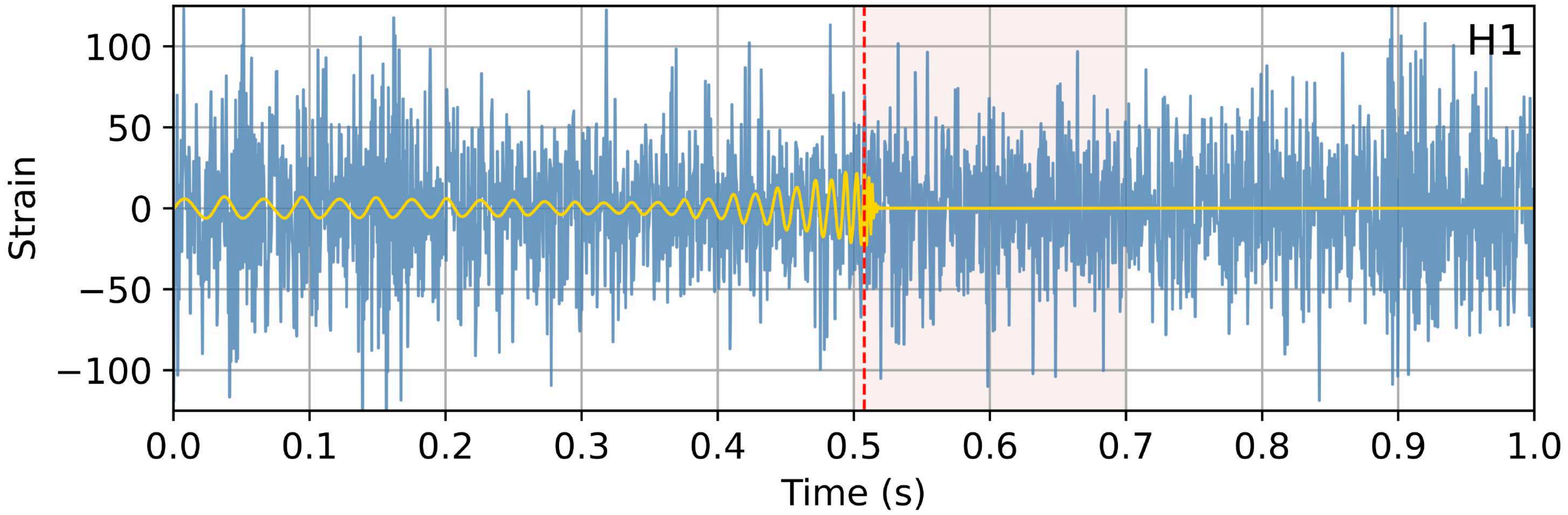
PHYSICAL REVIEW D **108,** 024022 (2023)

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)

AresGW

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

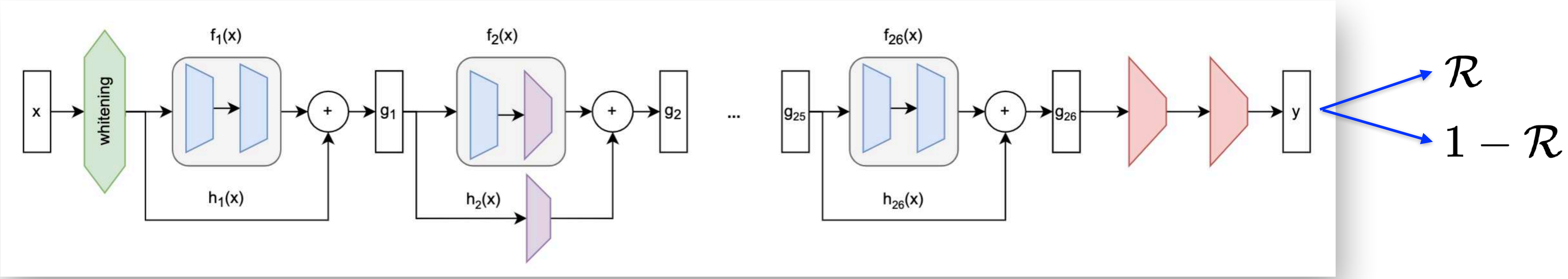
Mass range: $7M_{\odot} \leq M \leq 50M_{\odot}$

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

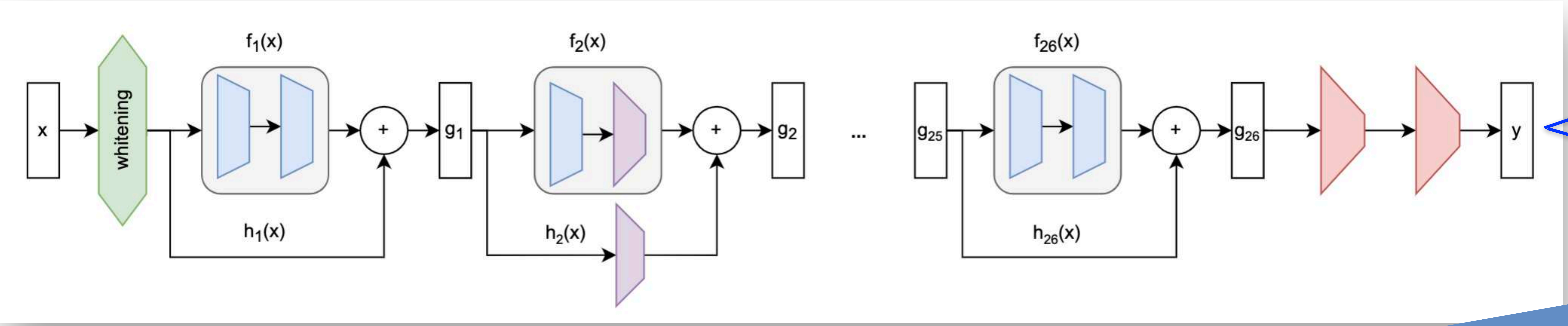
1-D ResNet-54 (27 residual blocks with 2 convolutional layers each and skip connections)!



- 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

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

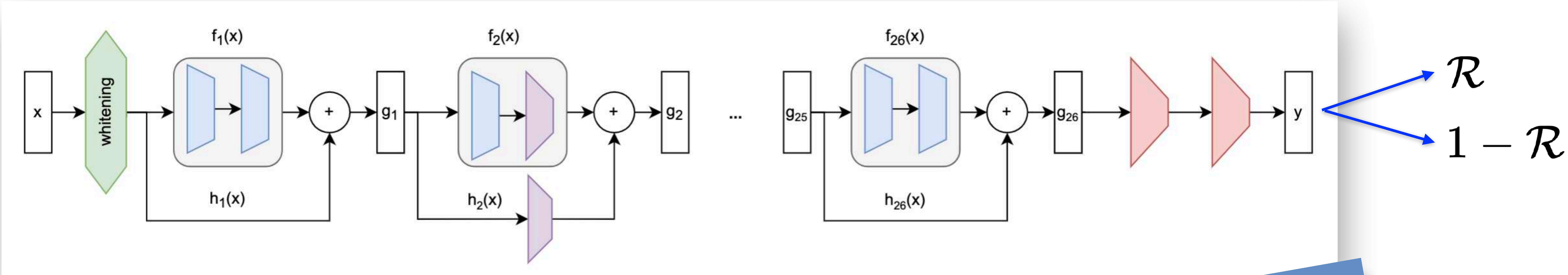


- Residual blocks with skip connections: g(**x**) = f(
- h(**x**) = **x** *or* a strided convolutional layer
- After each convolutional layer: batch nor
  ReLU activation
- Mini-batch size of 400 segments
- Adam optimizer for back propagation
- Objective function = regularized binary cross entropy

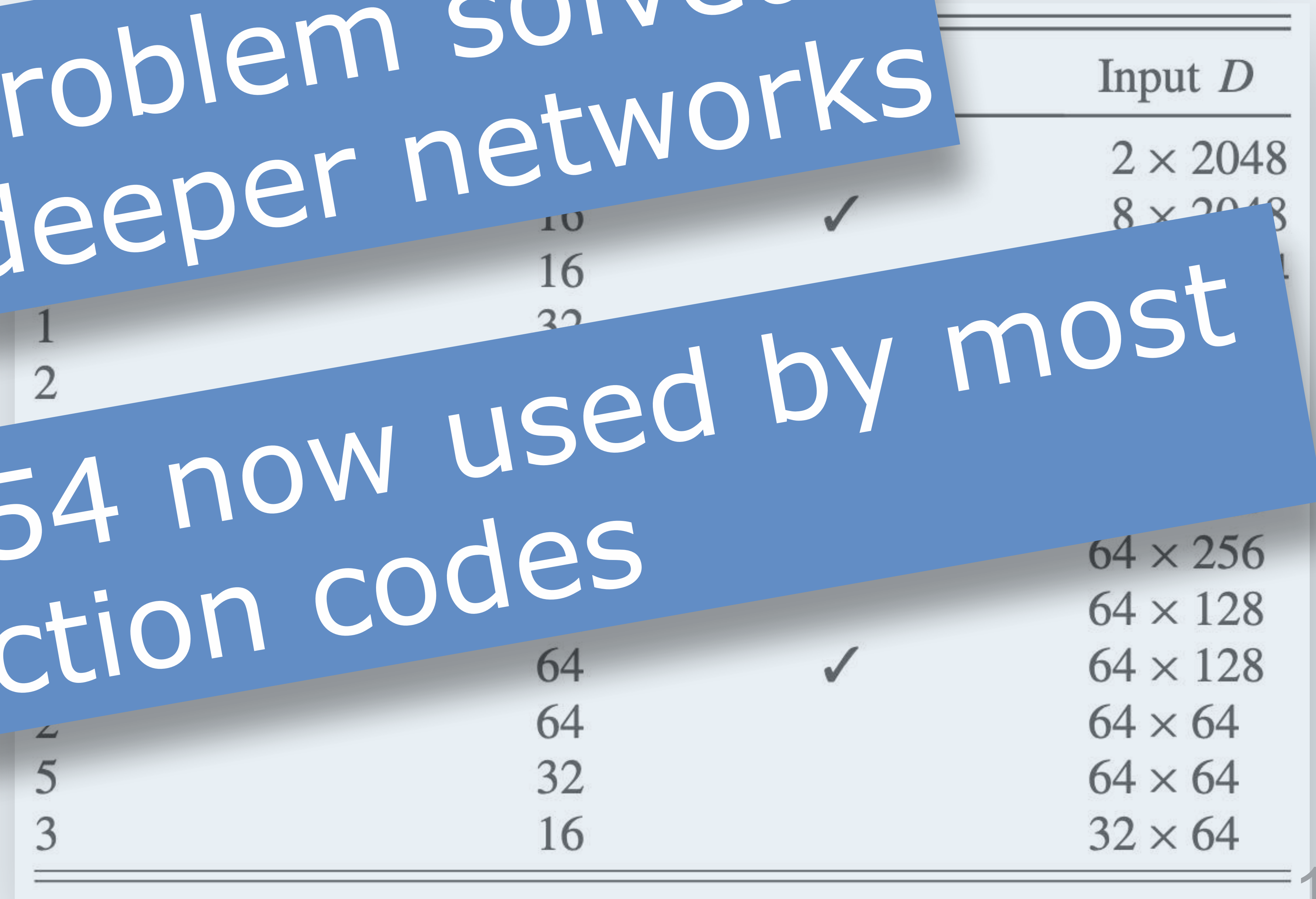
Gradient problem solved!
-> much deeper networks

			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

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

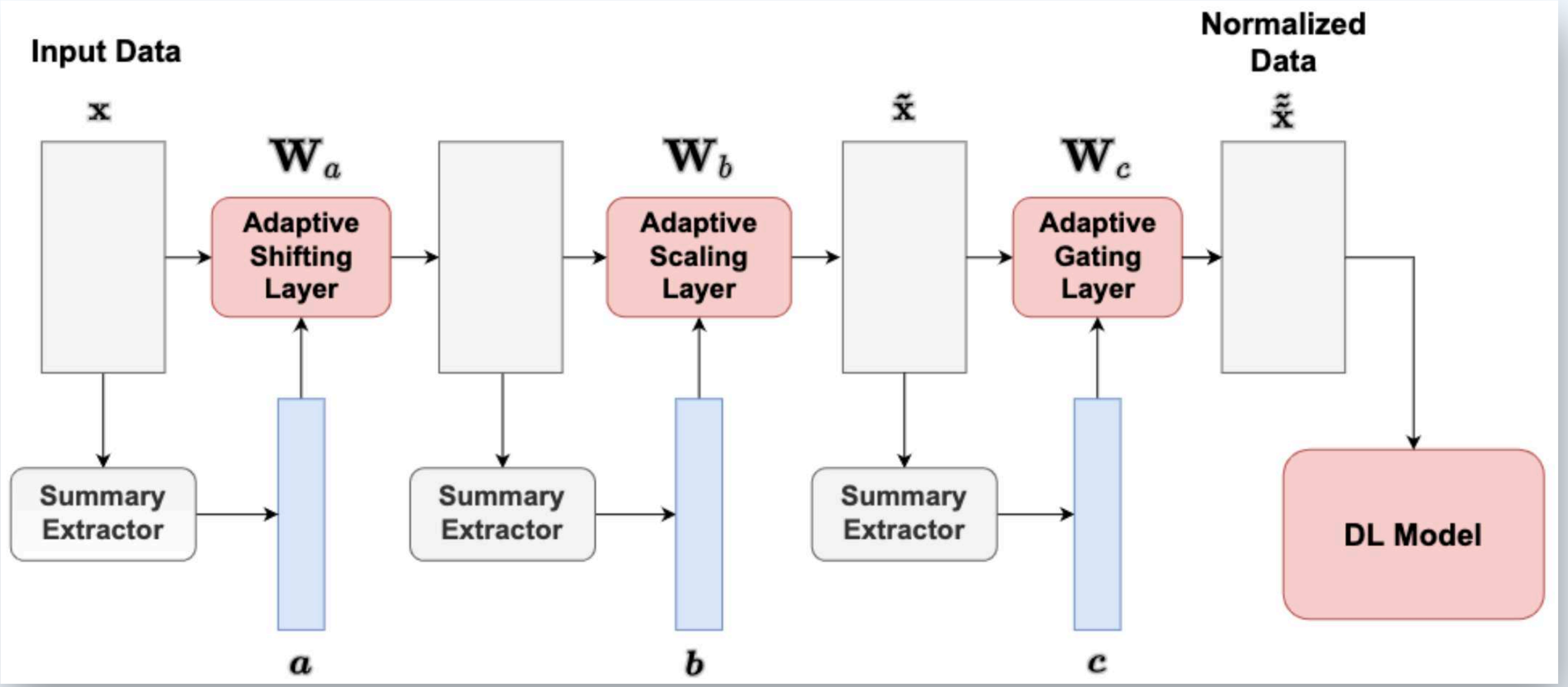


Gradient problem solved!
-> much deeper networks

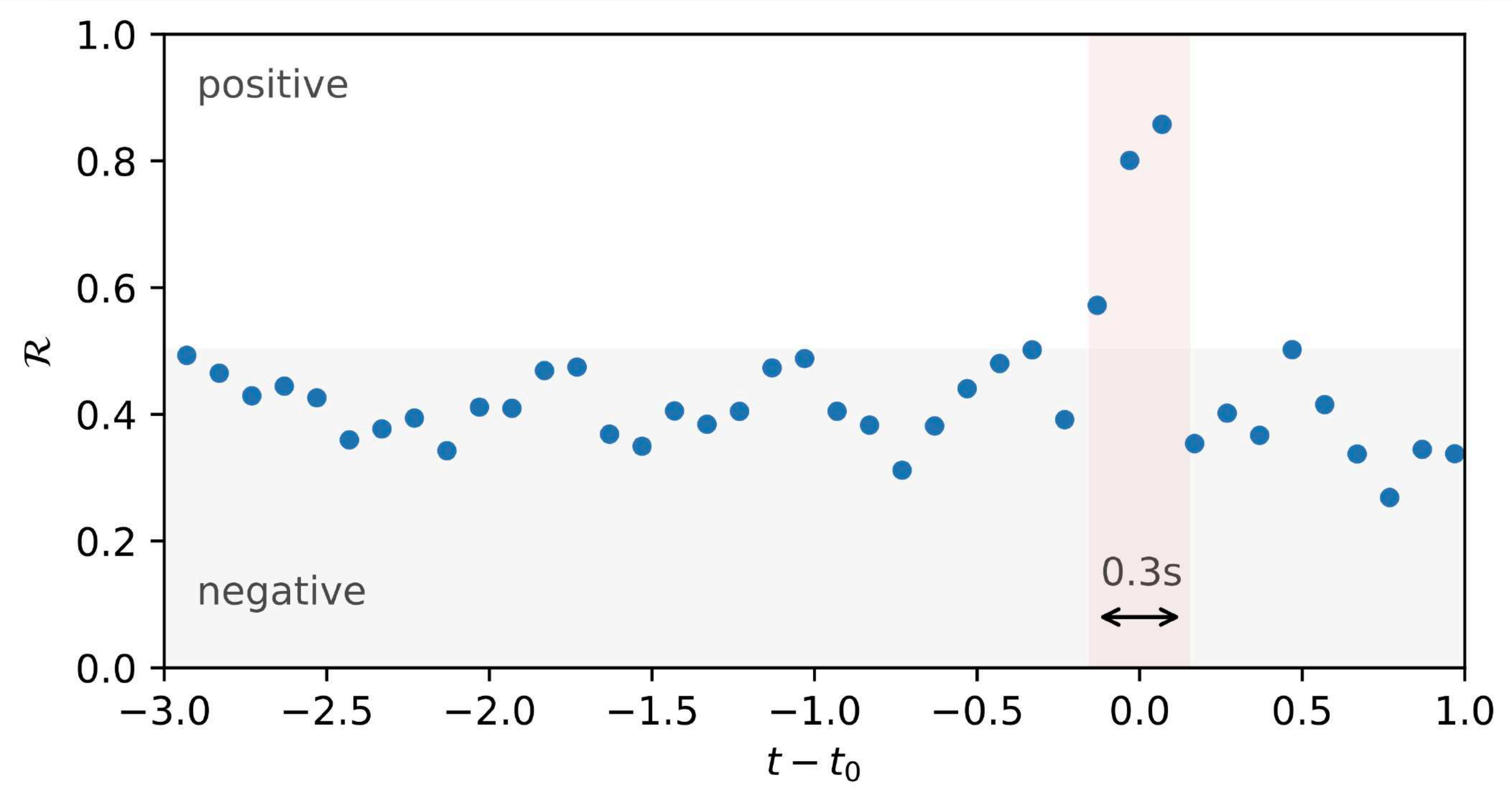
1-D ResNet-54 now used by most ML GW detection codes

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!



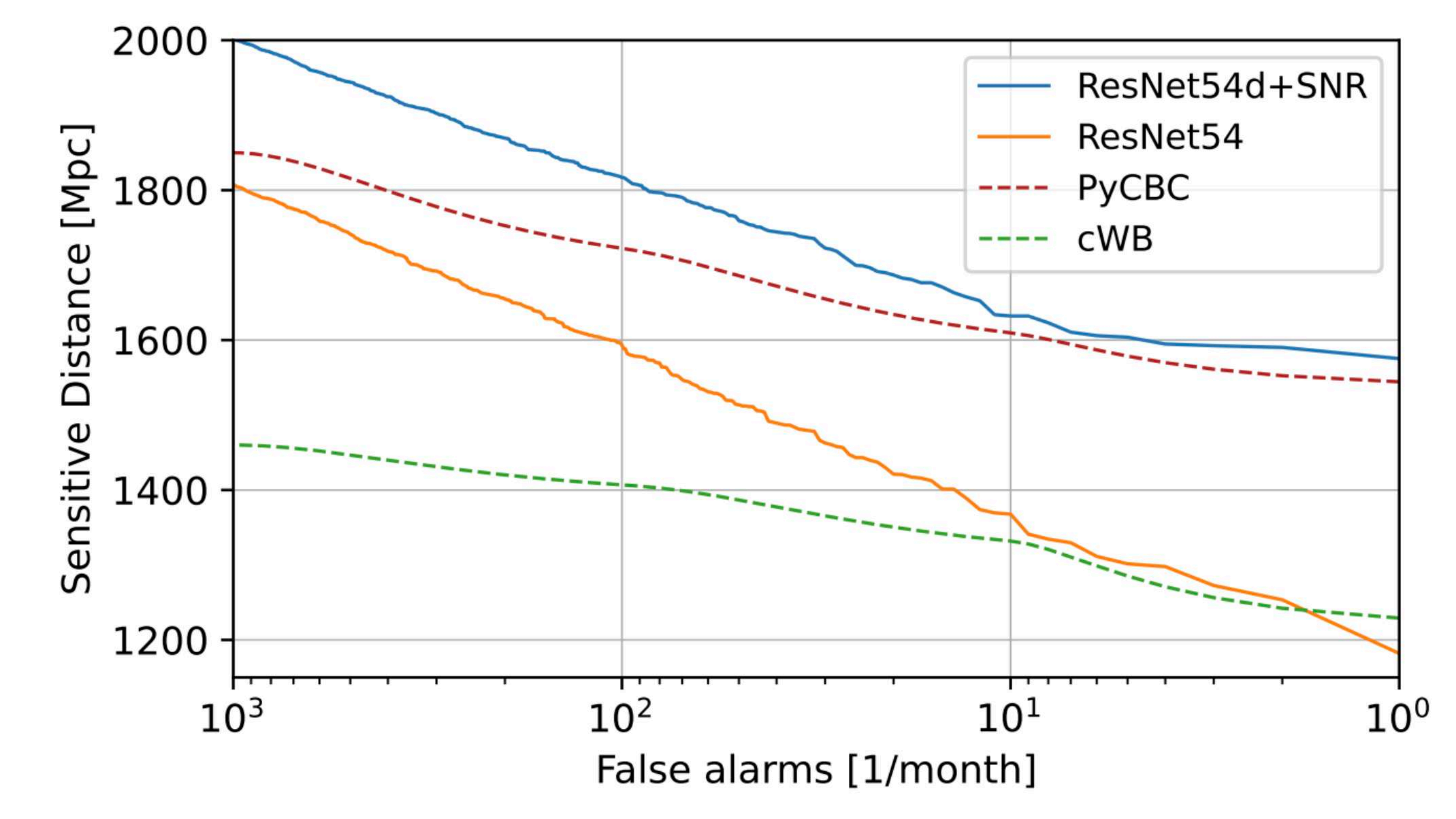
Neural Network Ranking Statistic R is reported every 0.1 s of input data



Logarithmic Ranking Statistic

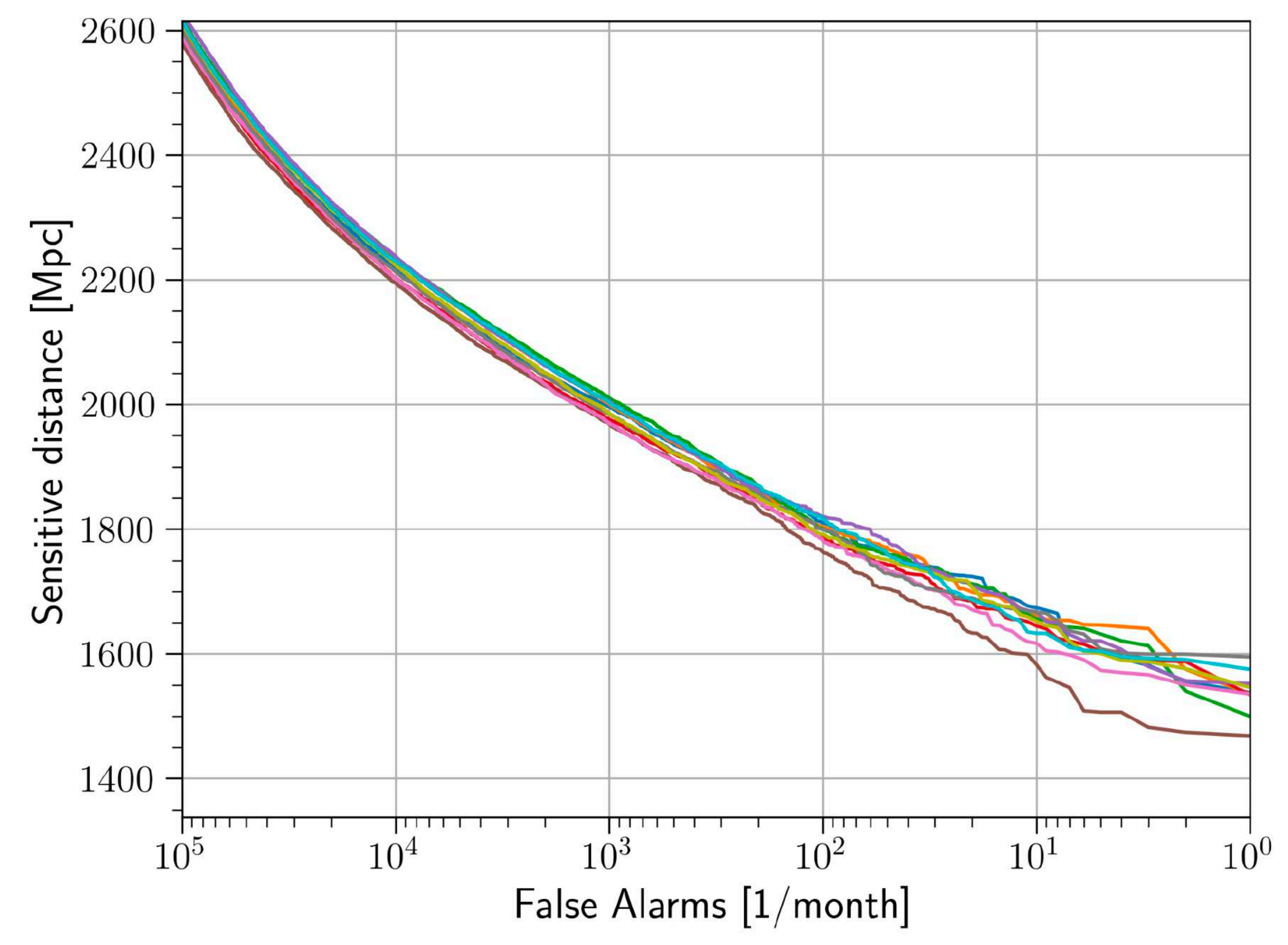
$$\boxed{\mathcal{R}_s = -\log_{10}\left(1 - \mathcal{R} + 10^{-16}\right)}$$

AresGW is the first ML code to exceed **the sensitivity of (non-optimized for this mass range) classical algorithms**



Variance with respect to different noise realizations and different injections:

Sensitive distance ~ 4 %

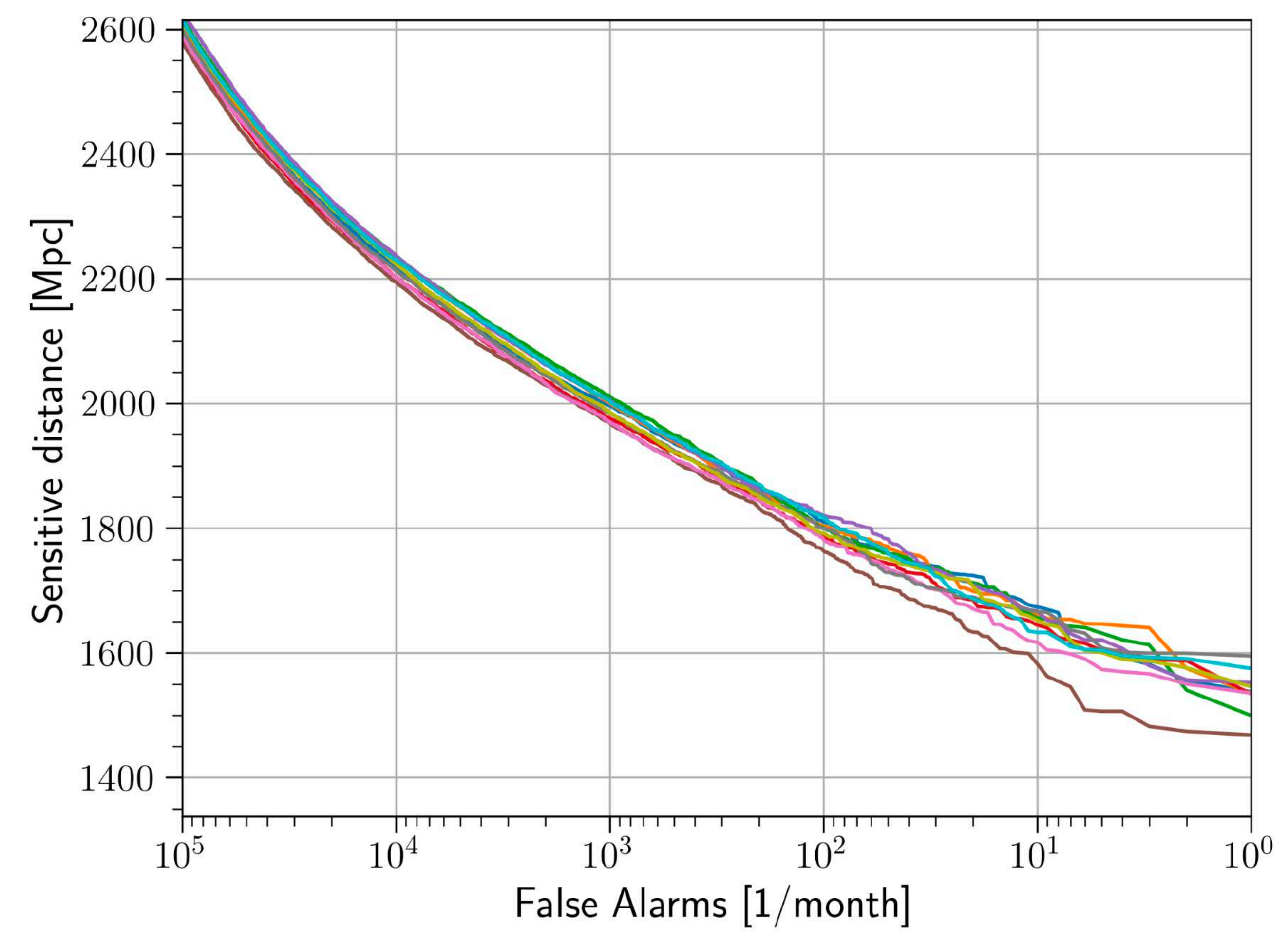


Koloniari, Paschalidis, Lazaridis, Stergioulas, in preparation (2025)

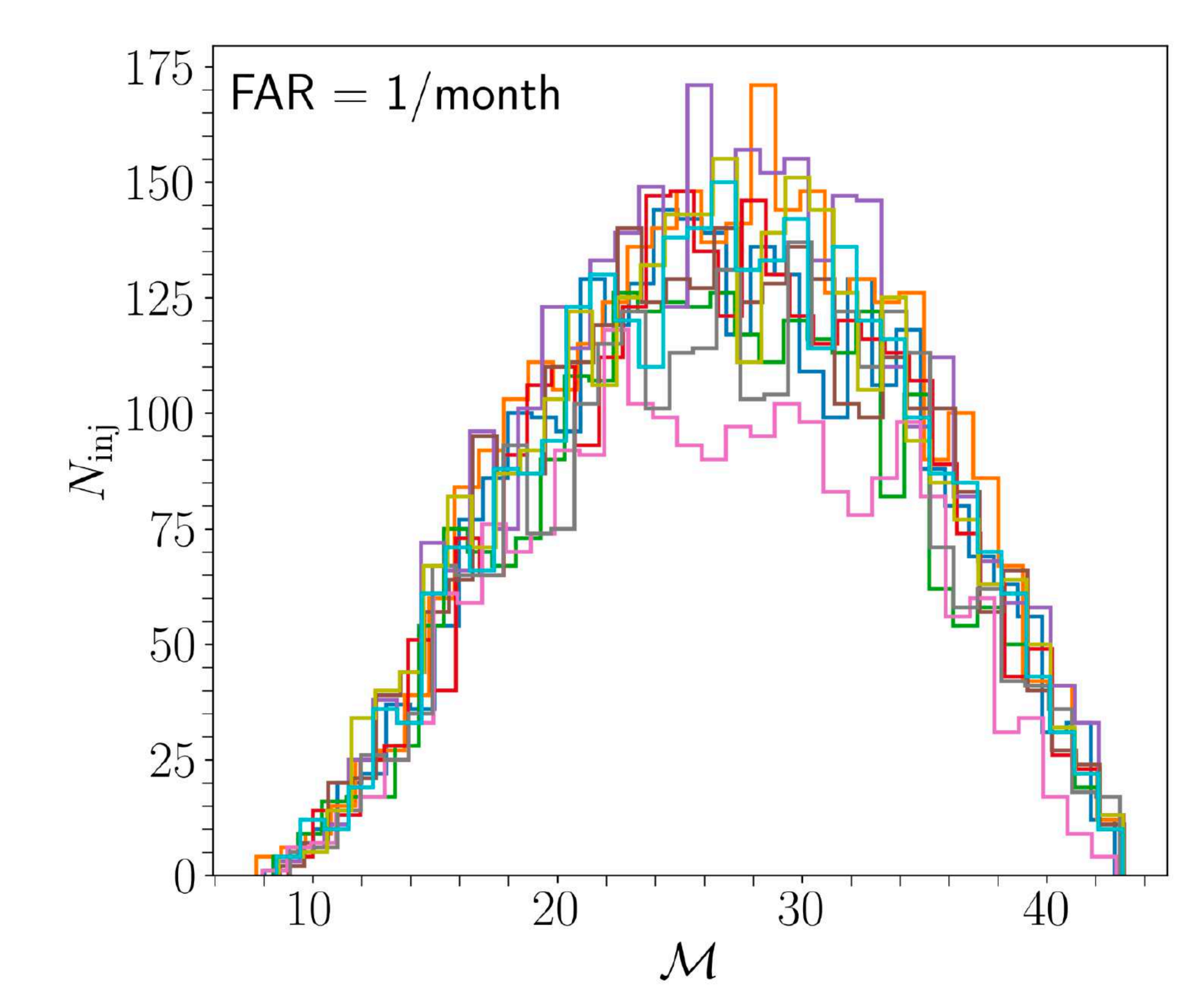
ROBUSTNESS OF SENSITIVITY METRICS FOR ML DETECTION CODES

Variance with respect to different noise realizations and different injections:

Sensitive distance ~ 4 %



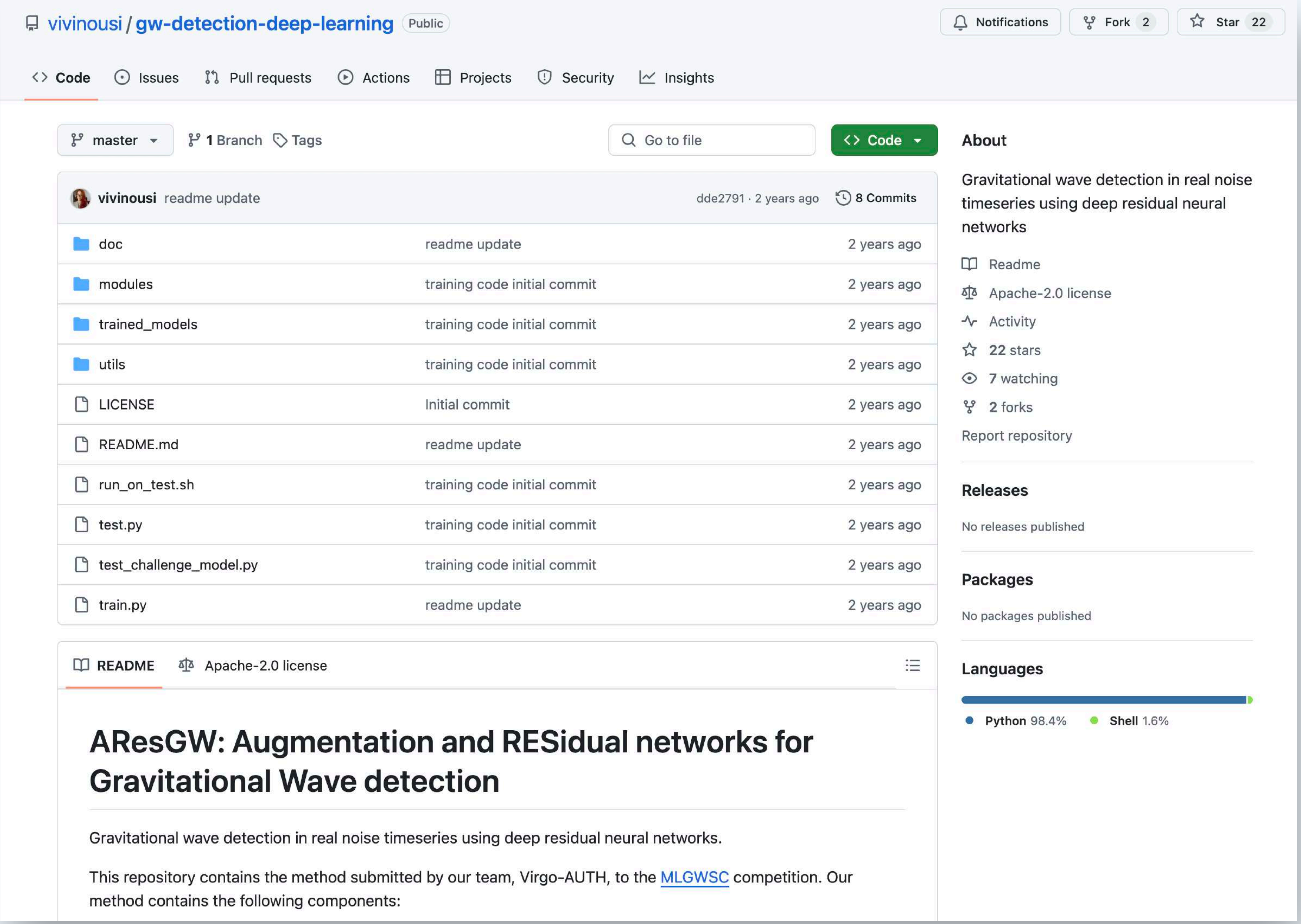
Number of injections ~ 19 %



Koloniari, Paschalidis, Lazaridis, Stergioulas, in preparation (2025)

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

CrossMark

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

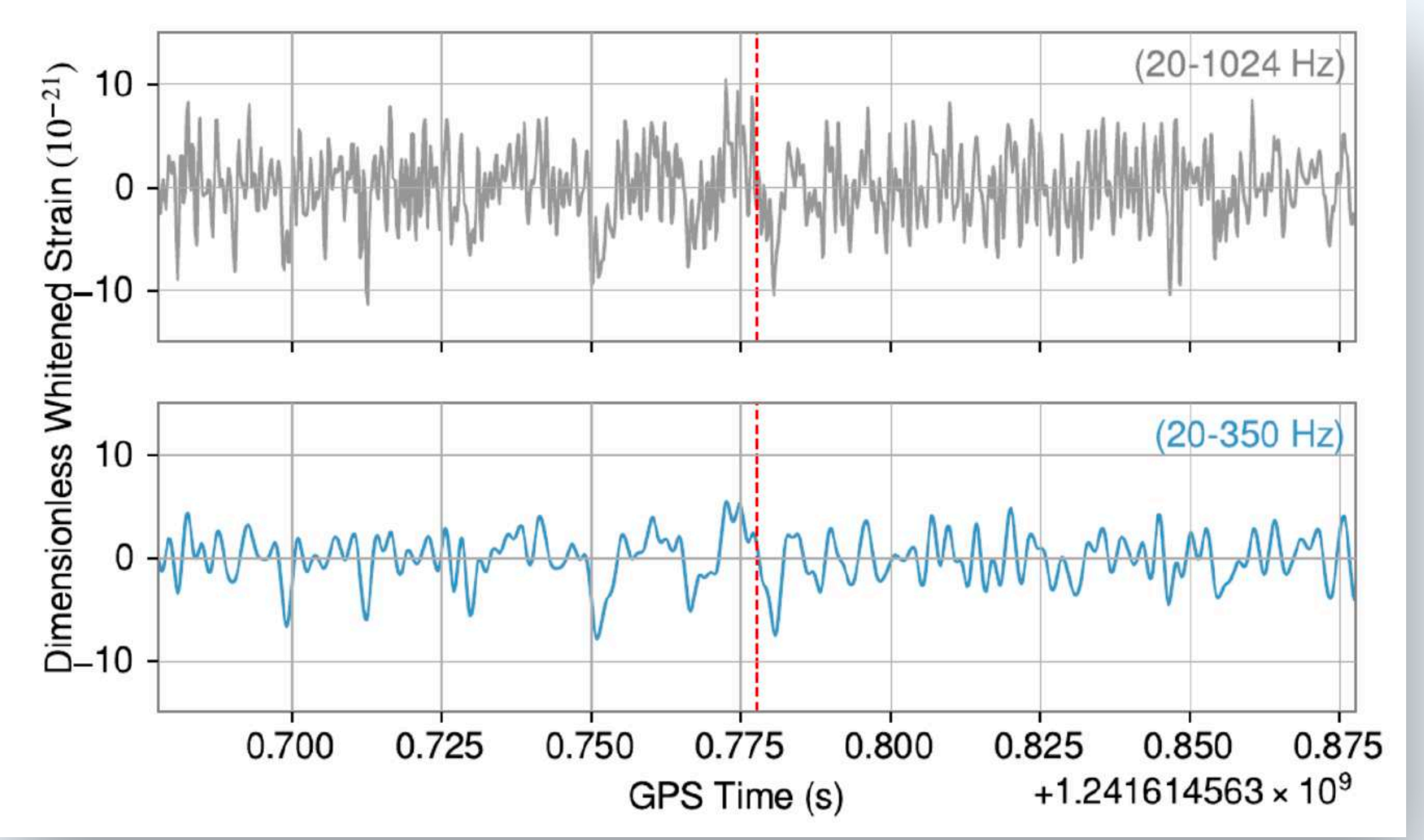
Alexandra E Koloniari[1,*], Evdokia C Koursoumpa[1], Paraskevi Nousi[2], Paraskevas Lampropoulos[1], Nikolaos Passalis[3], Anastasios Tefas[4] and Nikolaos Stergioulas[1]

[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

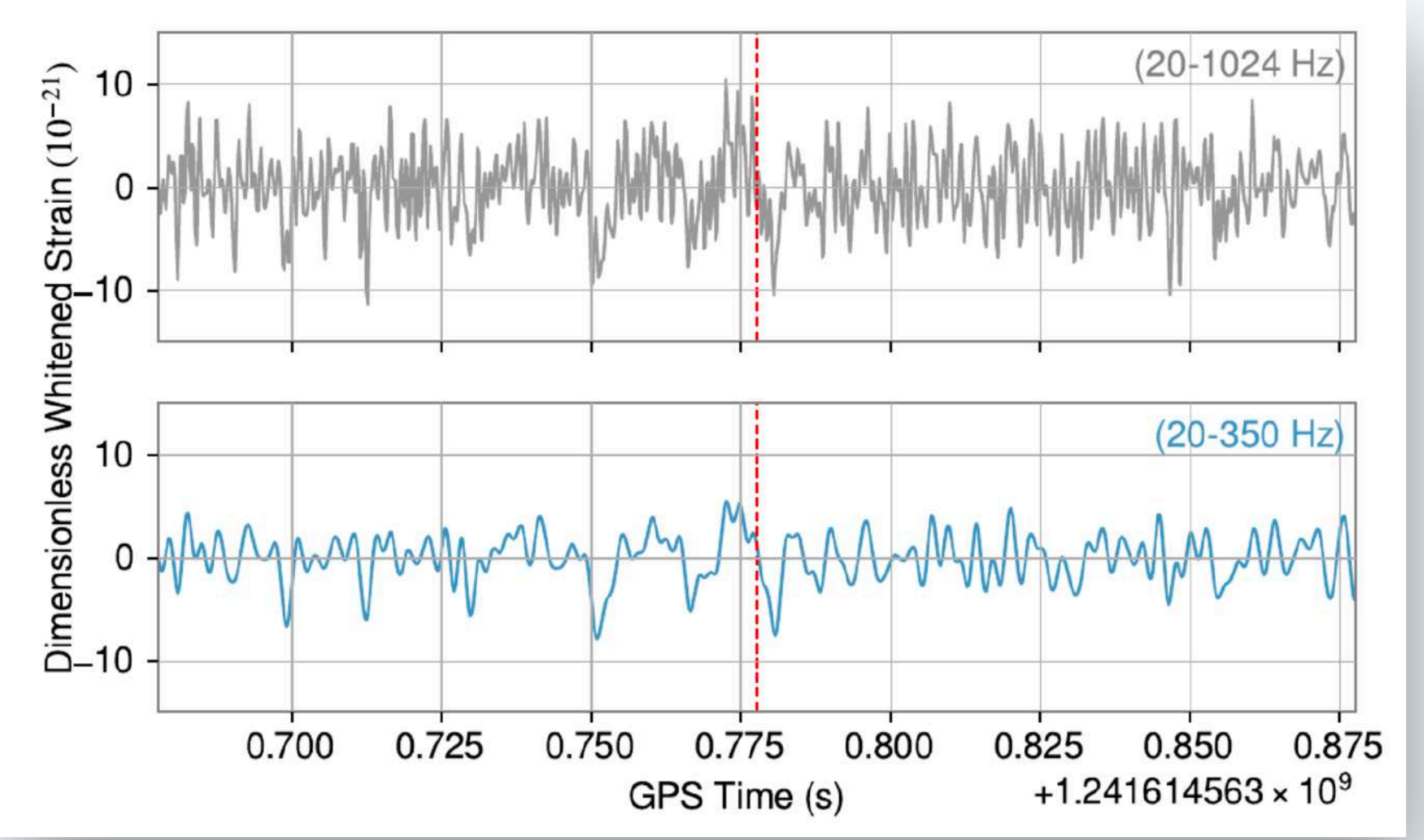
MAIN ENHANCEMENTS:

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

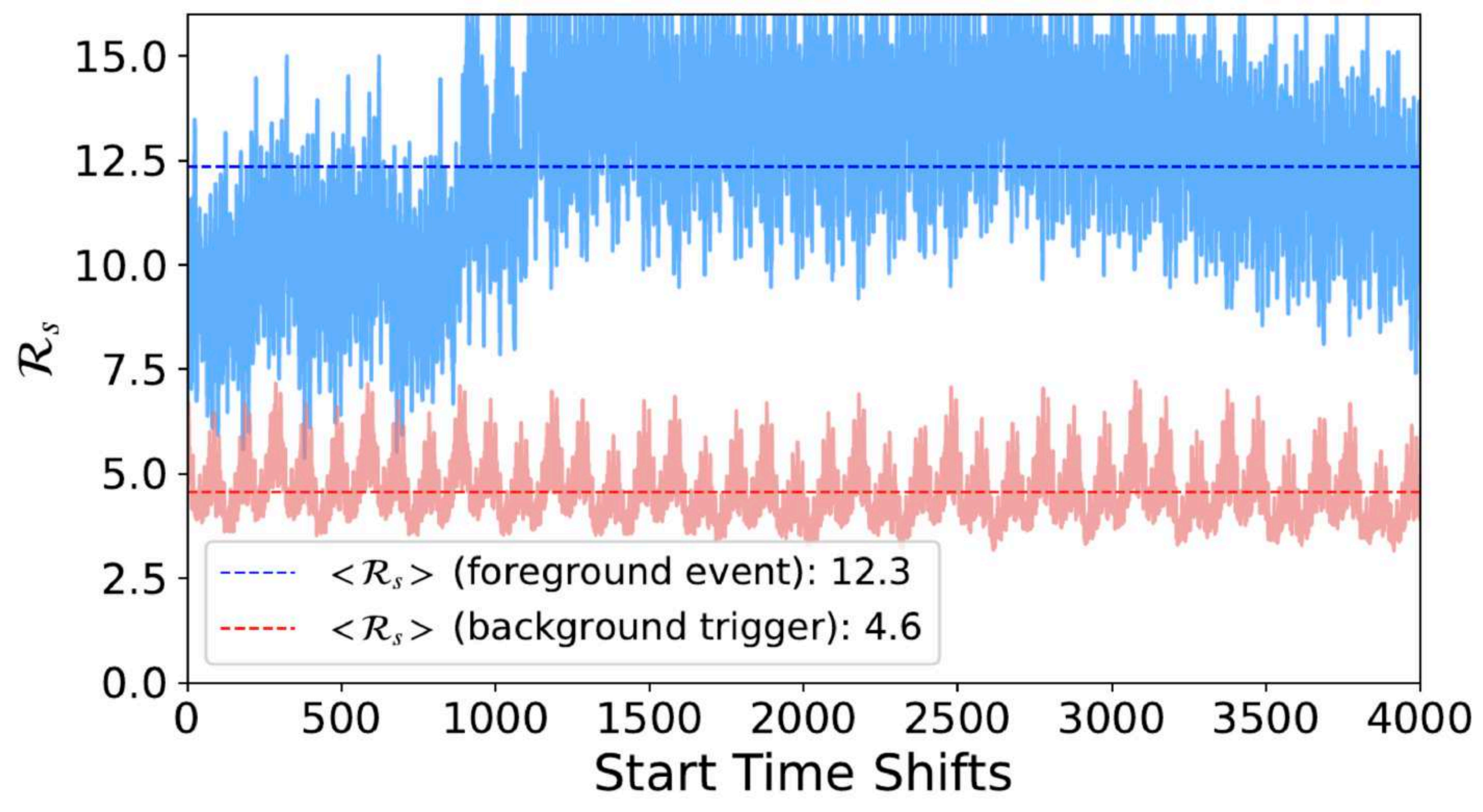


MAIN ENHANCEMENTS:

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

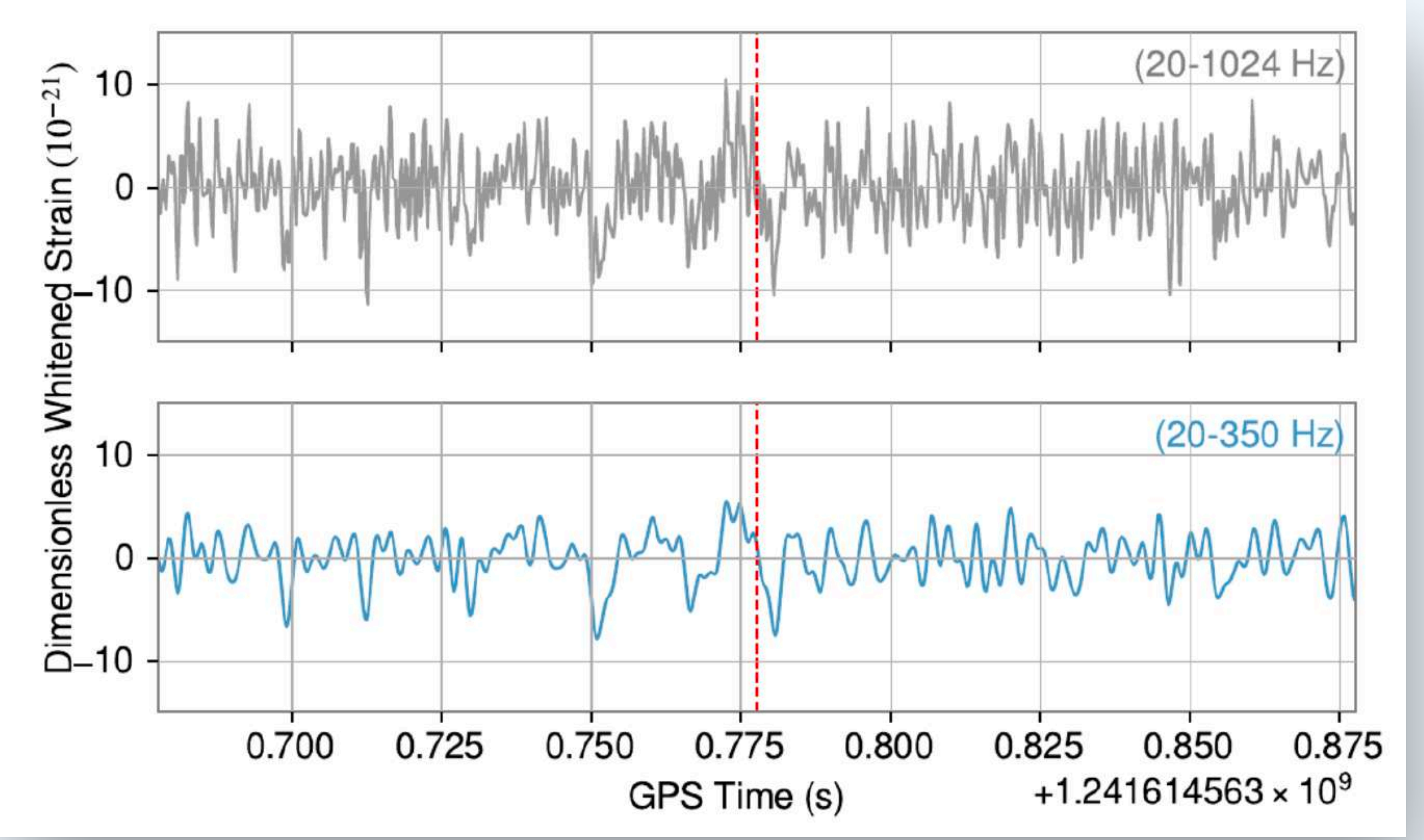


2) **Ensemble-averaged** ranking statistic

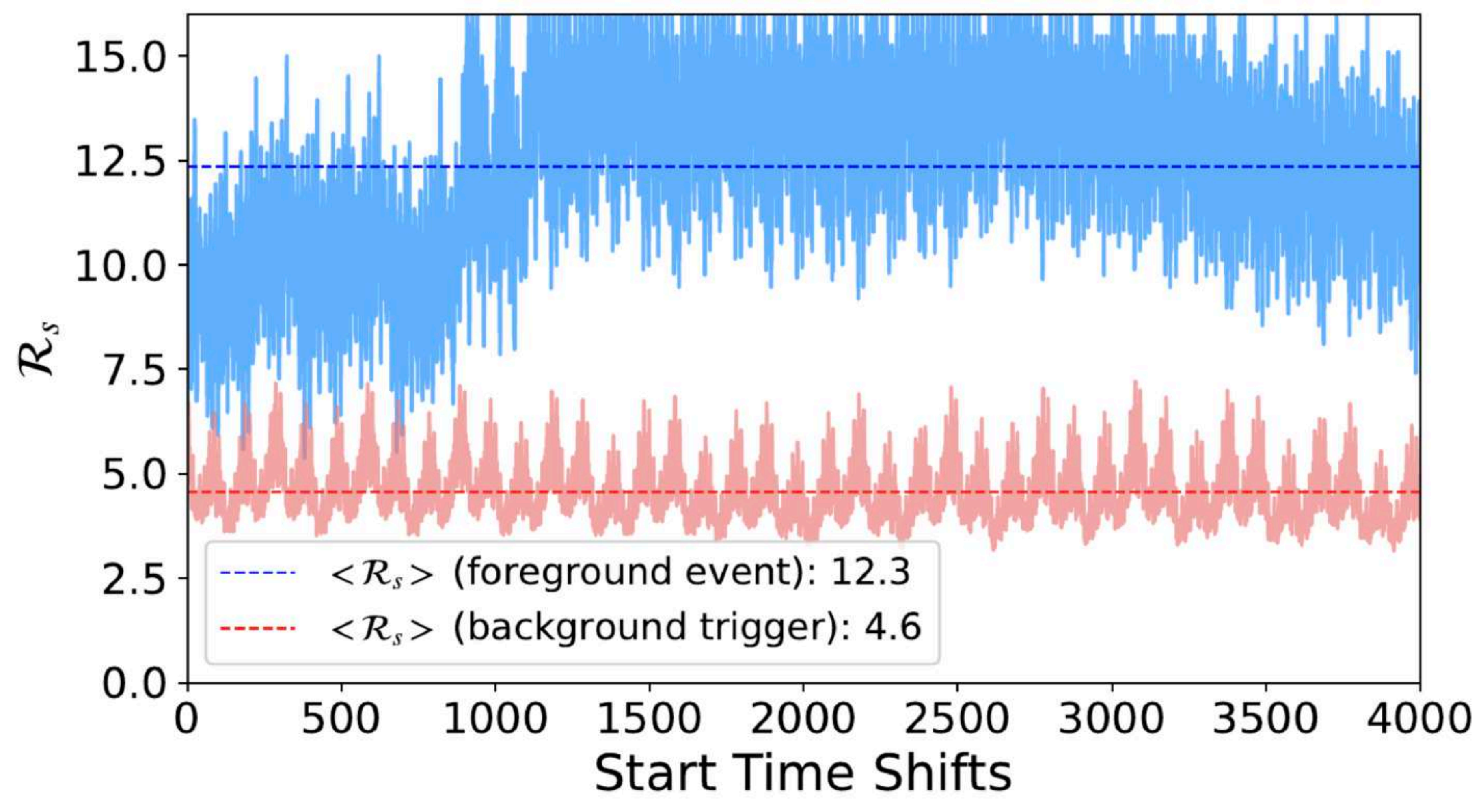


MAIN ENHANCEMENTS:

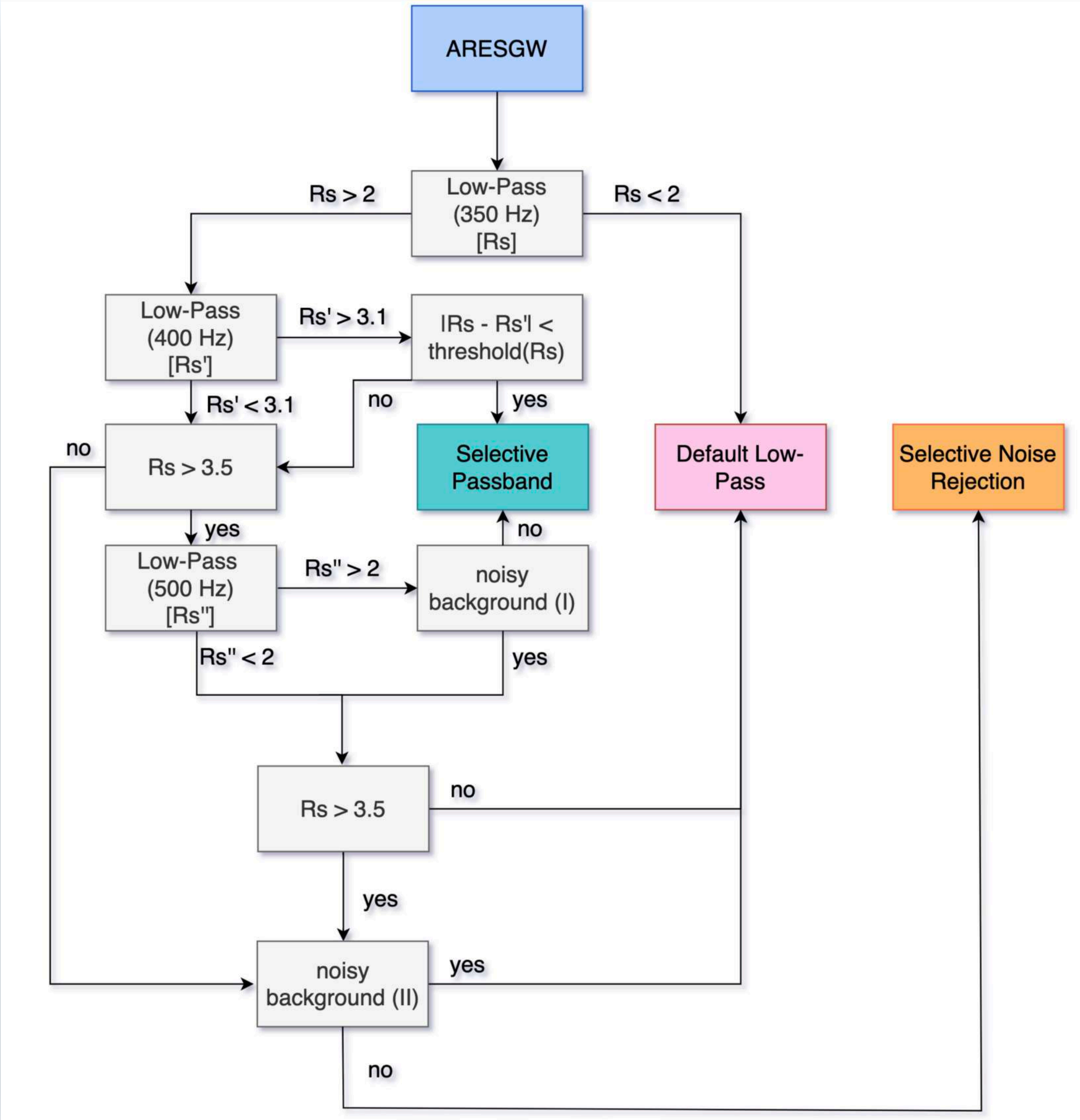
1) **Training** on **low-pass filtered** data



2) **Ensemble-averaged** ranking statistic



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



p_astro is calculated as

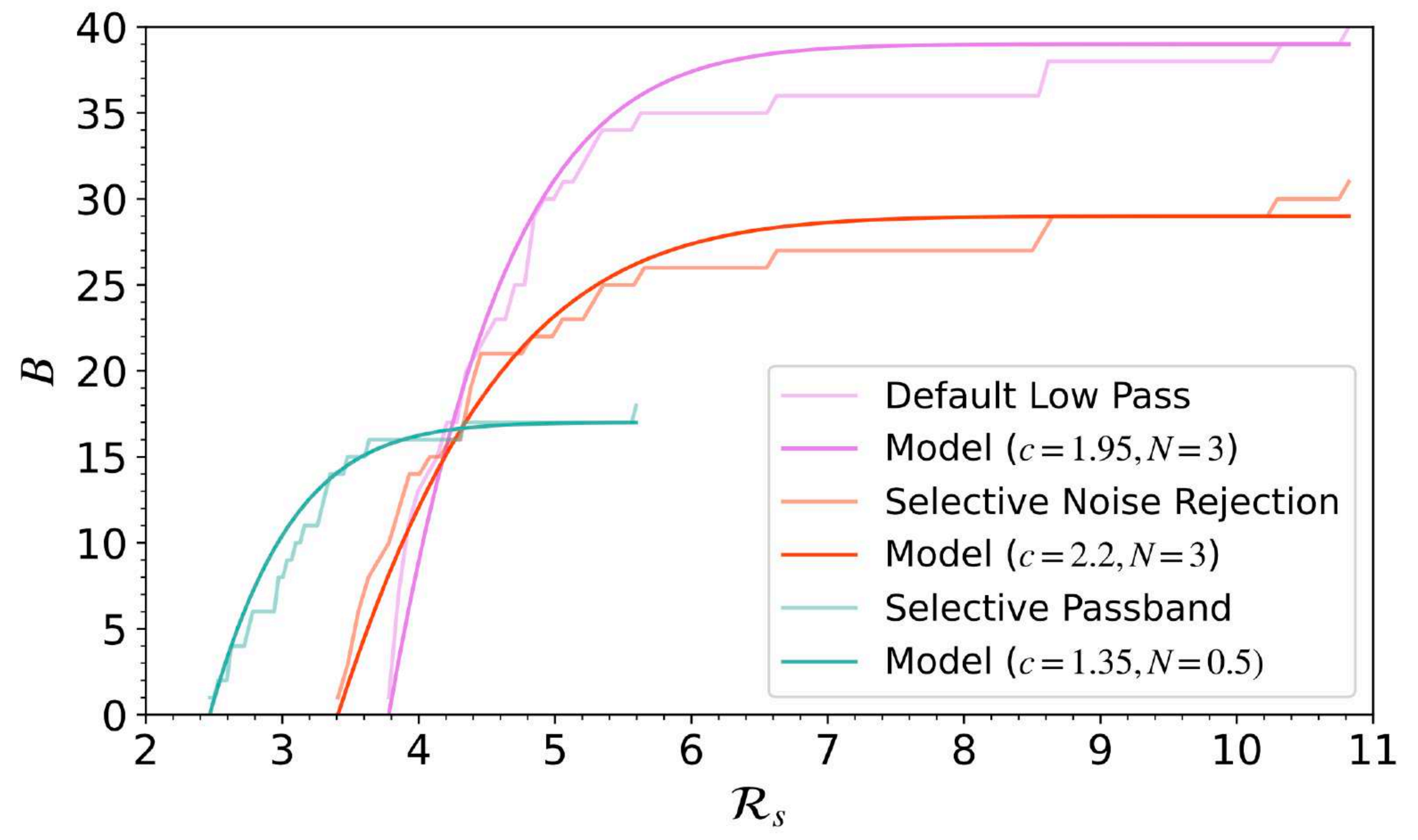
$$\boxed{p_{\text{astro}} = \frac{f(\mathcal{R}_s)}{b(\mathcal{R}_s) + f(\mathcal{R}_s)}}$$

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

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

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

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

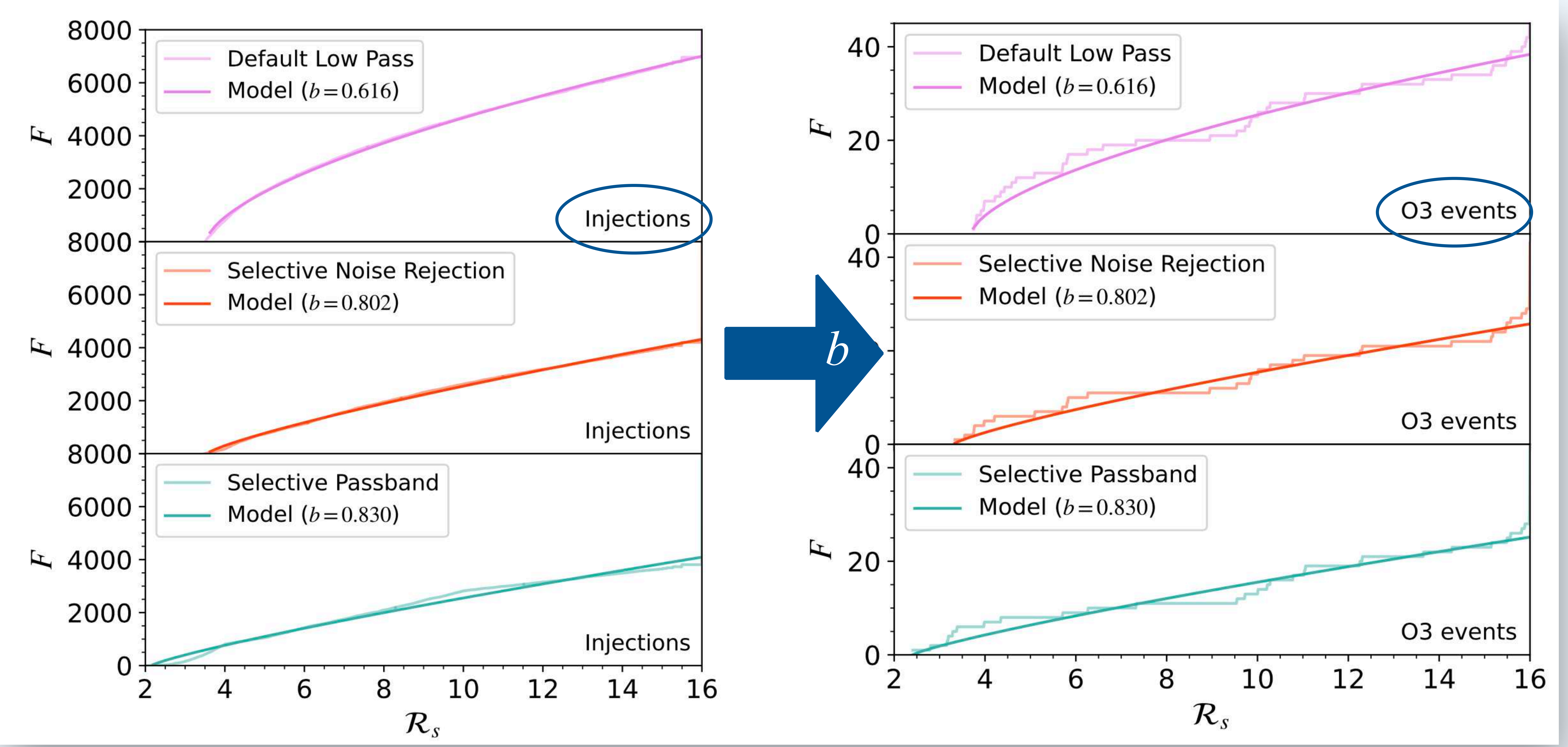


ASTROPHYSICAL PROBABILITY

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

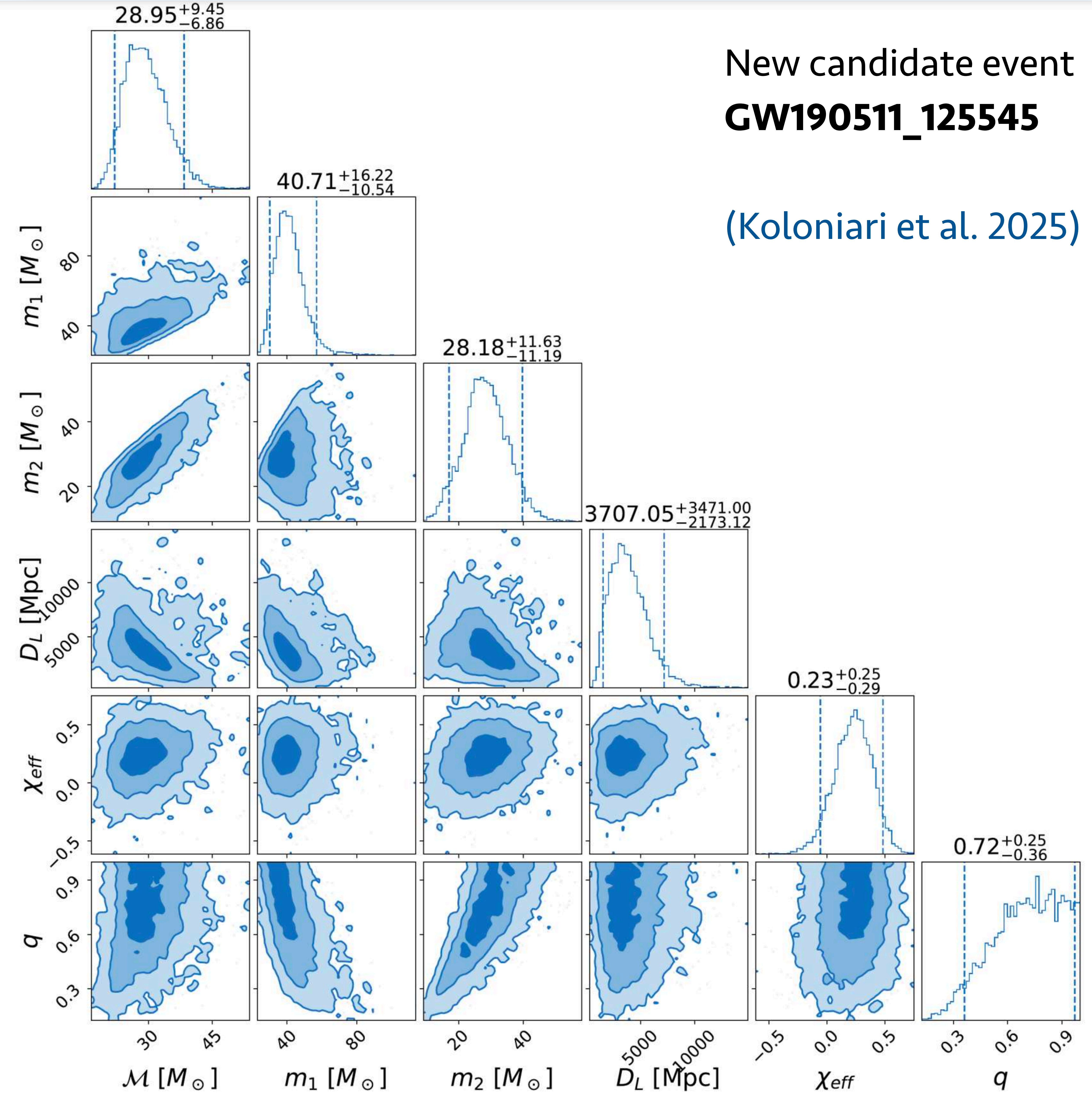
AresGW detects **42 out of 51** known O3 events (*most sensitive pipeline in this mass range*)

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

PARAMETER ESTIMATION USING BILBY

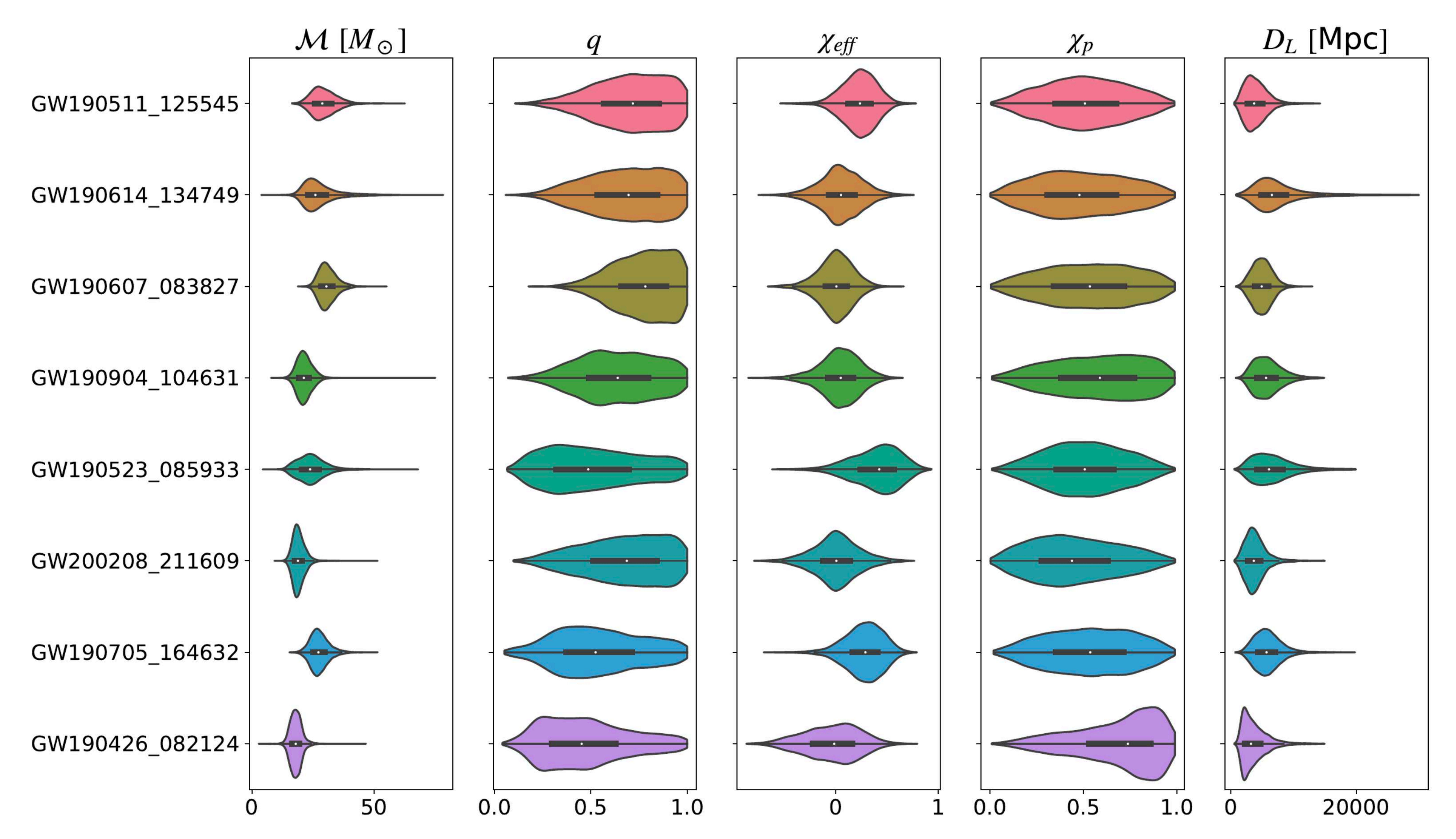


New candidate event

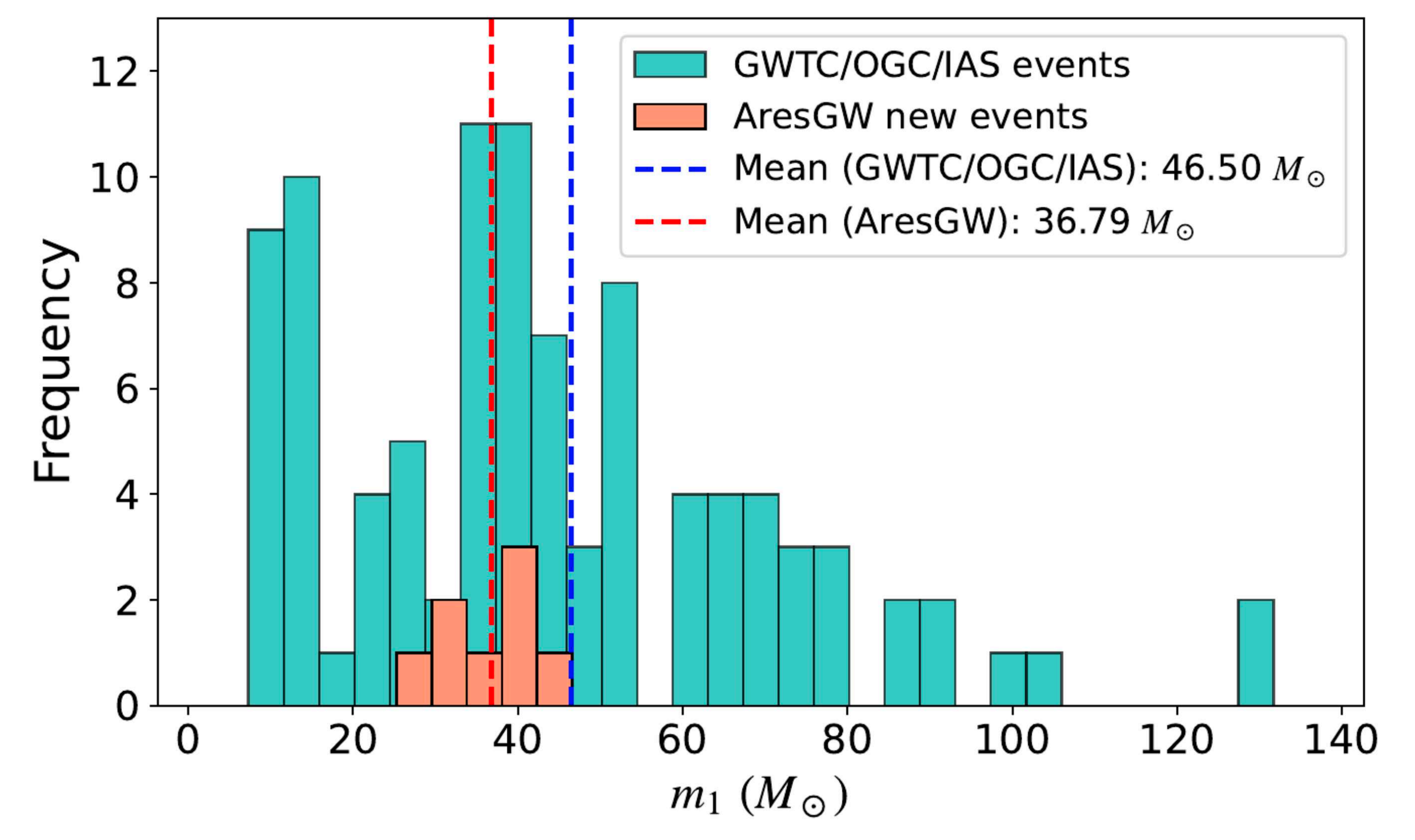
GW190511_125545

(Koloniari et al. 2025)

PARAMETER ESTIMATION USING BILBY



PARAMETER ESTIMATION USING BILBY



RECONSTRUCTED WAVEFORMS FOR NEW EVENTS

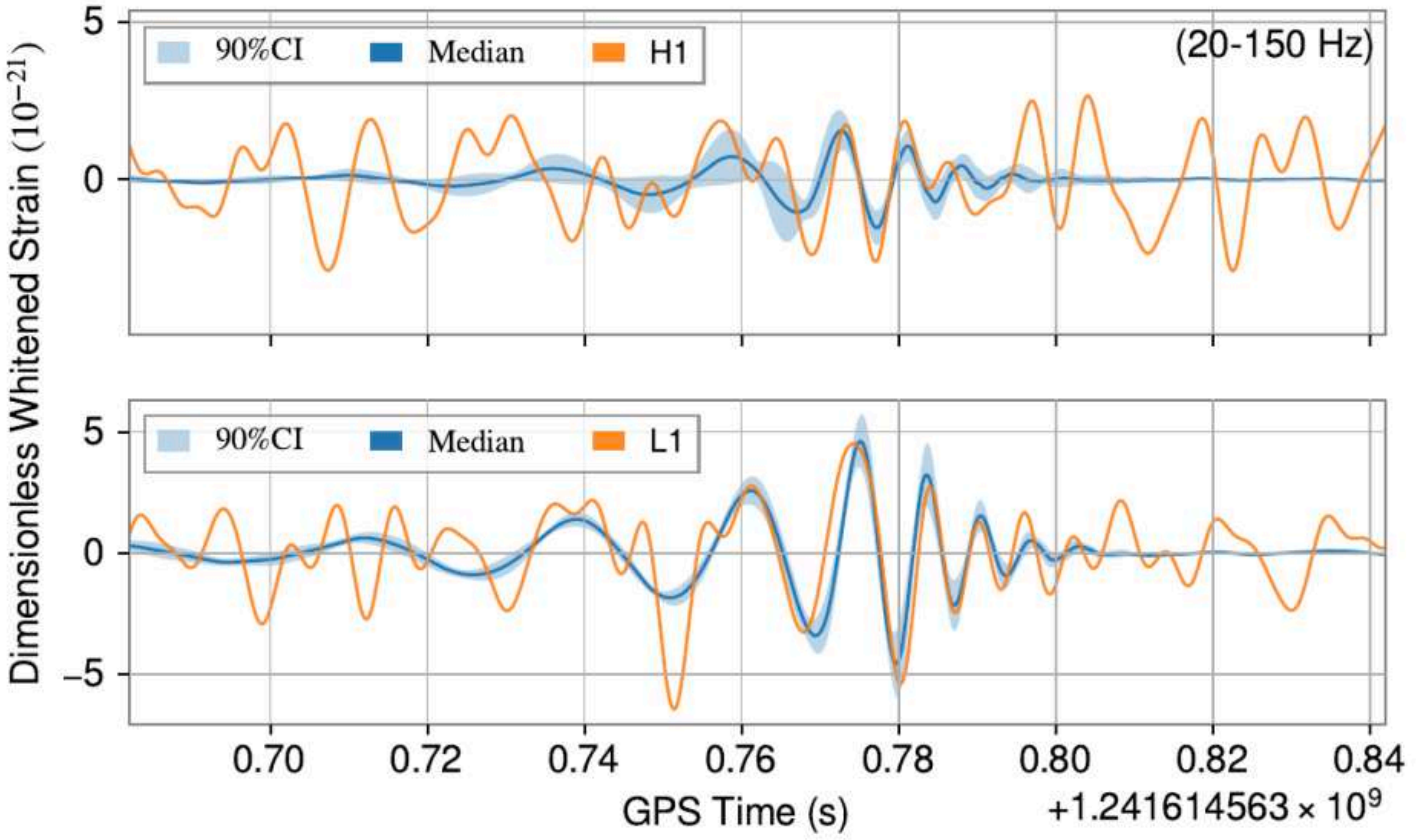


FIG. 35: Whitened, bandpassed strain data and reconstructed waveform for the new event GW190511_125545 identified by AresGW.

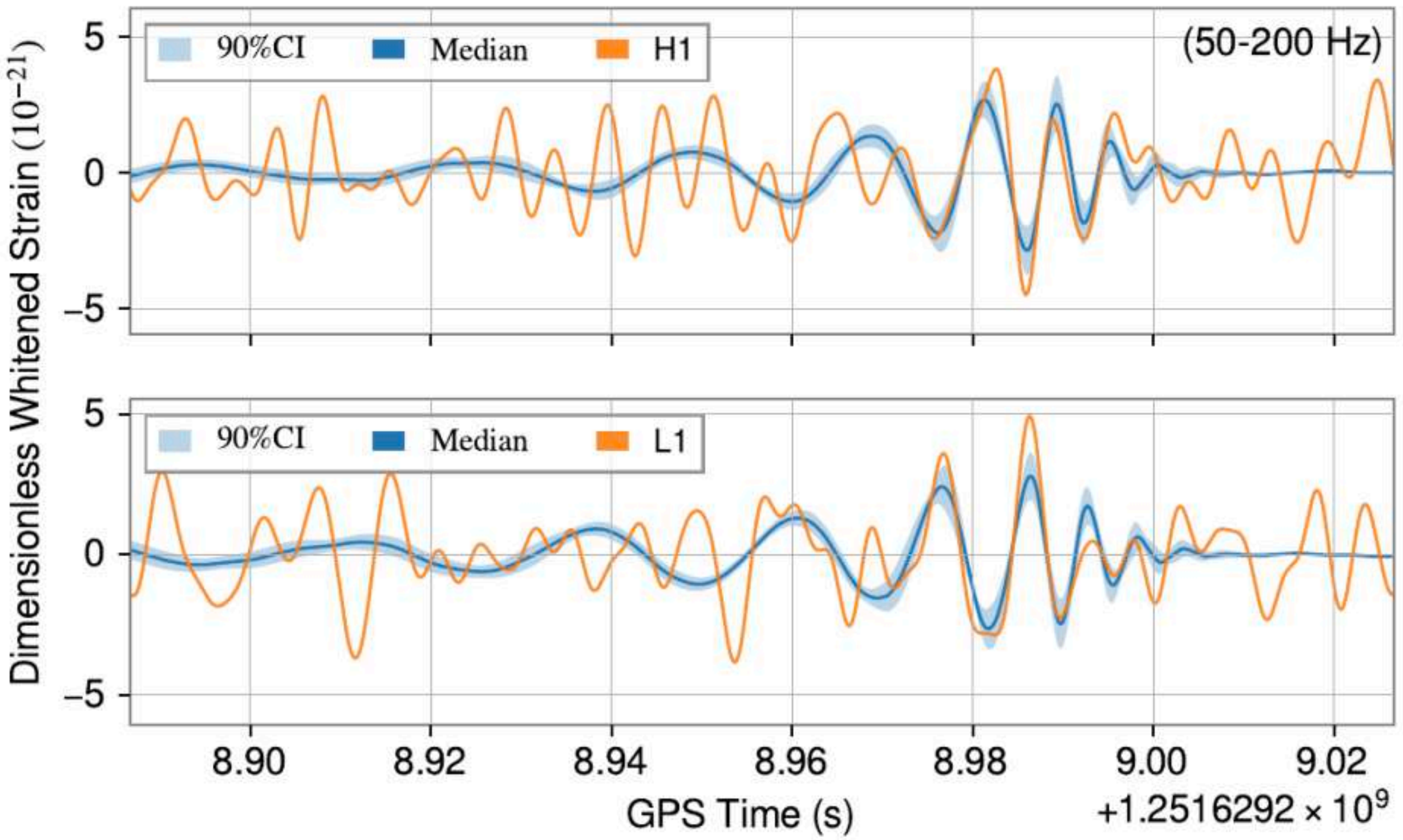
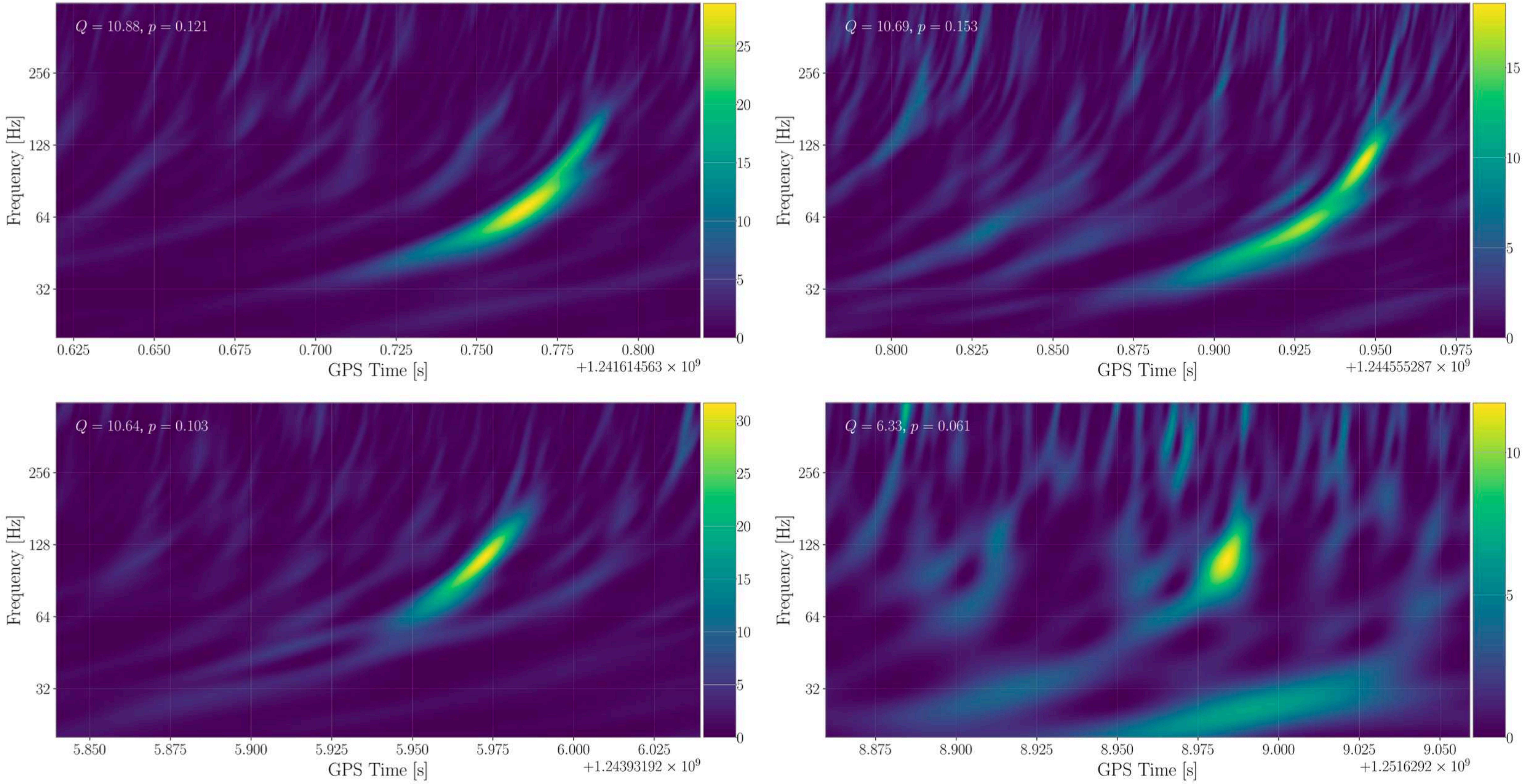


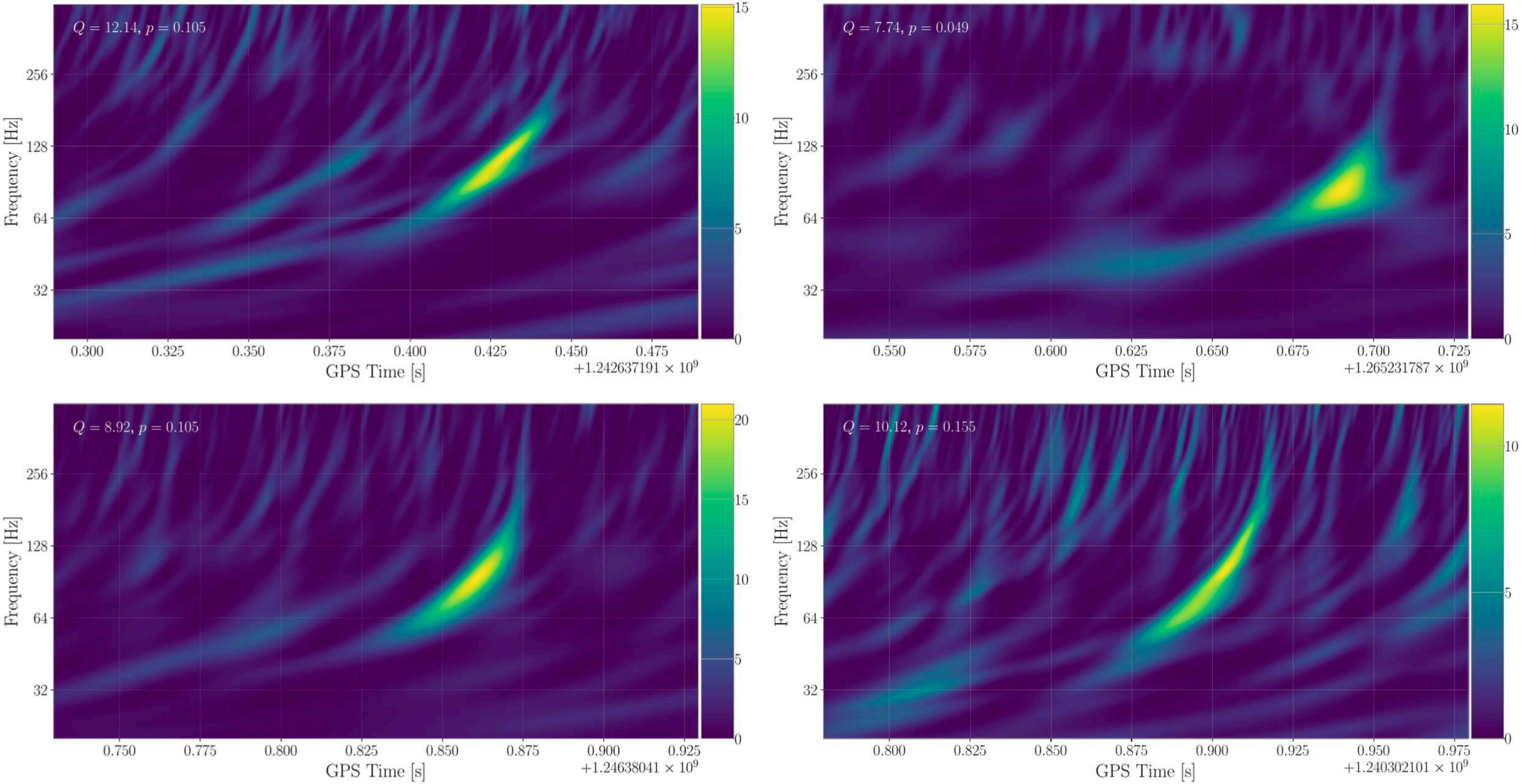
FIG. 38: Same as Fig. 35, but for the new event GW190904_104631.

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

ARESGW NEW CANDIDATE EVENTS

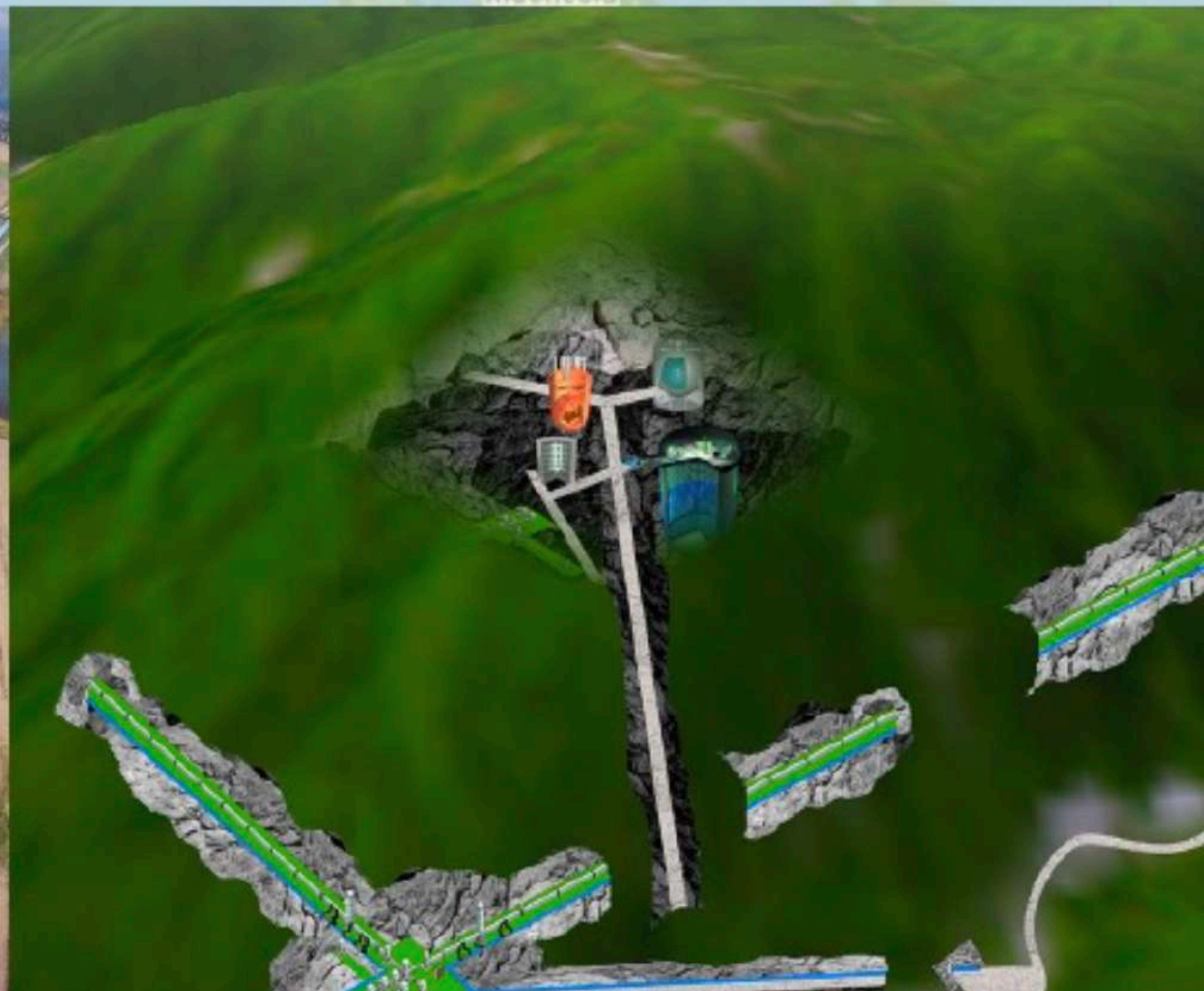


ARESGW NEW CANDIDATE EVENTS

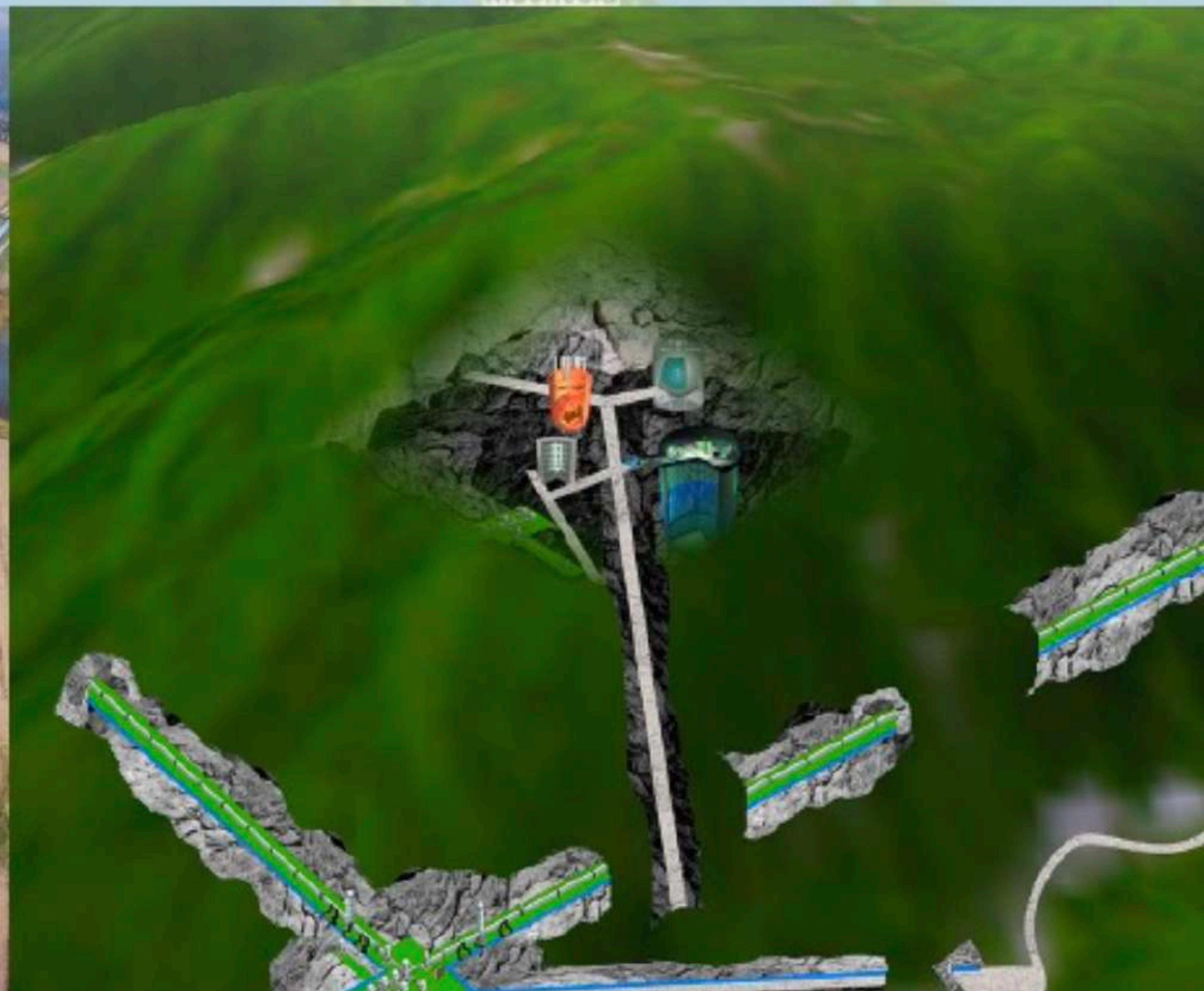


PART B.
BINARY NEUTRON STAR POST-MERGER PHASE

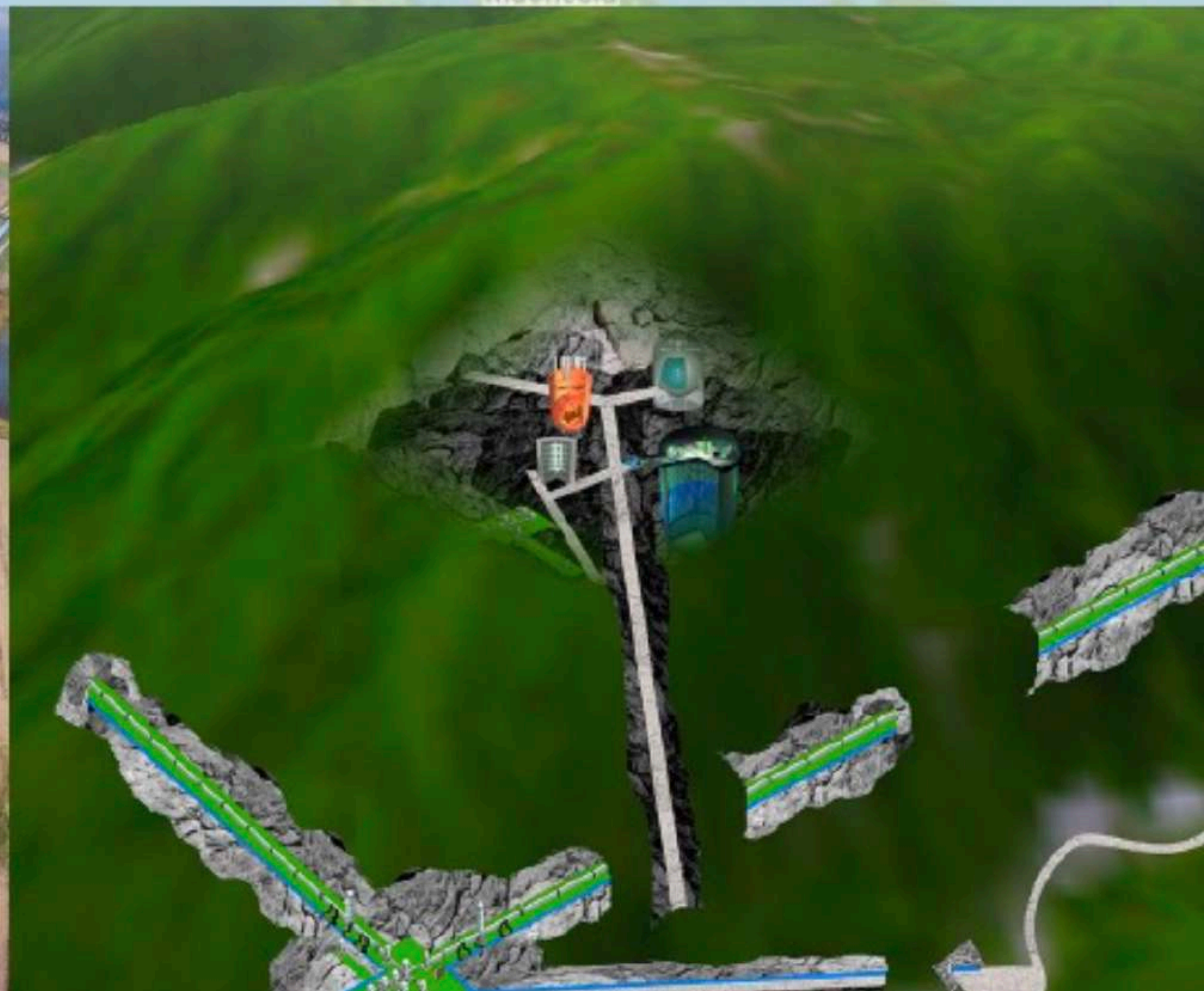
INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



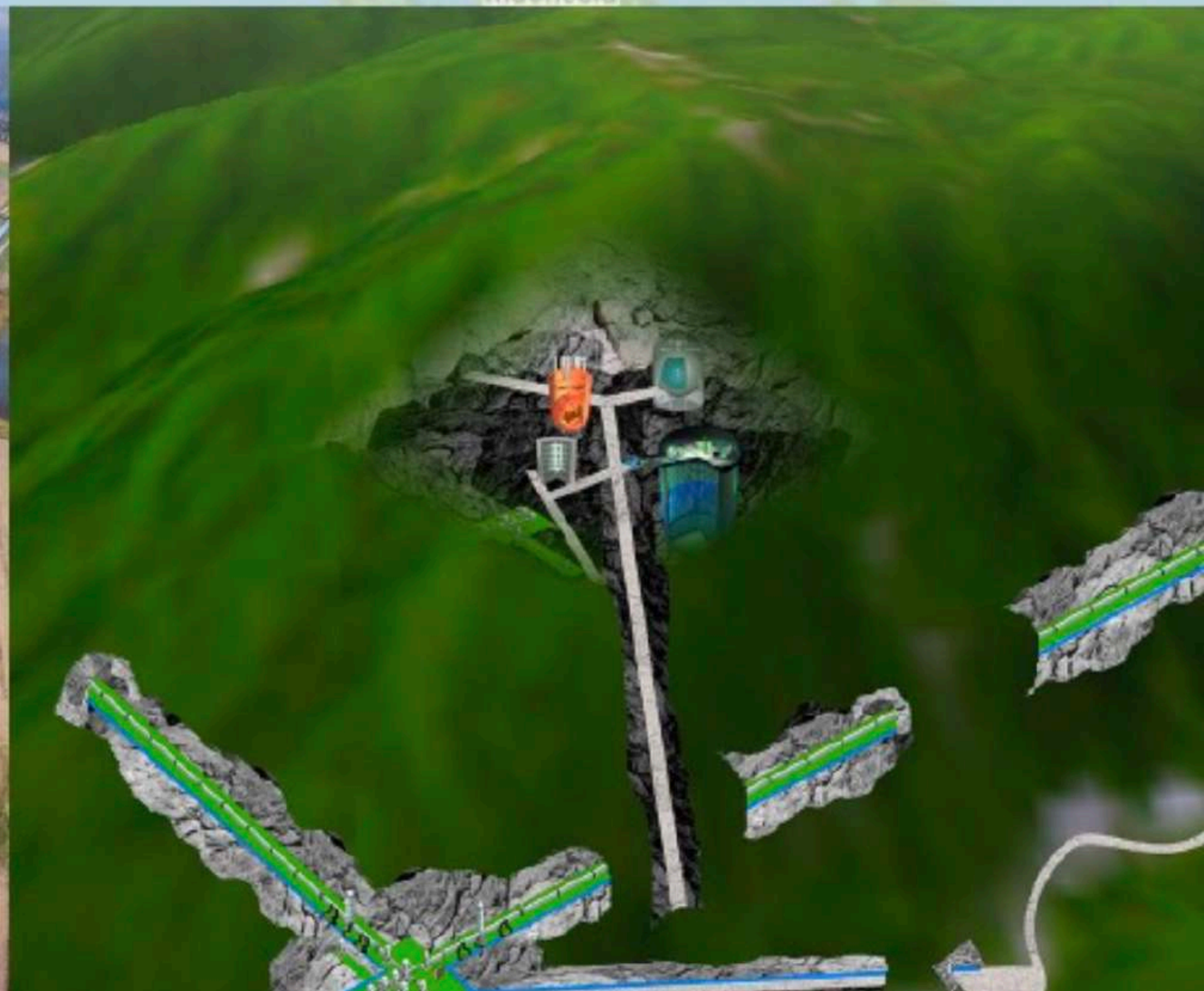
INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)



INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)

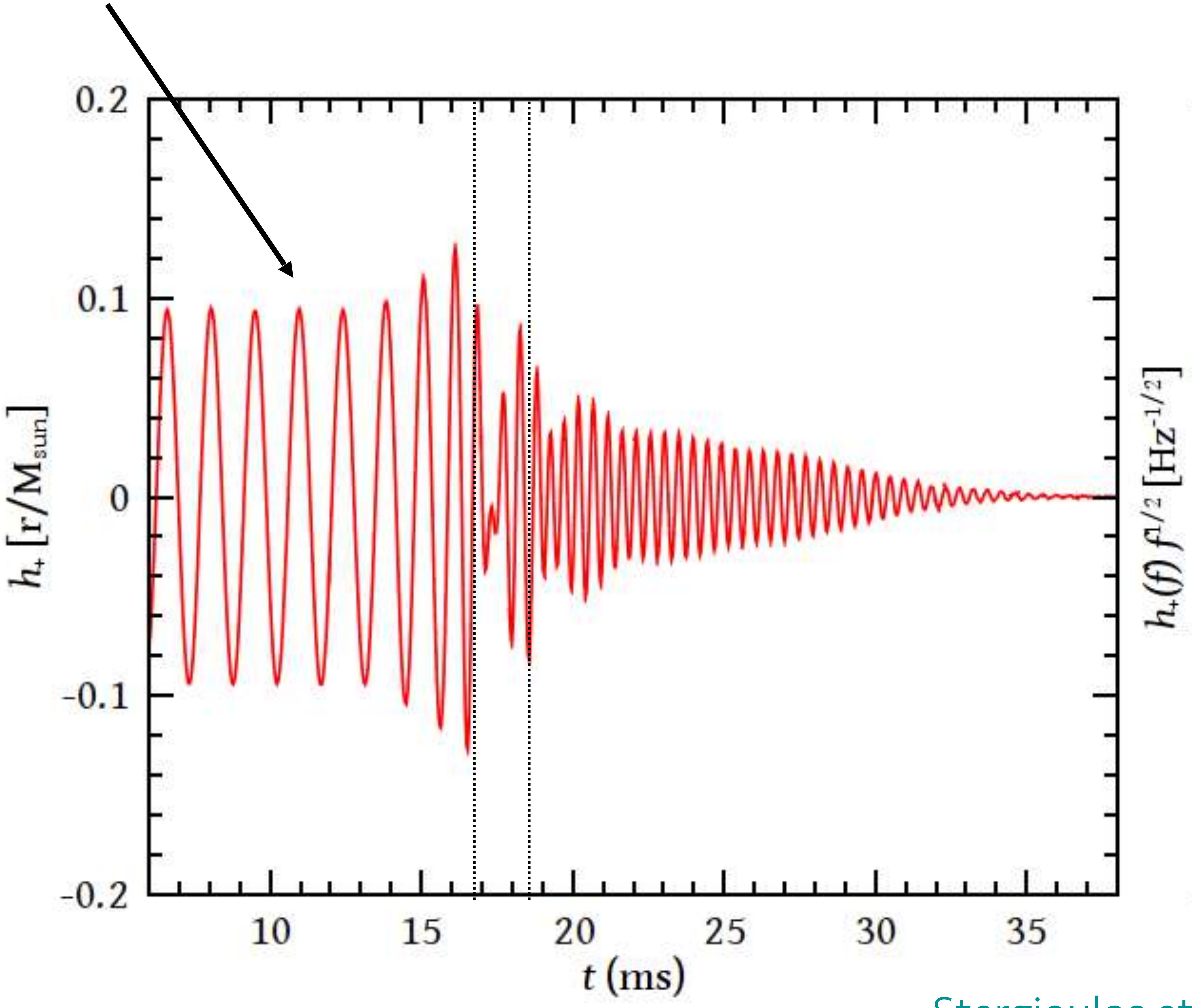


INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)

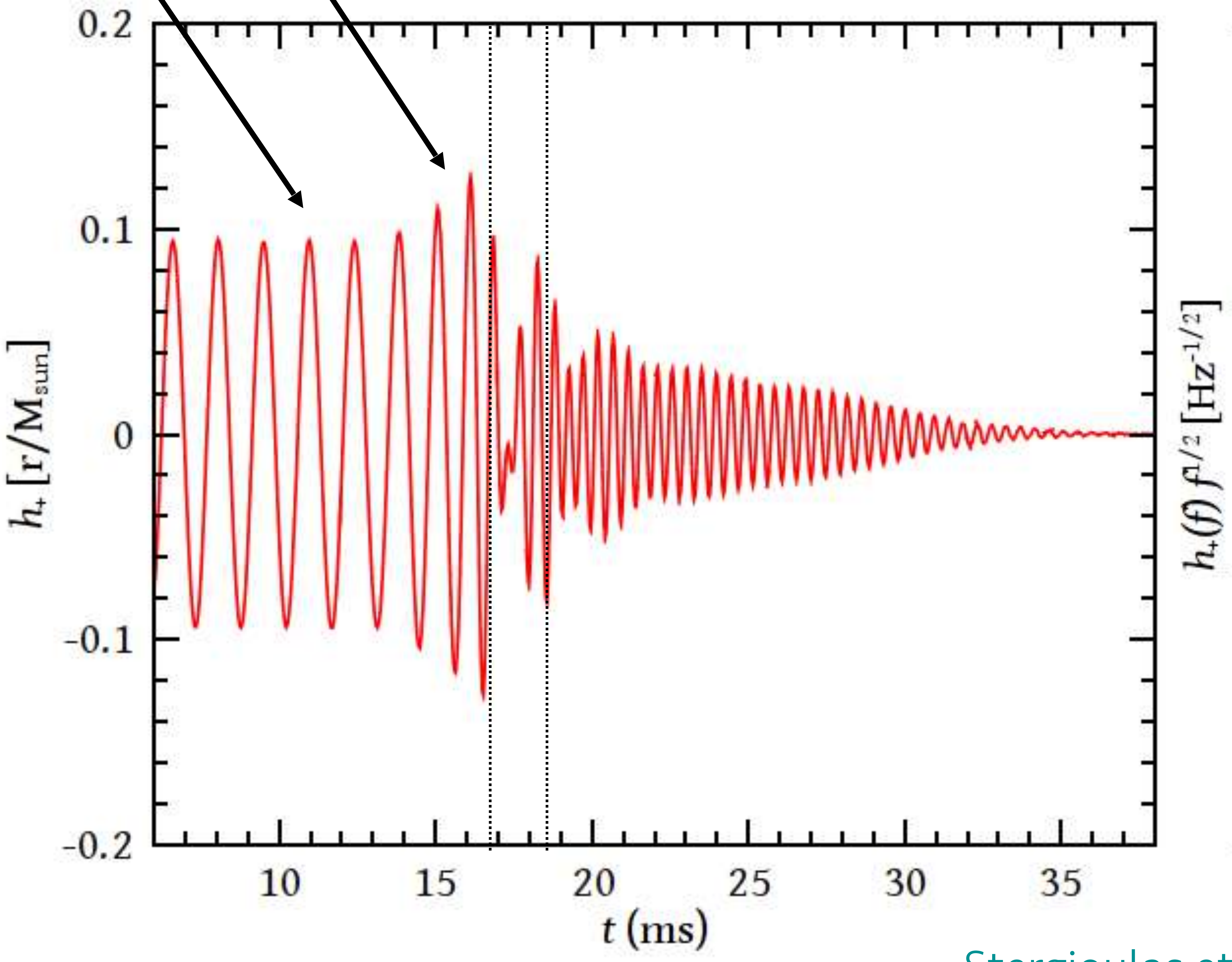
Time domain: three distinct phases of the GW signal:

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



Stergioulas et al. (2011)

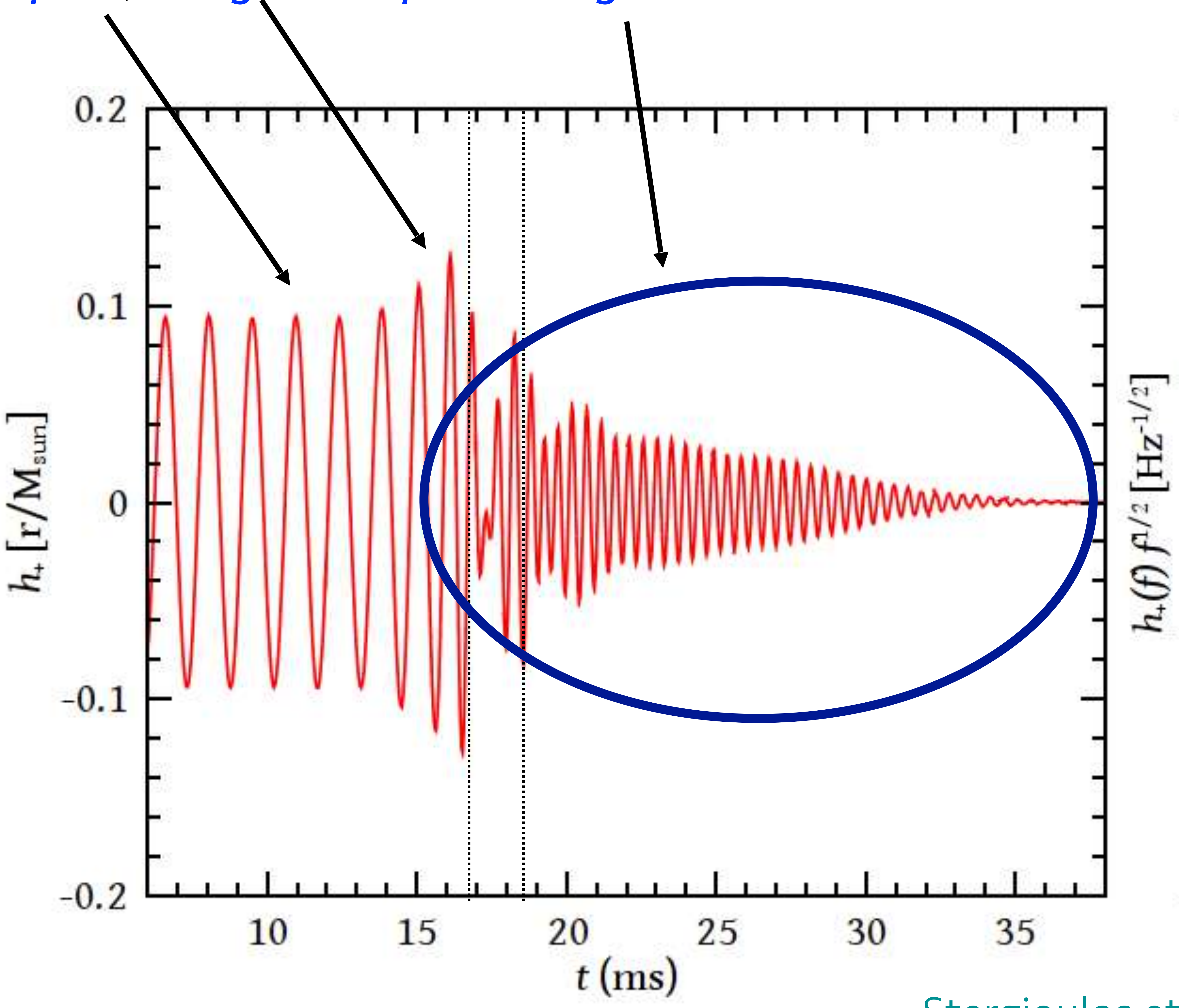
Time domain: three distinct phases of the GW signal:
inspiral, *merger* and *post-merger oscillations*.



Stergioulas et al. (2011)

Time domain: three distinct phases of the GW signal:

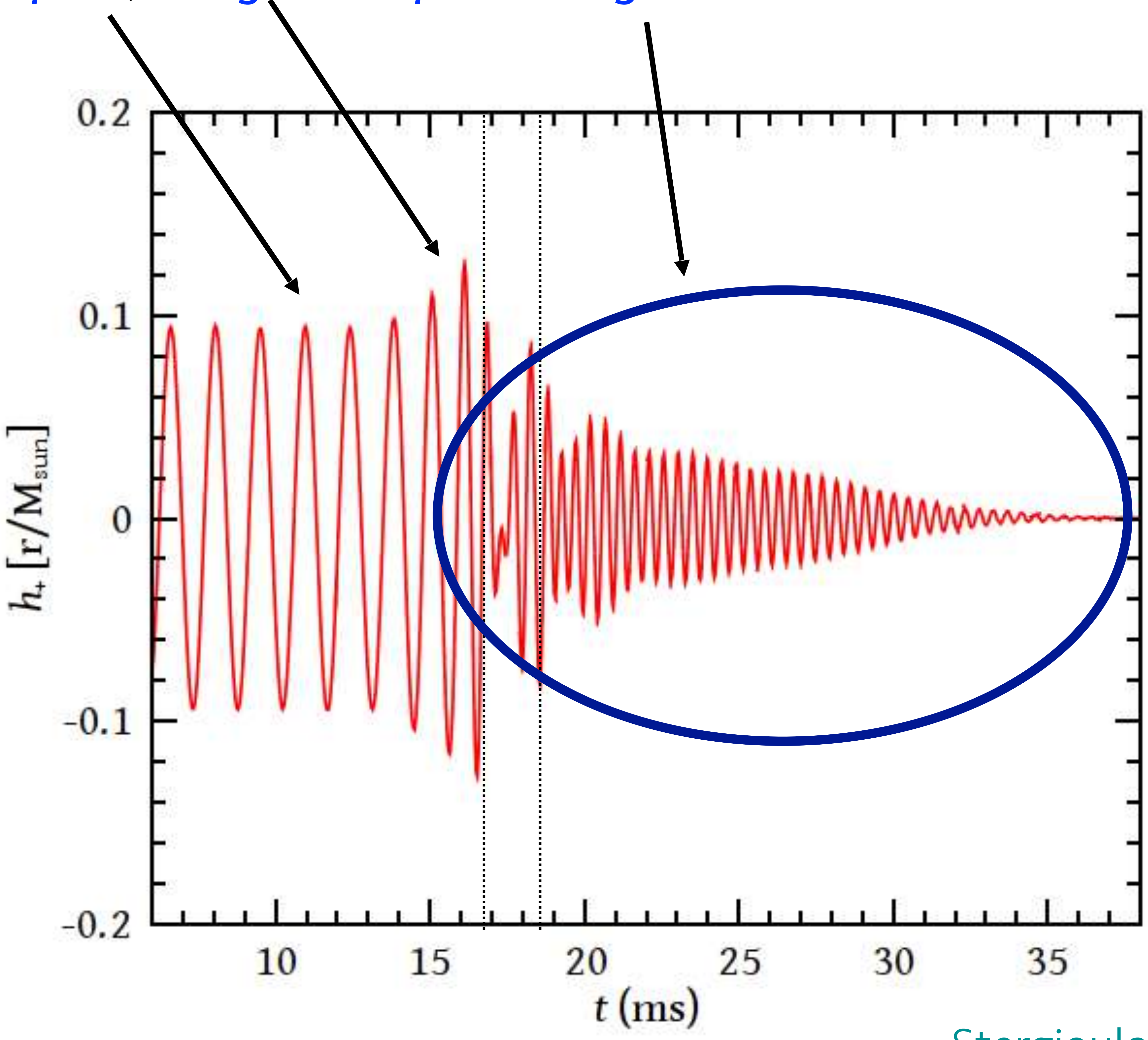
inspiral, *merger* and *post-merger oscillations*.



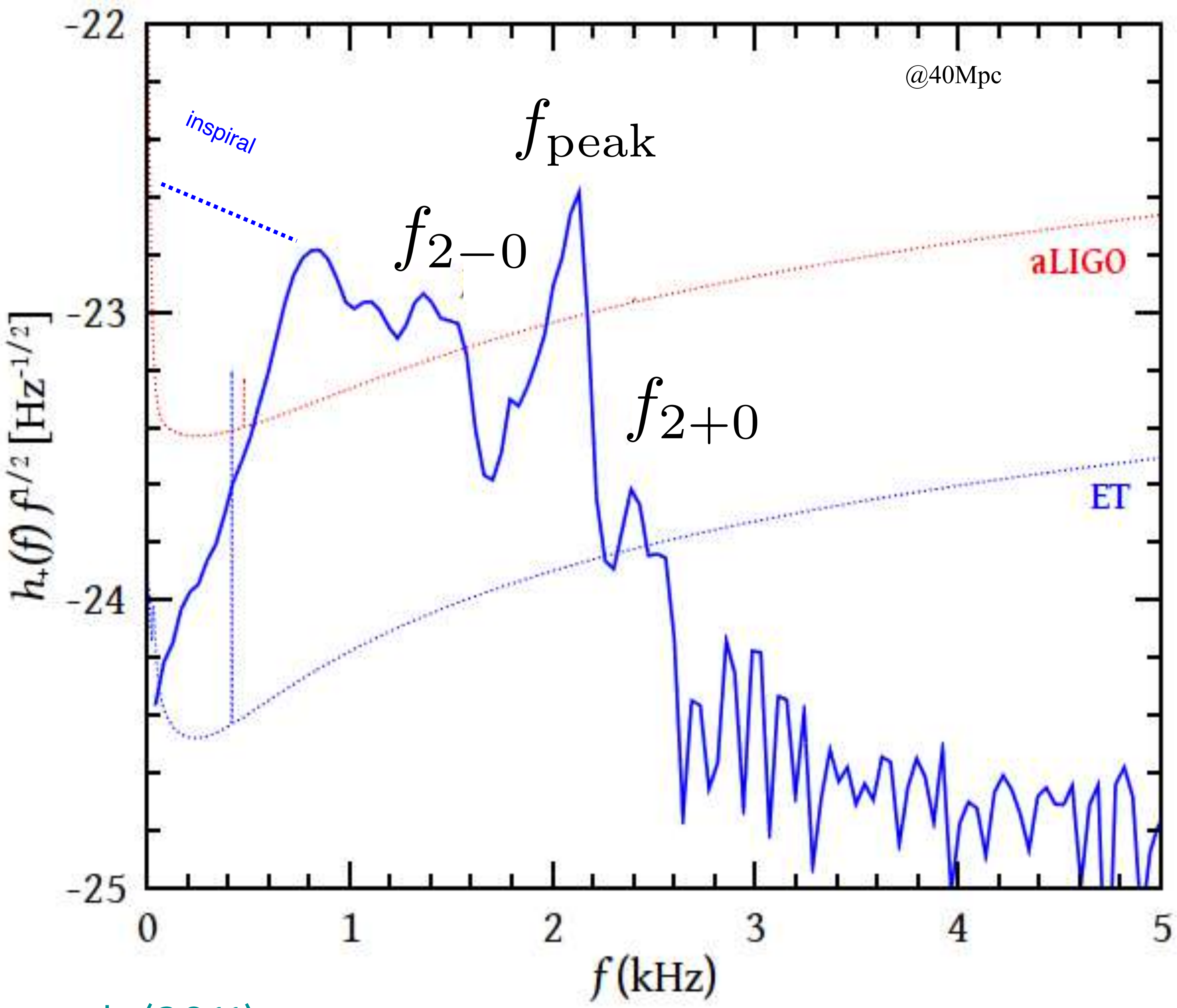
Stergioulas et al. (2011)

POST-MERGER PHASE IN BNS MERGERS

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

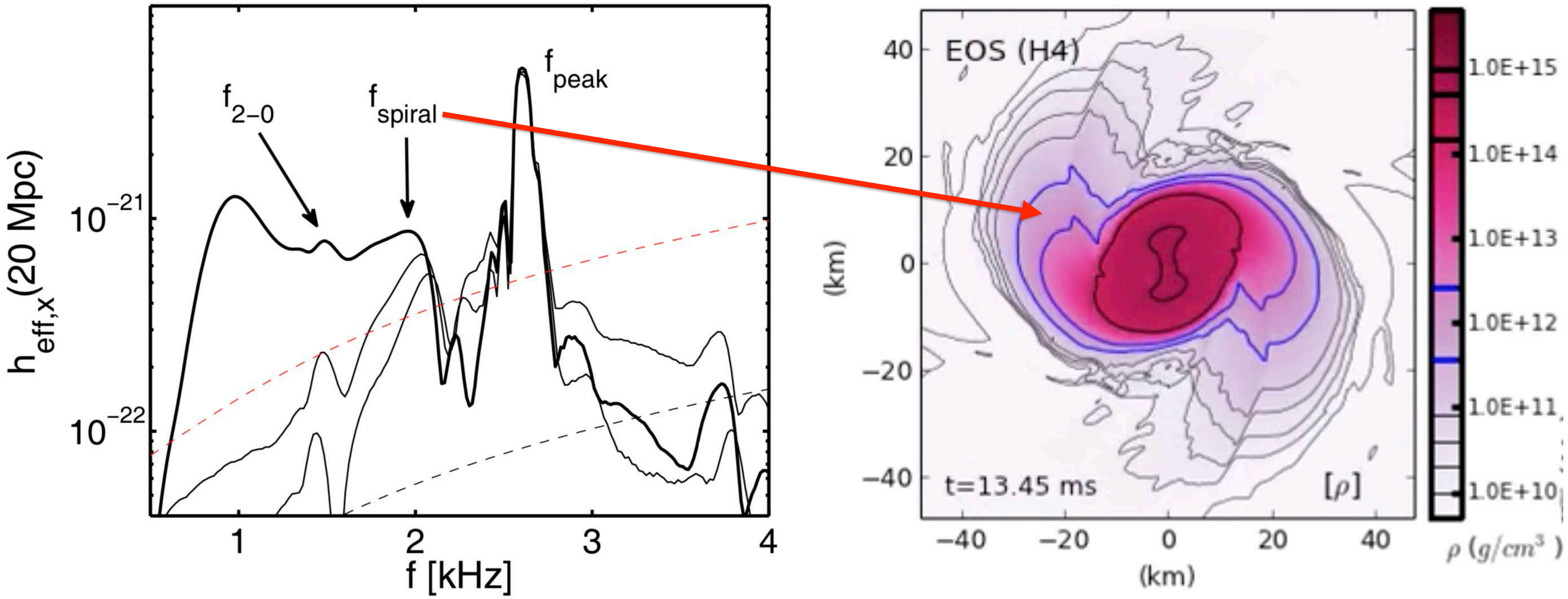


Frequency domain:
f_{peak} : *l=m=2 fundamental mode.*
$f_{2\text{-}0}, f_{2+0}$: *nonlinear combination tones*



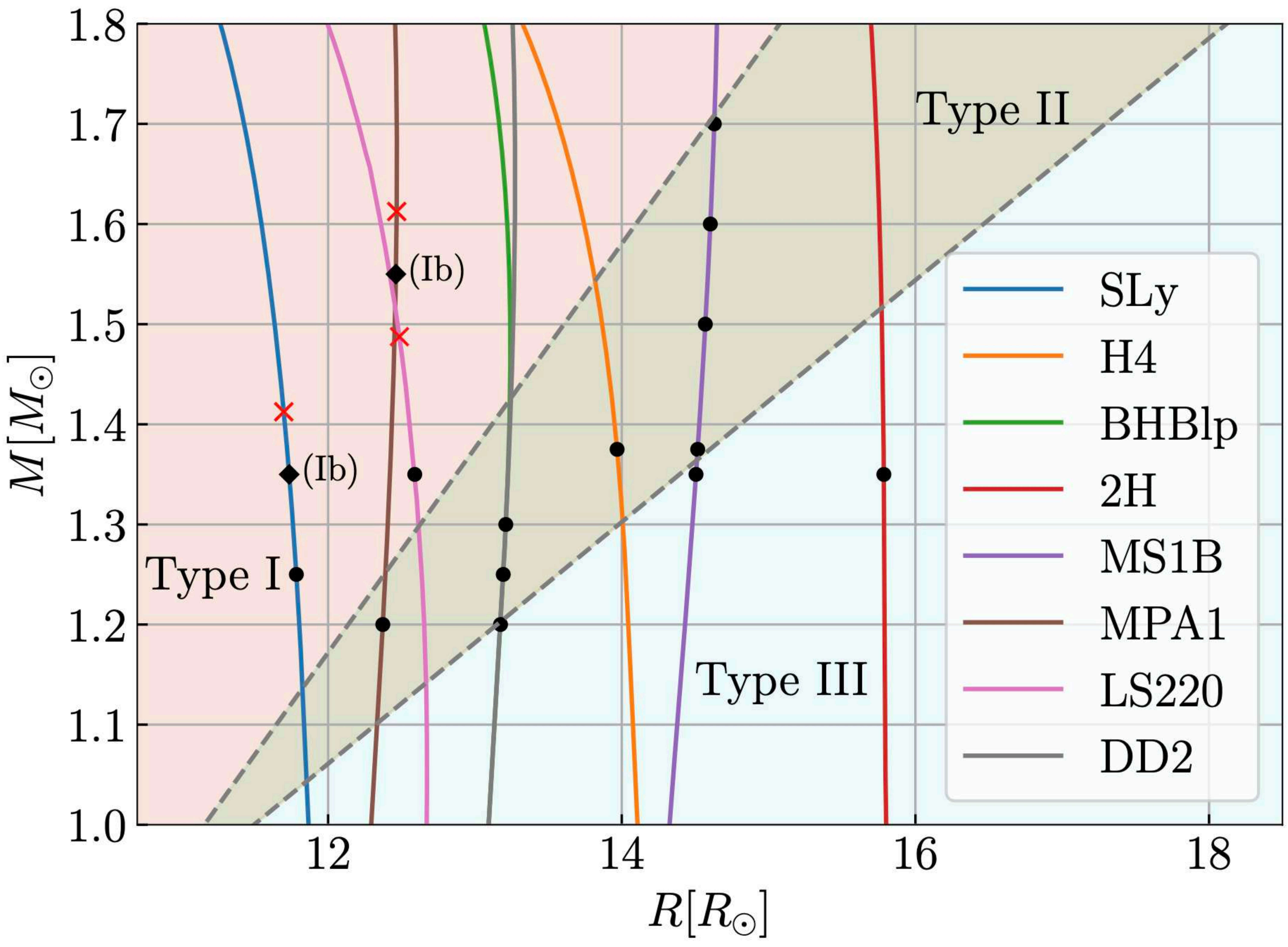
Stergioulas et al. (2011)

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

$$f_{\text{peak}}\, M_{\text{chirp}} = 1.392 - 0.108 M_{\text{chirp}} + 51.70 \tilde{\Lambda}^{-1/2}$$

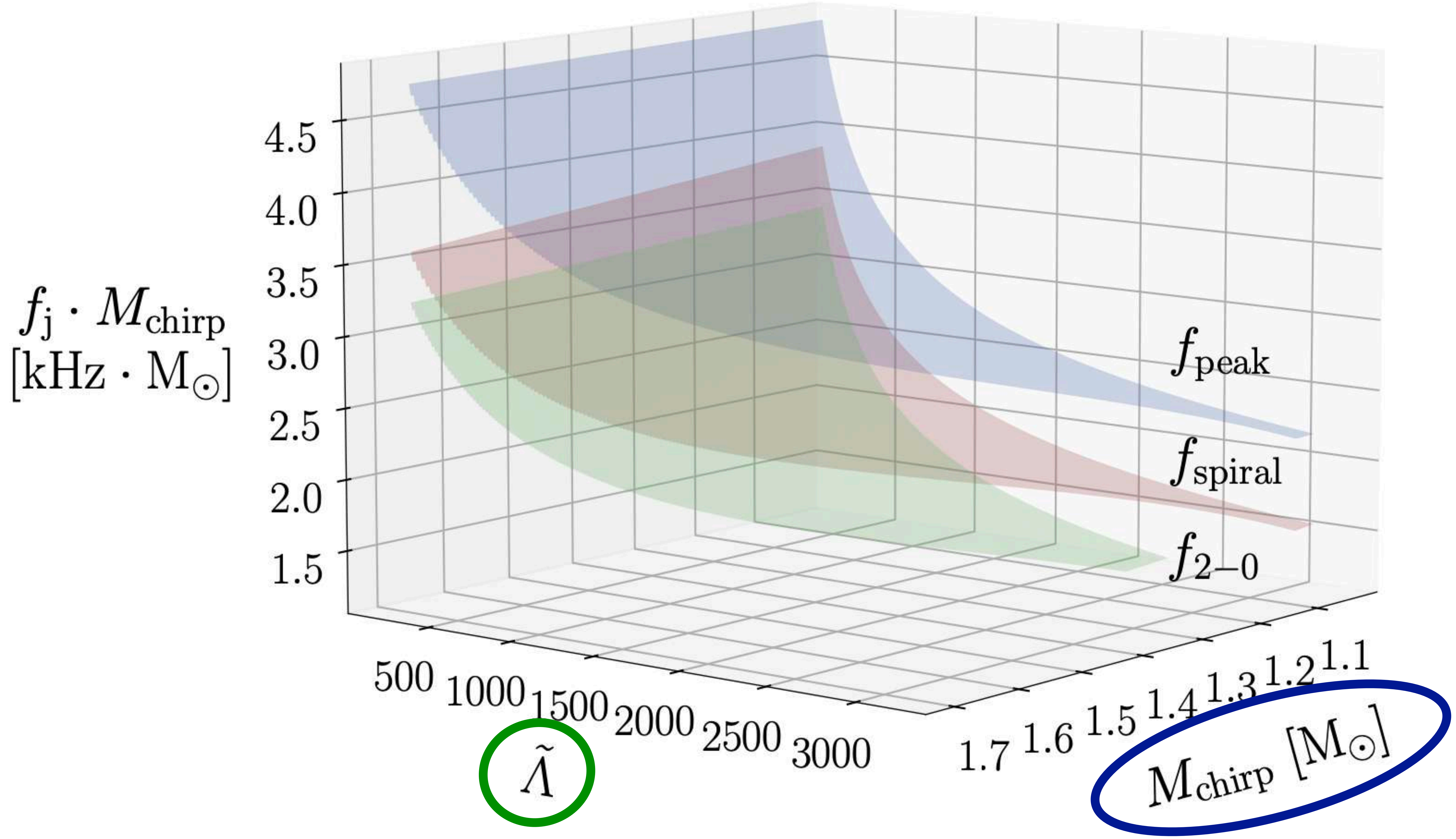
Vretinaris, Bauswein & Stergioulas (2020)

$$f_{2-0} M_{\text{chirp}} = 0.558 - 0.406 M_{\text{chirp}} + 48.696 \tilde{\Lambda}^{-1/2}$$

Vretinaris et al. (2025)

$$f_{\text{spiral}}\, M_{\text{chirp}} = 1.2 - 0.442 M_{\text{chirp}} + 45.819 \tilde{\Lambda}^{-1/2}$$

Vretinaris et al. (2025)

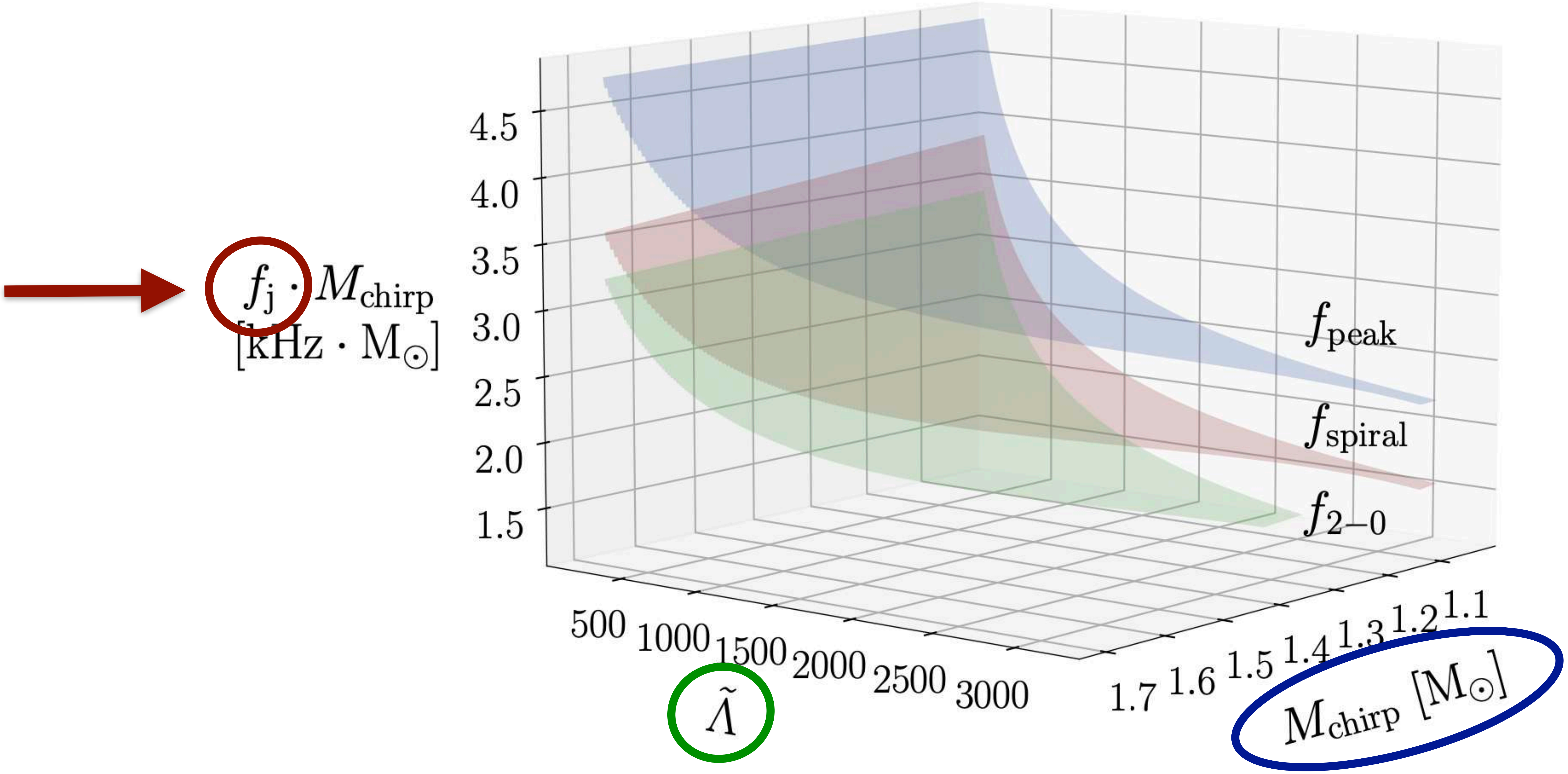


EMPIRICAL RELATIONS OF POST-MERGER FREQUENCIES

$$f_{\mathrm{peak}}\, M_{\mathrm{chirp}} = 1.392 - 0.108 M_{\mathrm{chirp}} + 51.70 \tilde{\Lambda}^{-1/2}$$ 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)

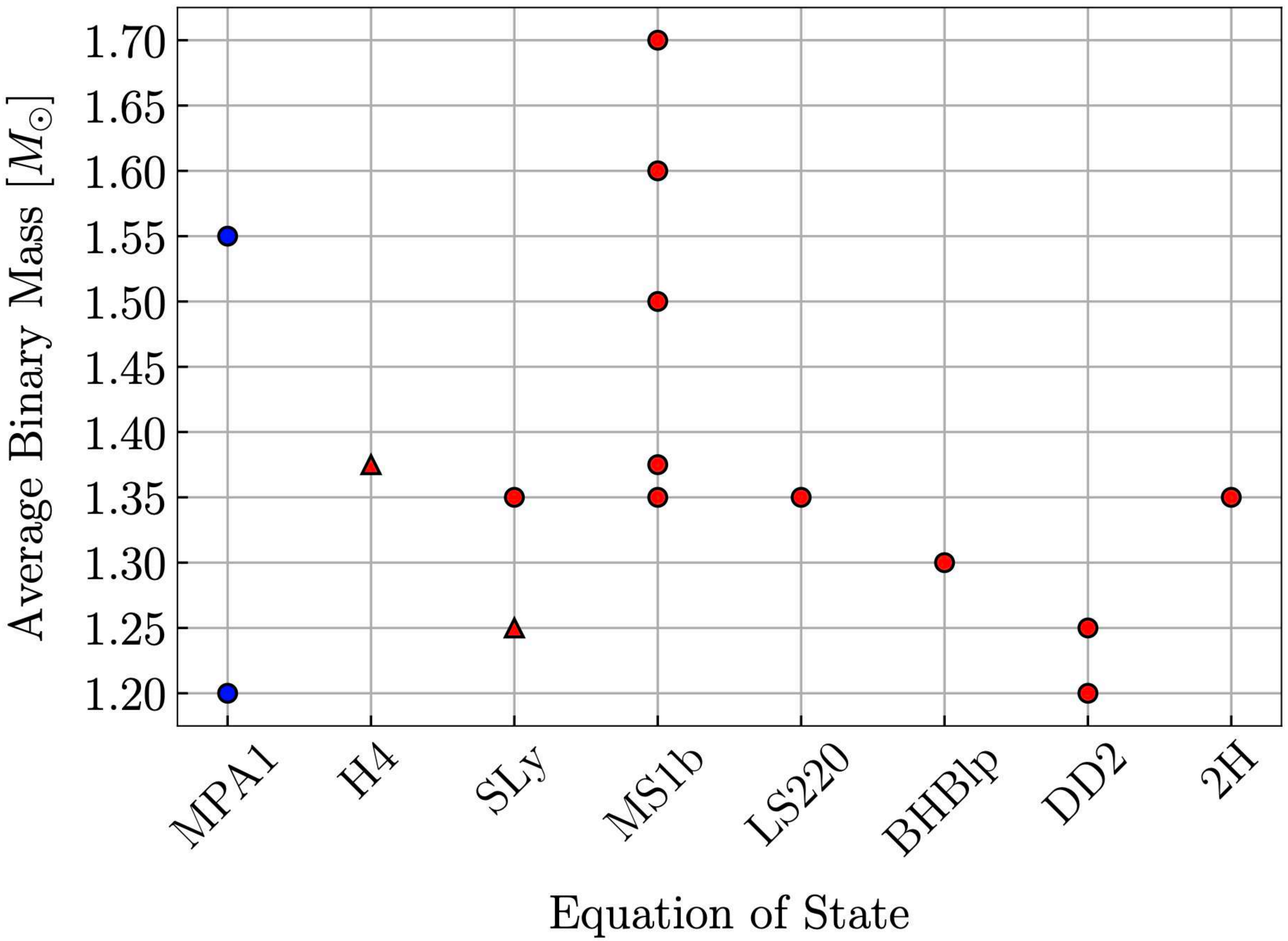
$$f_{\mathrm{spiral}}\, M_{\mathrm{chirp}} = 1.2 - 0.442 M_{\mathrm{chirp}} + 45.819 \tilde{\Lambda}^{-1/2}$$ Vretinaris et al. (2025)



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]



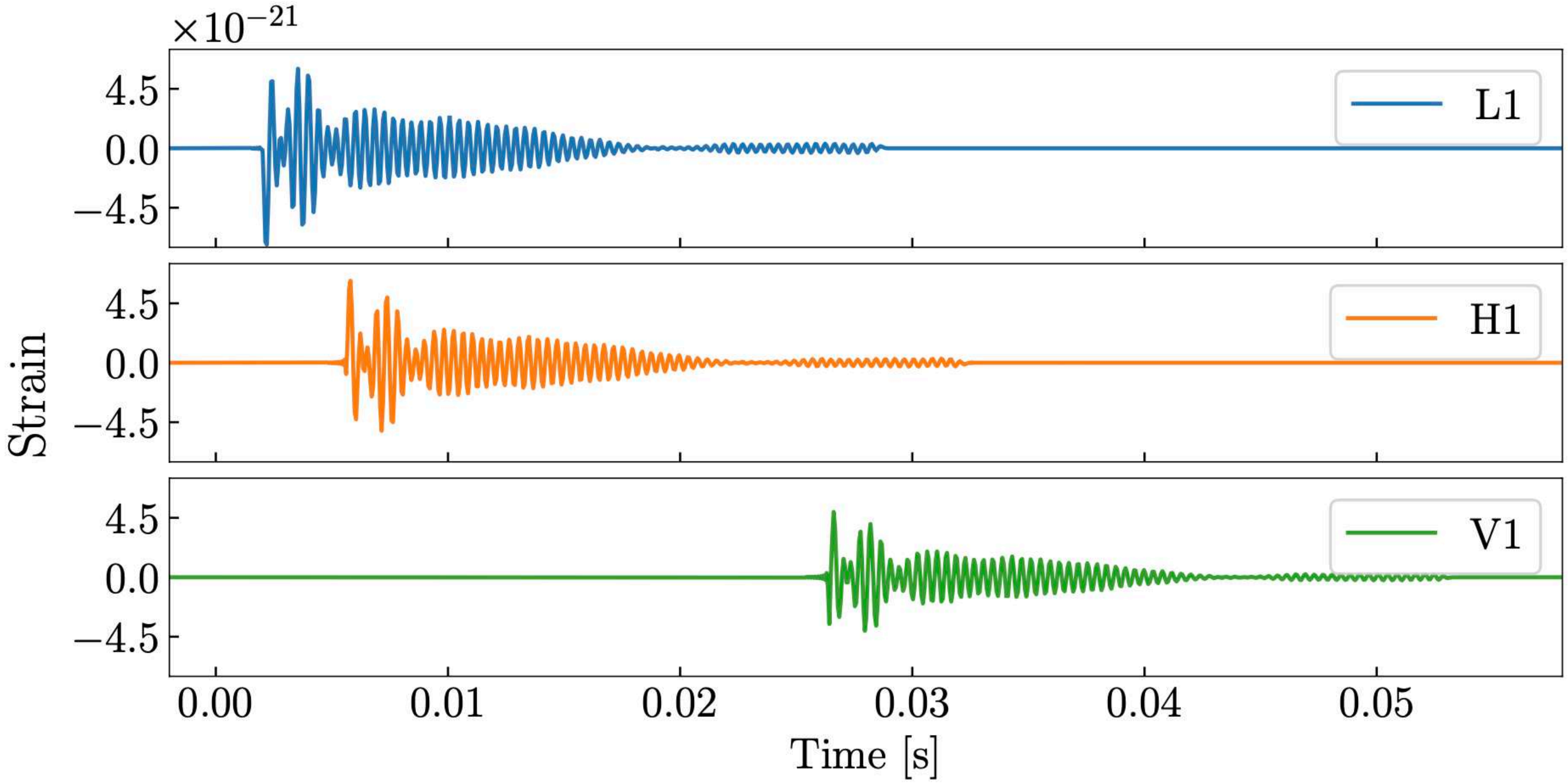
Extension of the set of 9 waveforms used in previous work by Easter et al. (2020)

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:

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

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}.

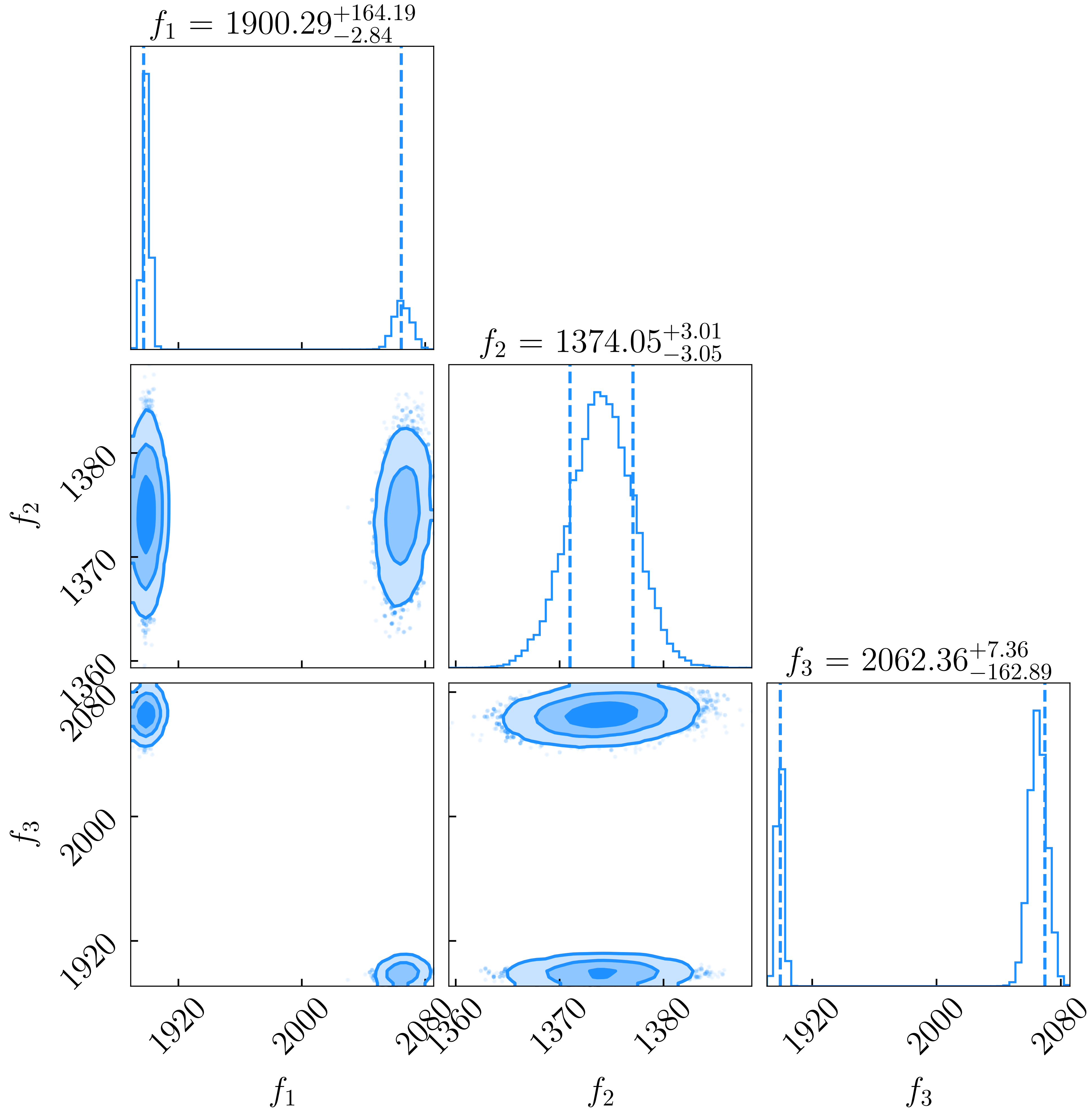
where σ = 3 x max error of empirical relations. For $f_{\rm peak}$: Gaussian ; for f_4: flat in [$f_{\rm peak}$+0.3kHz, 5 kHz].

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)

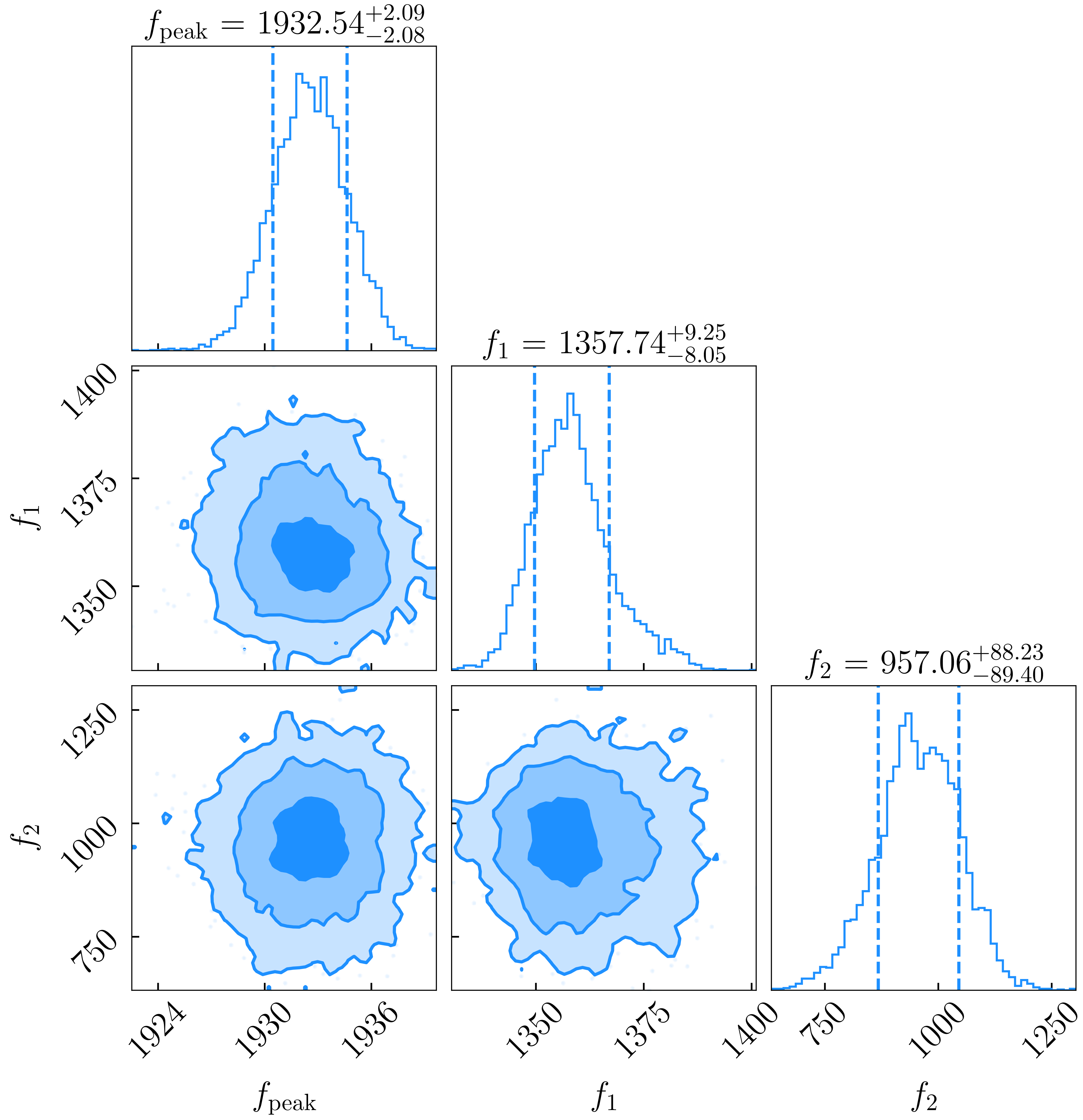
POSTERIOR DISTRIBUTIONS

EOS 2H 1.35+1.35, SNR = 50

using flat priors



using informed priors (through empirical relations)

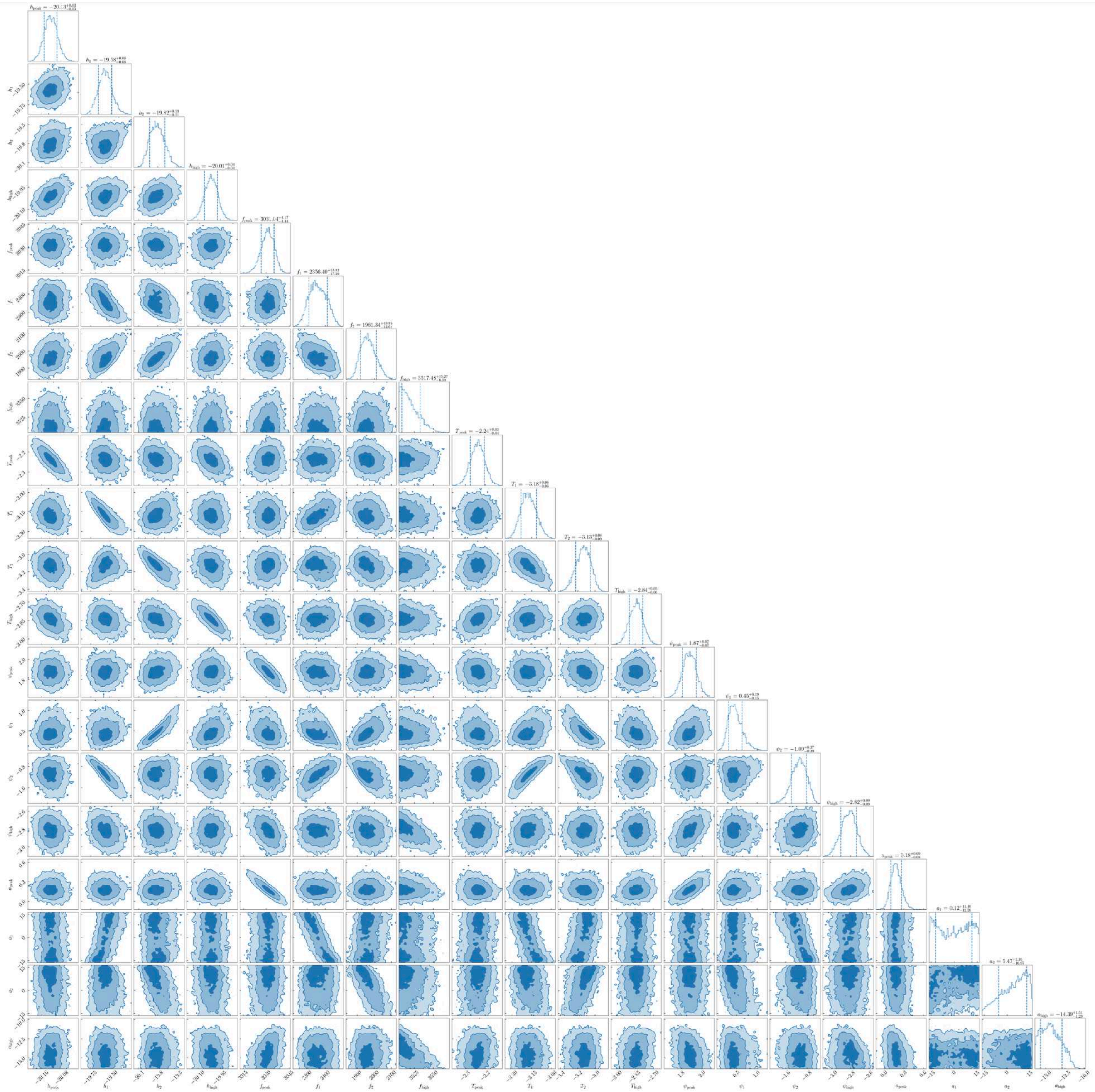


POSTERIOR DISTRIBUTIONS

EOS MPA 1.55+1.55, SNR = 50

using informed priors

all parameters



Preconditioned Monte Carlo Sampling (Karamanis et al. 2022)

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

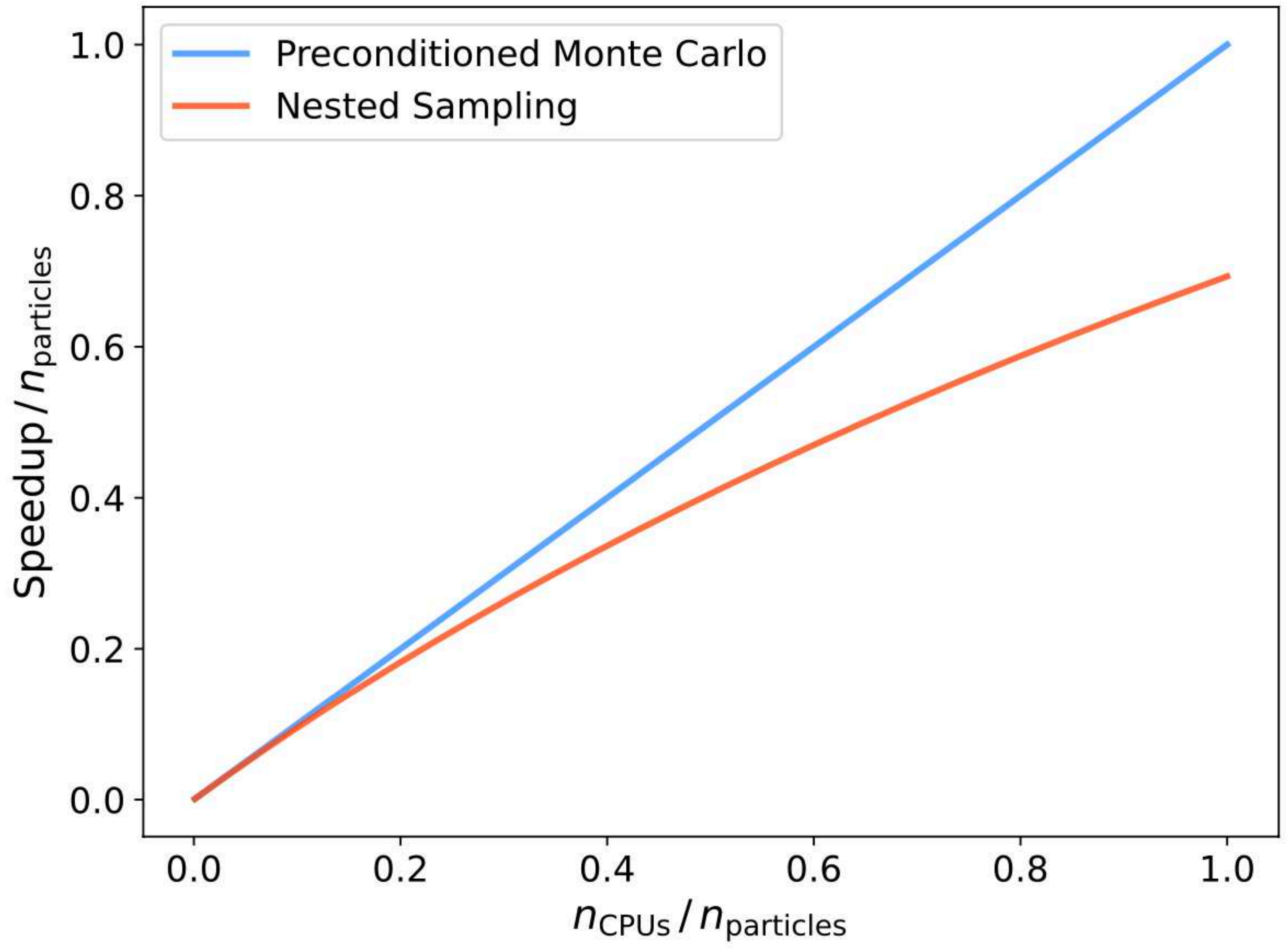
$$p_t(\boldsymbol{\theta} \mid d) \propto \mathcal{L}^{\beta_t}(\boldsymbol{\theta} \mid d)\pi(\boldsymbol{\theta})$$

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.

pocoMC

https://github.com/minaskar/pocomc



PocoMC is highly parallelizable

Preconditioned Monte Carlo Sampling (Karamanis et al. 2022)

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

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

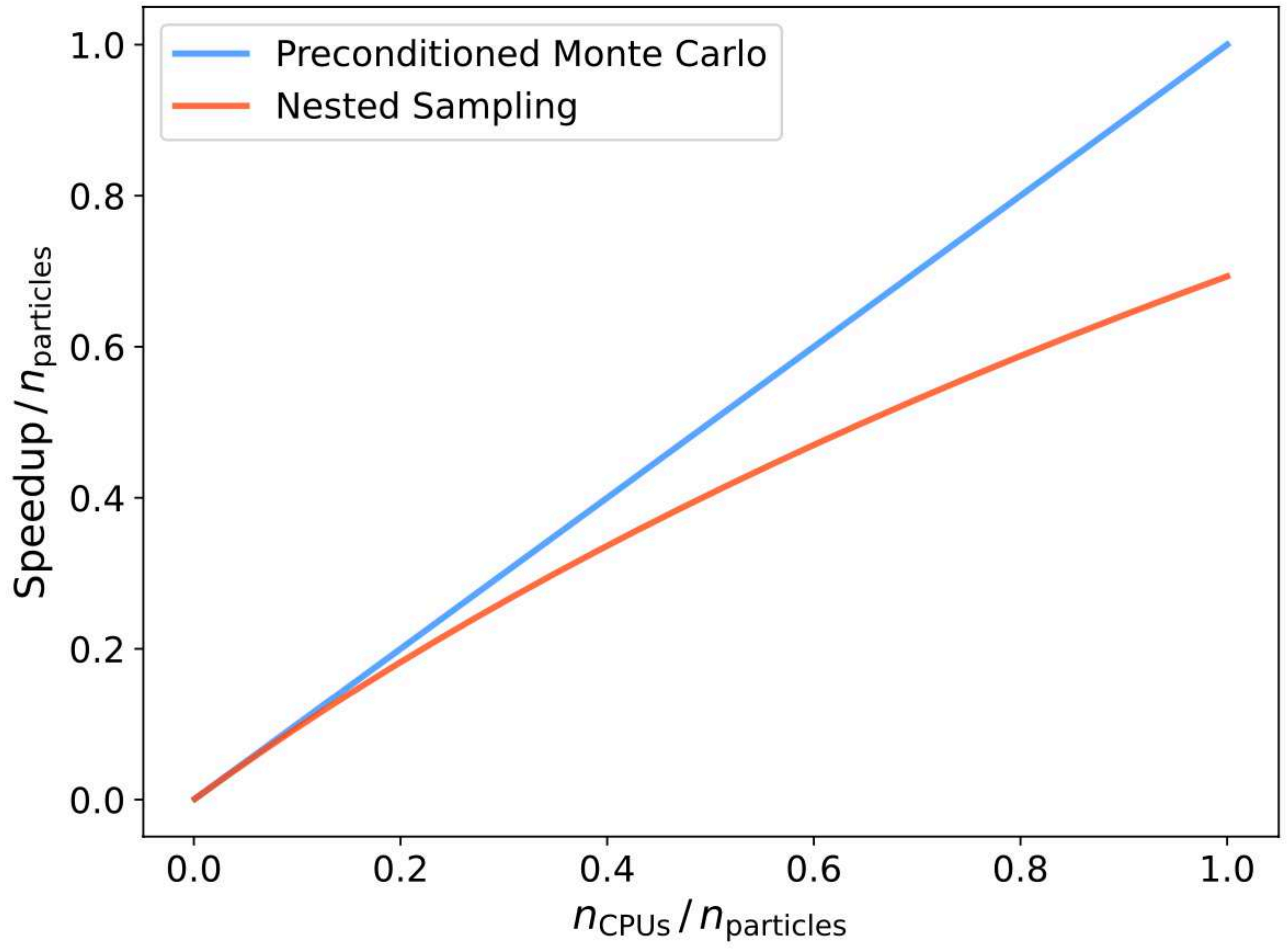
where β_t a parameter (inverse temperature).

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.

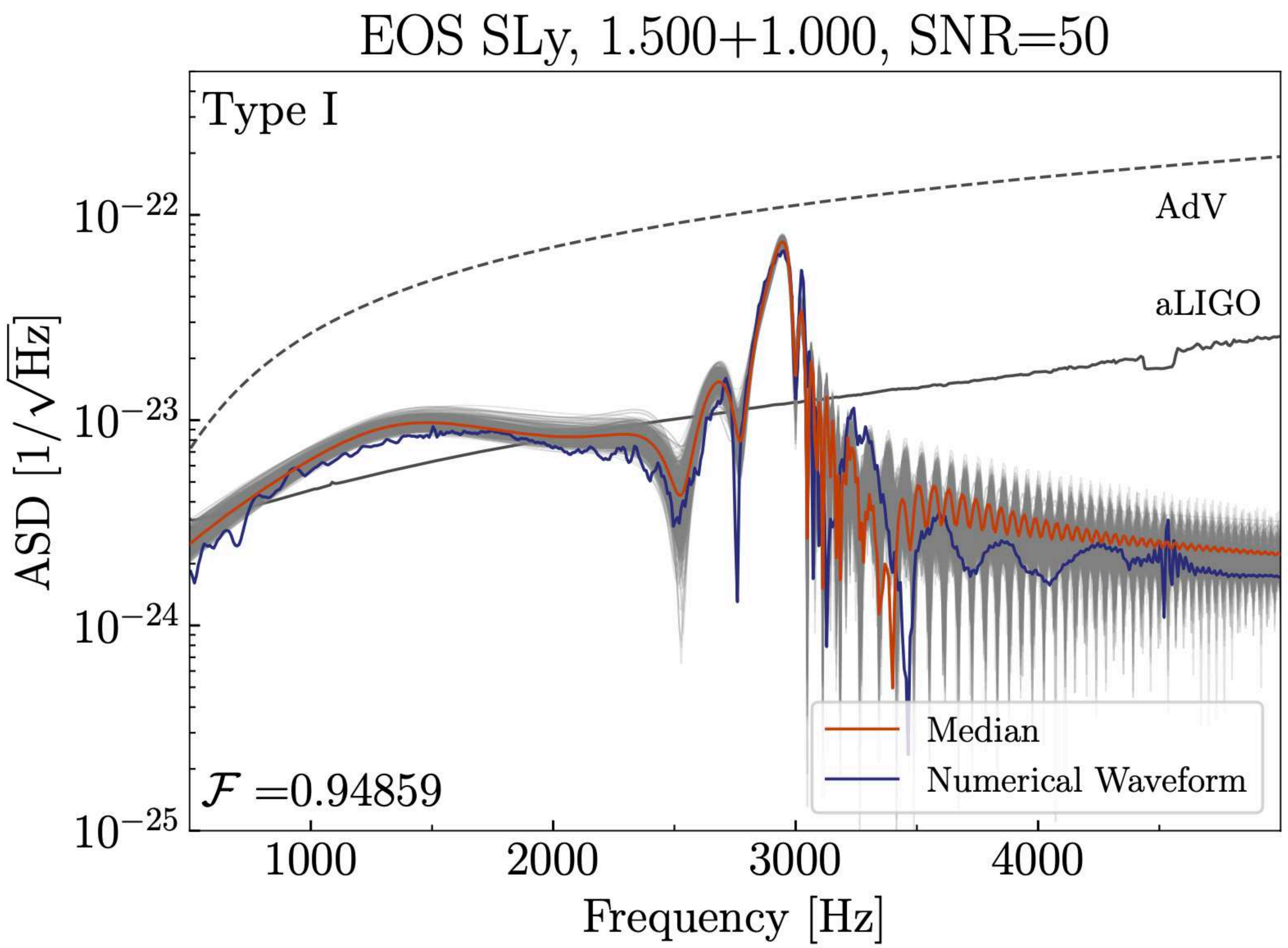
On same number of CPUs: pocoMC is ~**10 times faster** than dynesty.

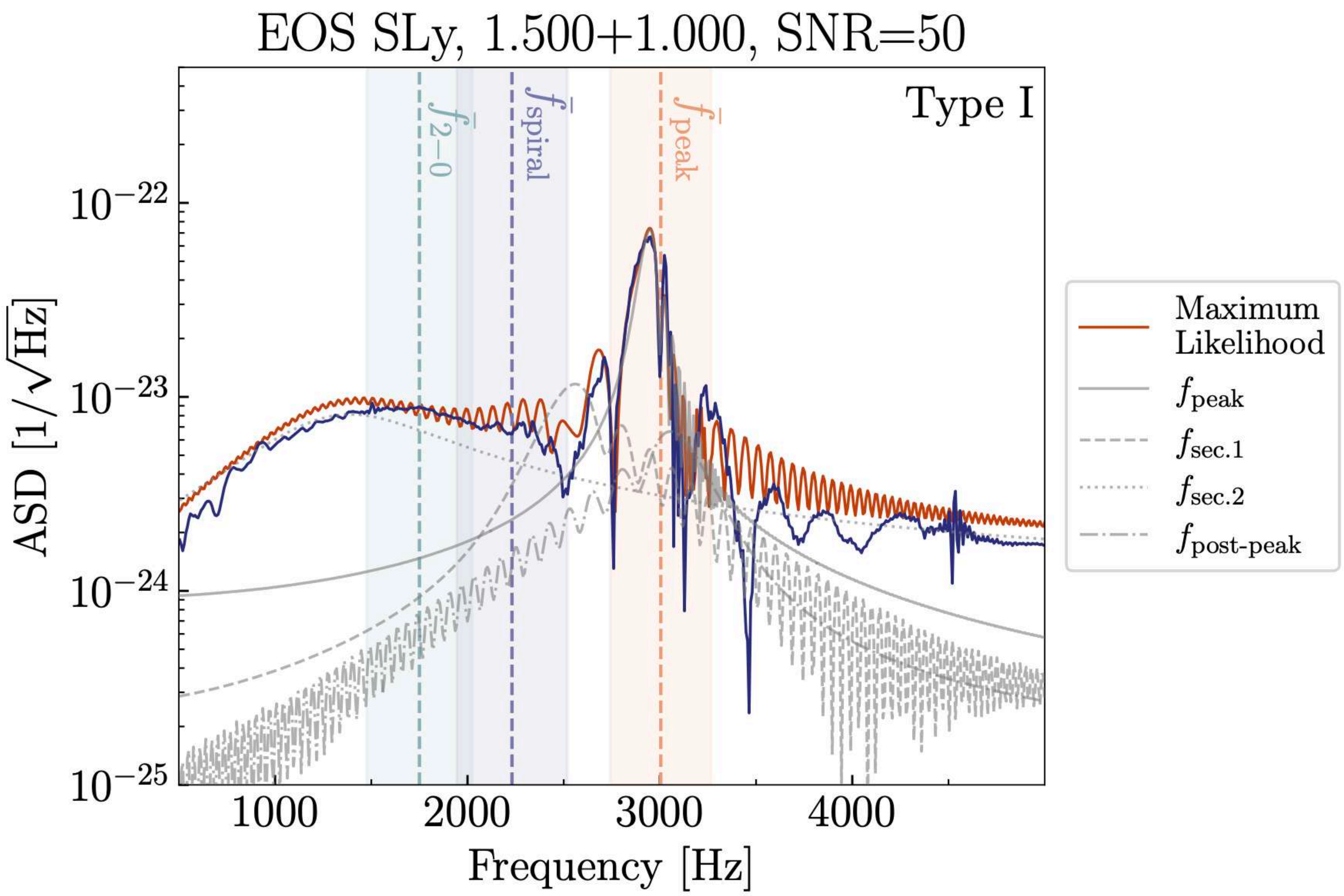
https://github.com/minaskar/pocomc



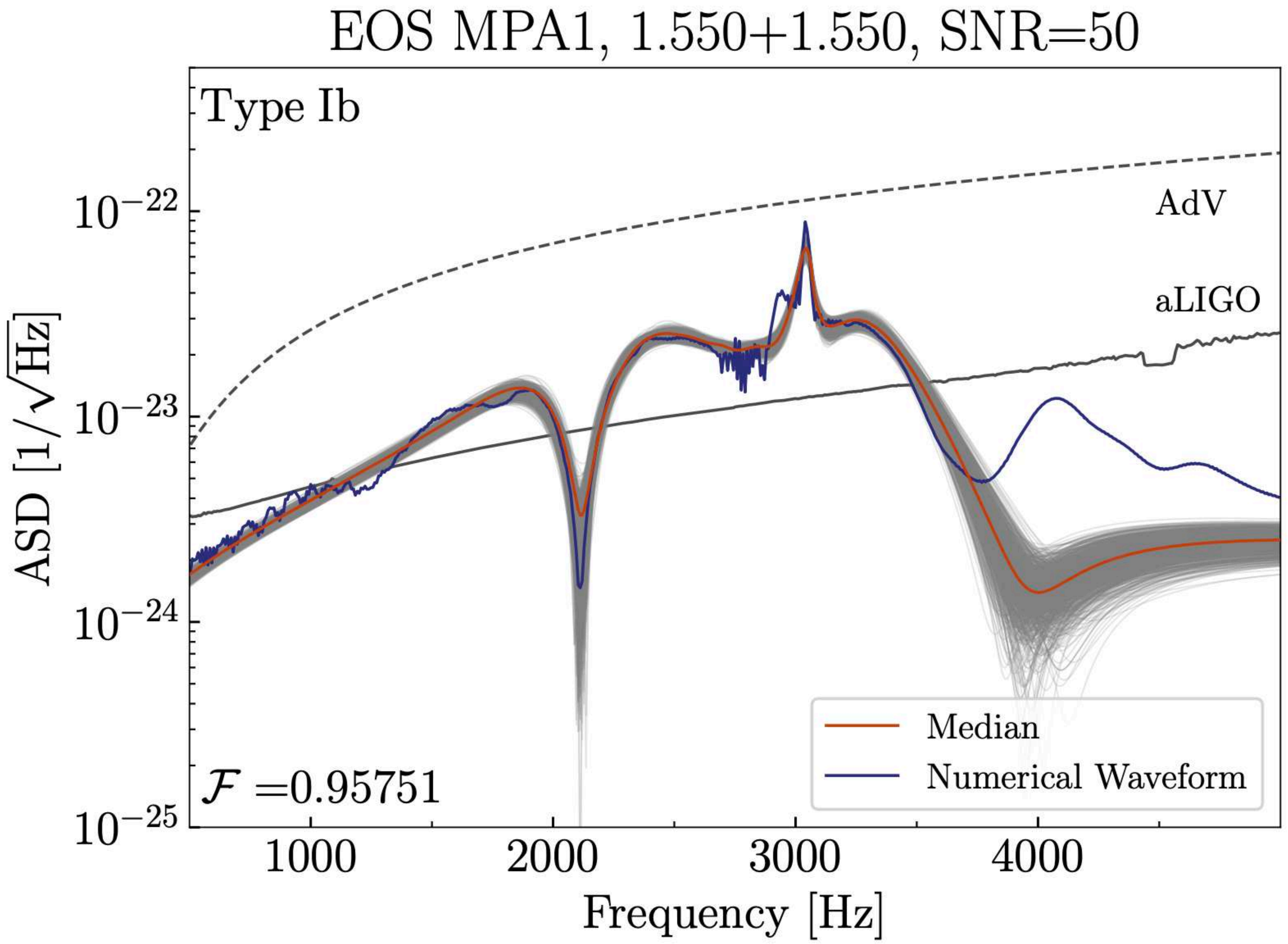
PocoMC is highly parallelizable

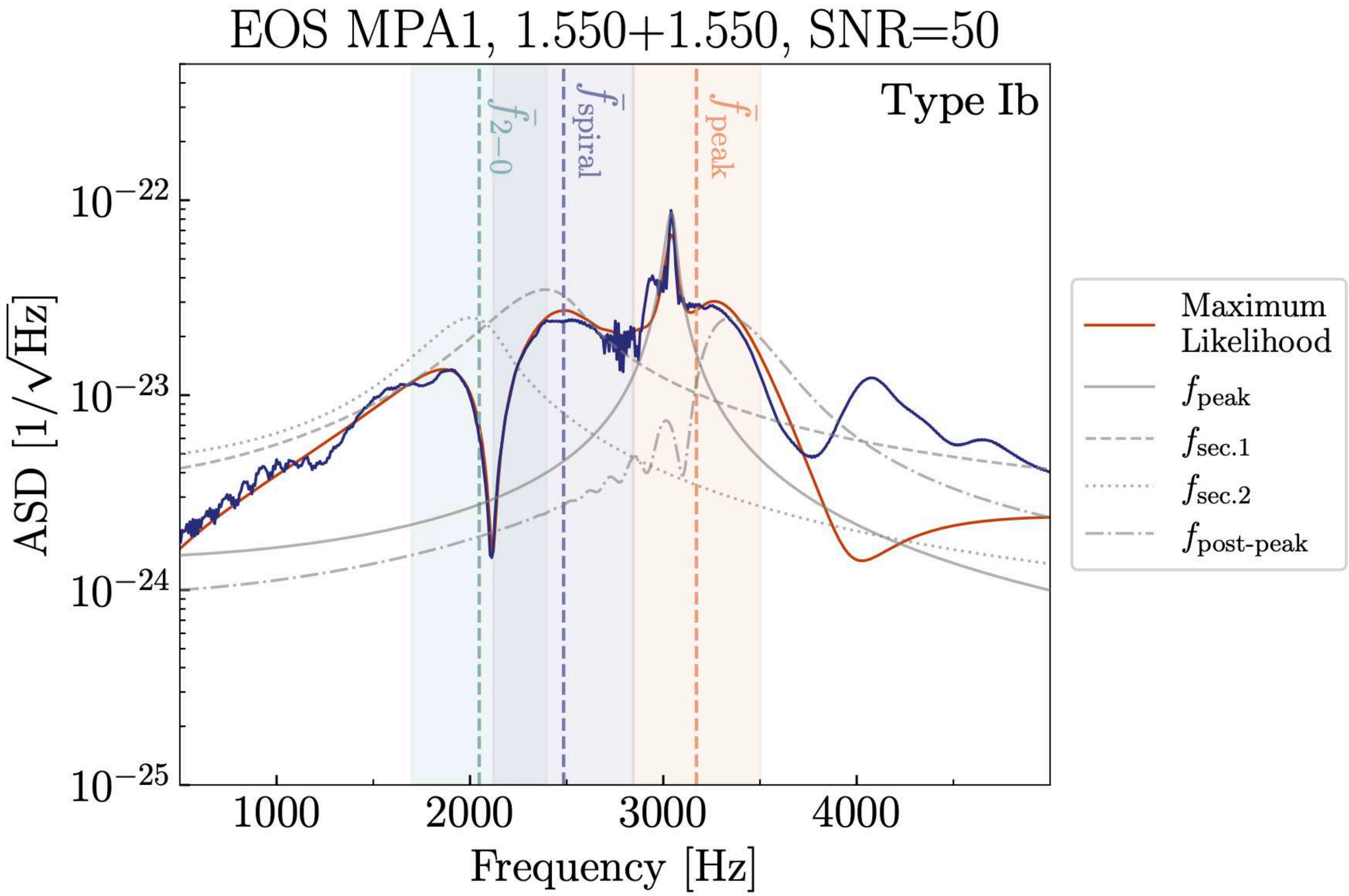
RECONSTRUCTION IN THE FREQUENCY DOMAIN



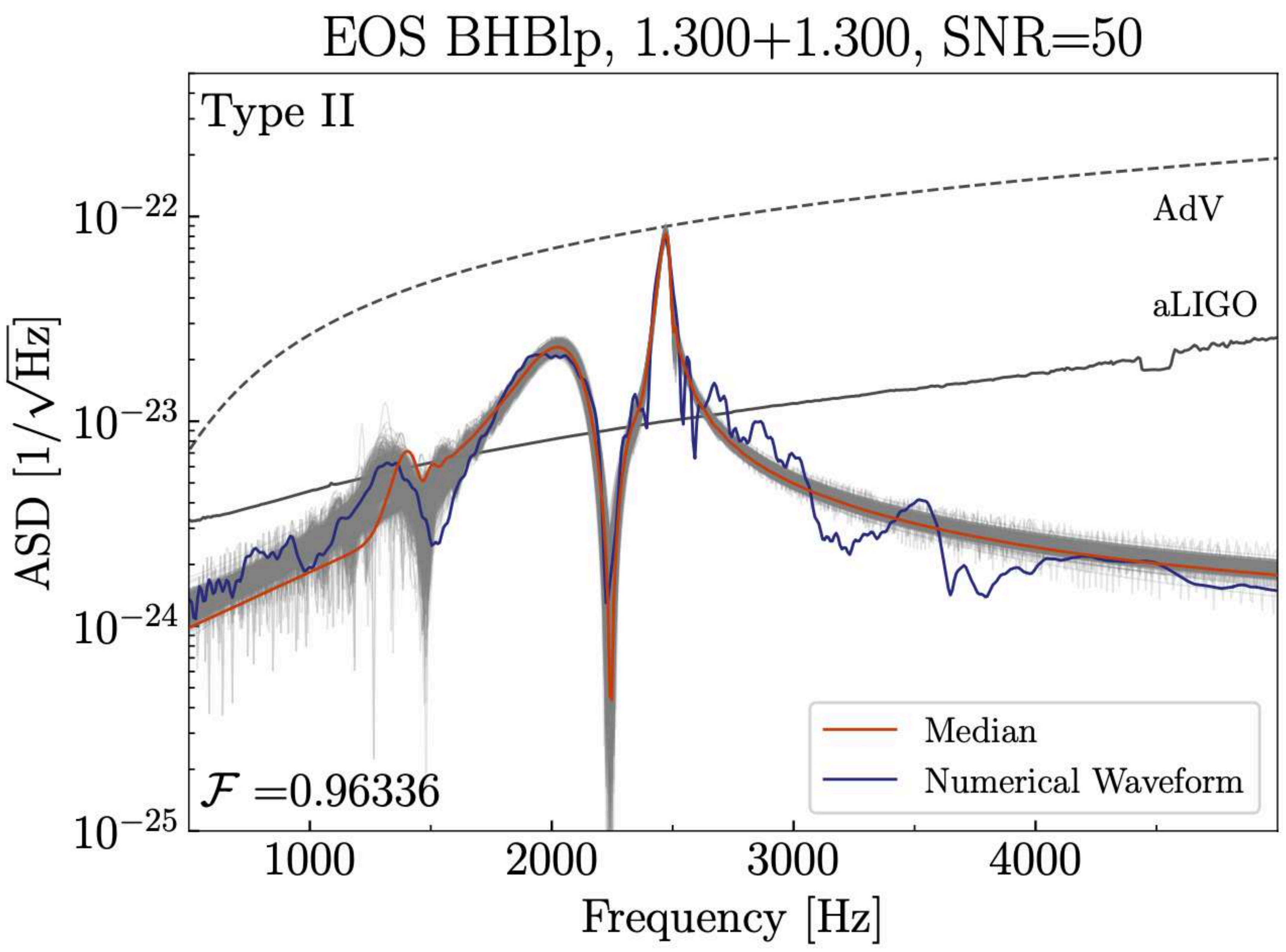


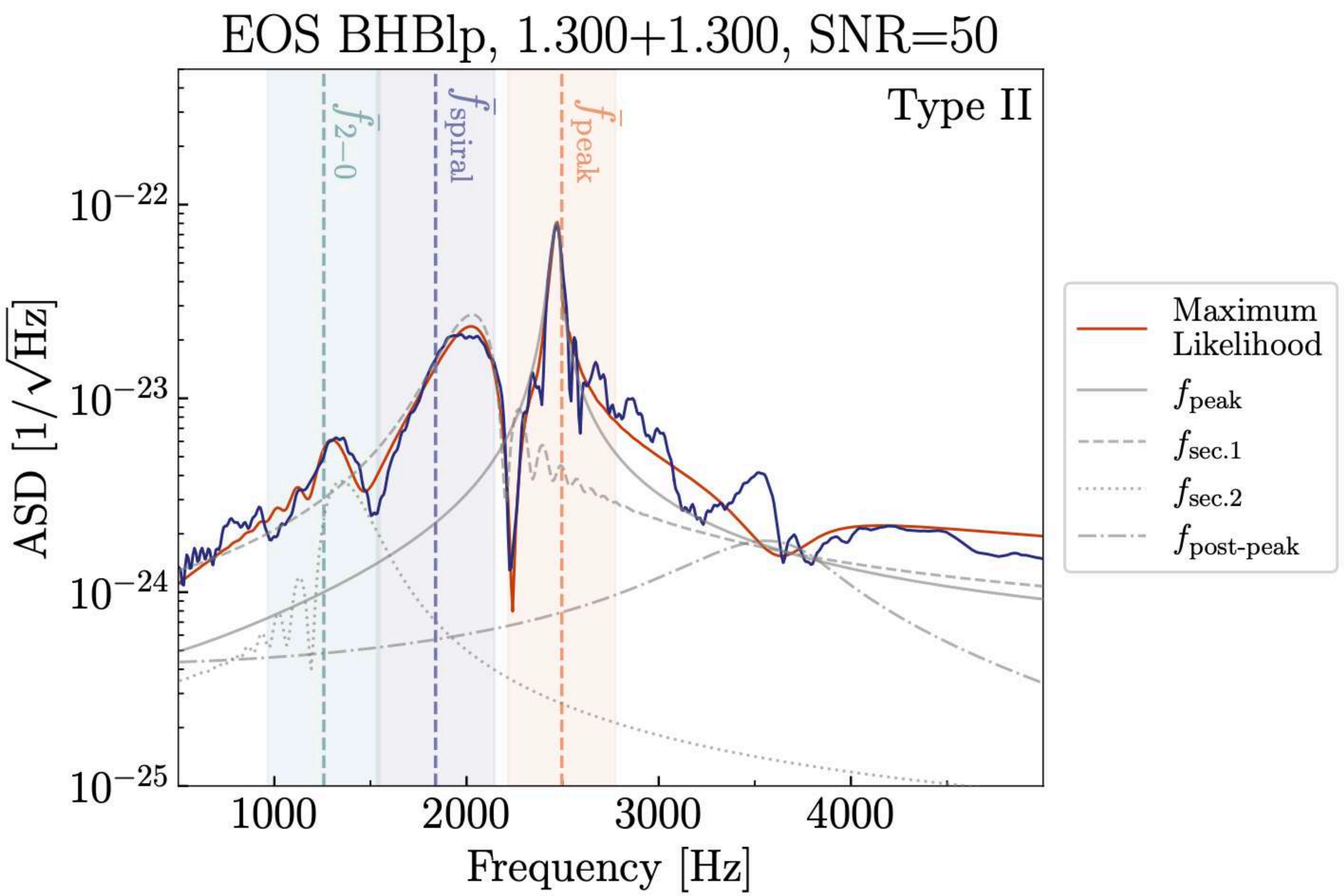
RECONSTRUCTION IN THE FREQUENCY DOMAIN



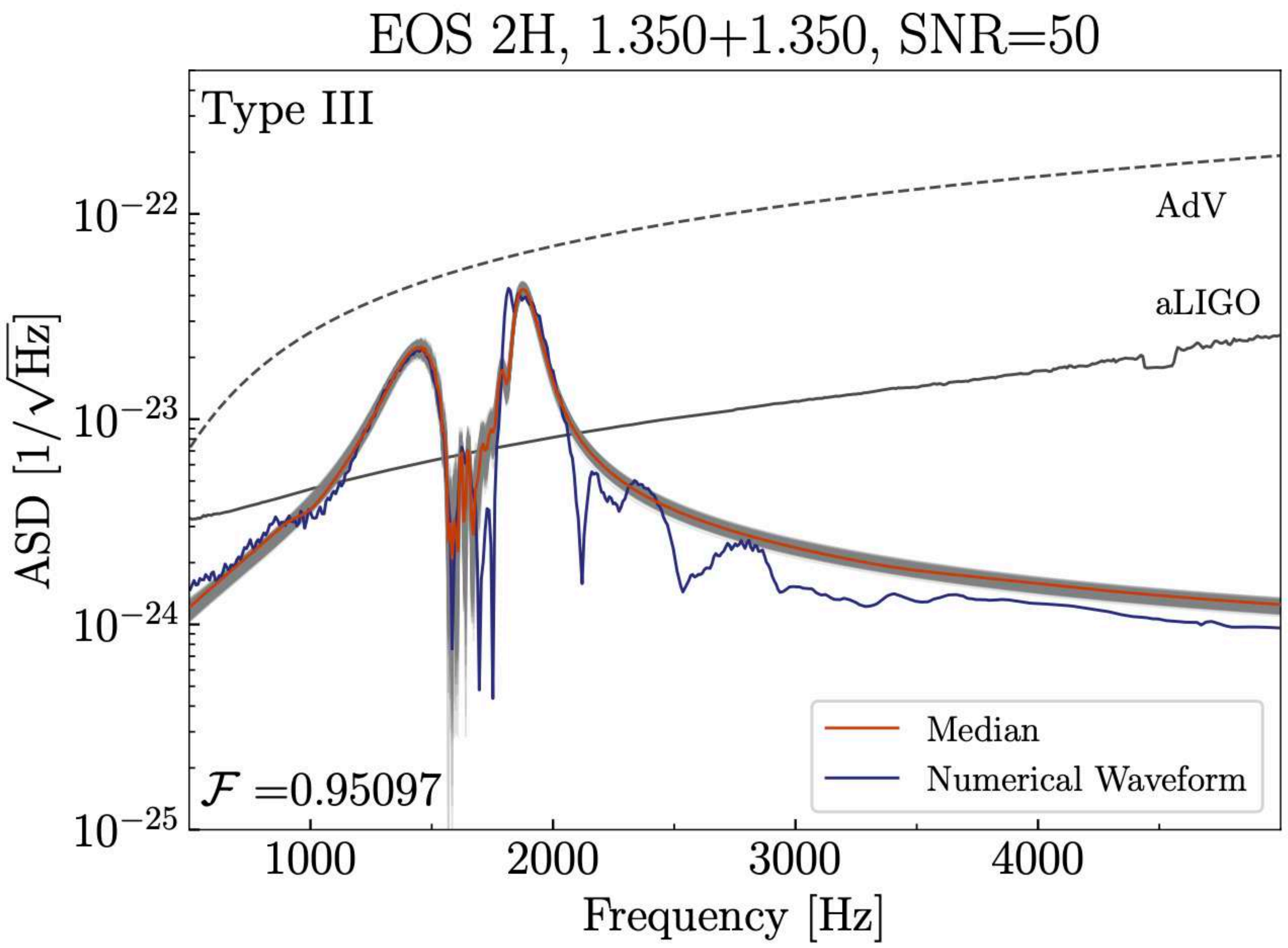


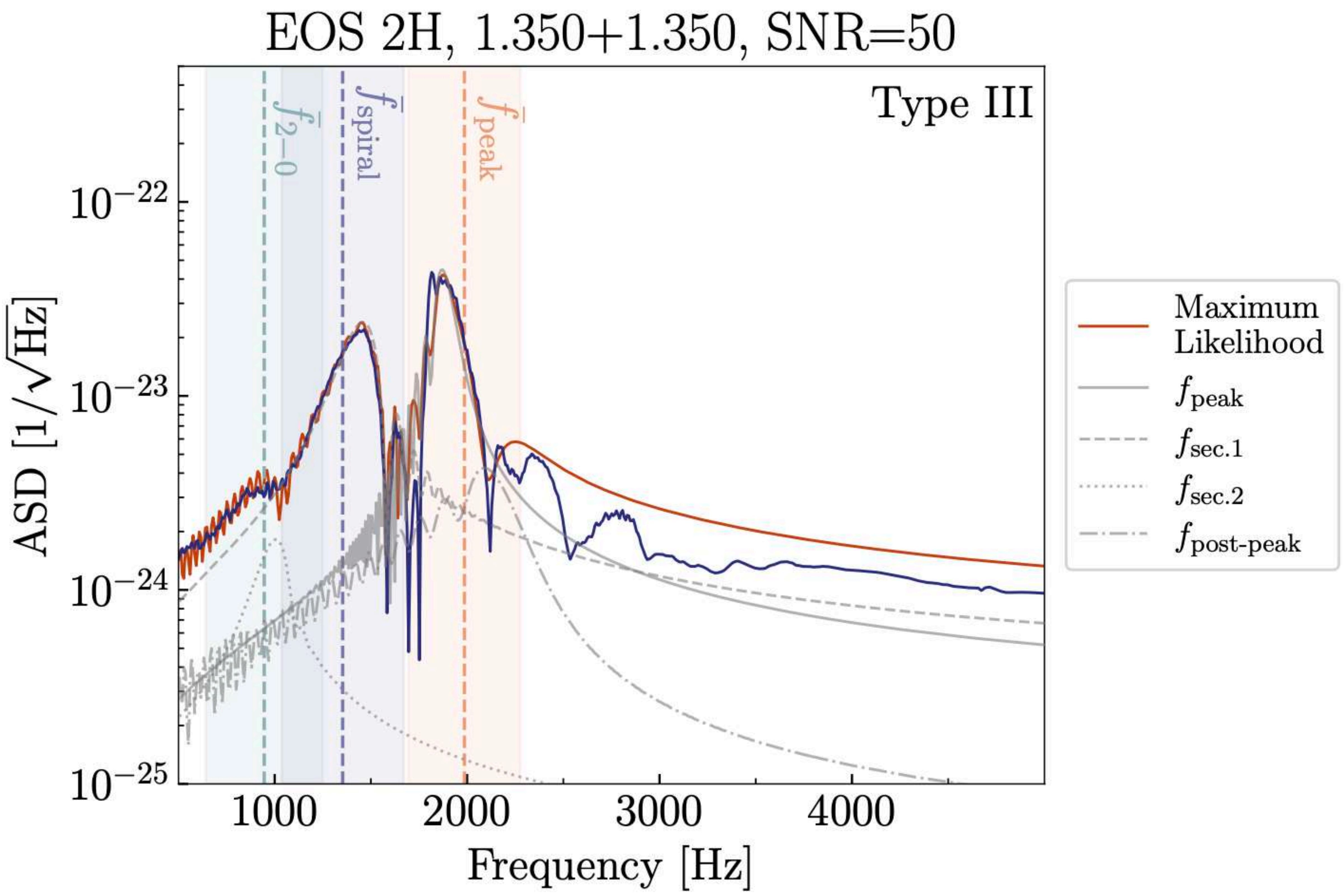
RECONSTRUCTION IN THE FREQUENCY DOMAIN



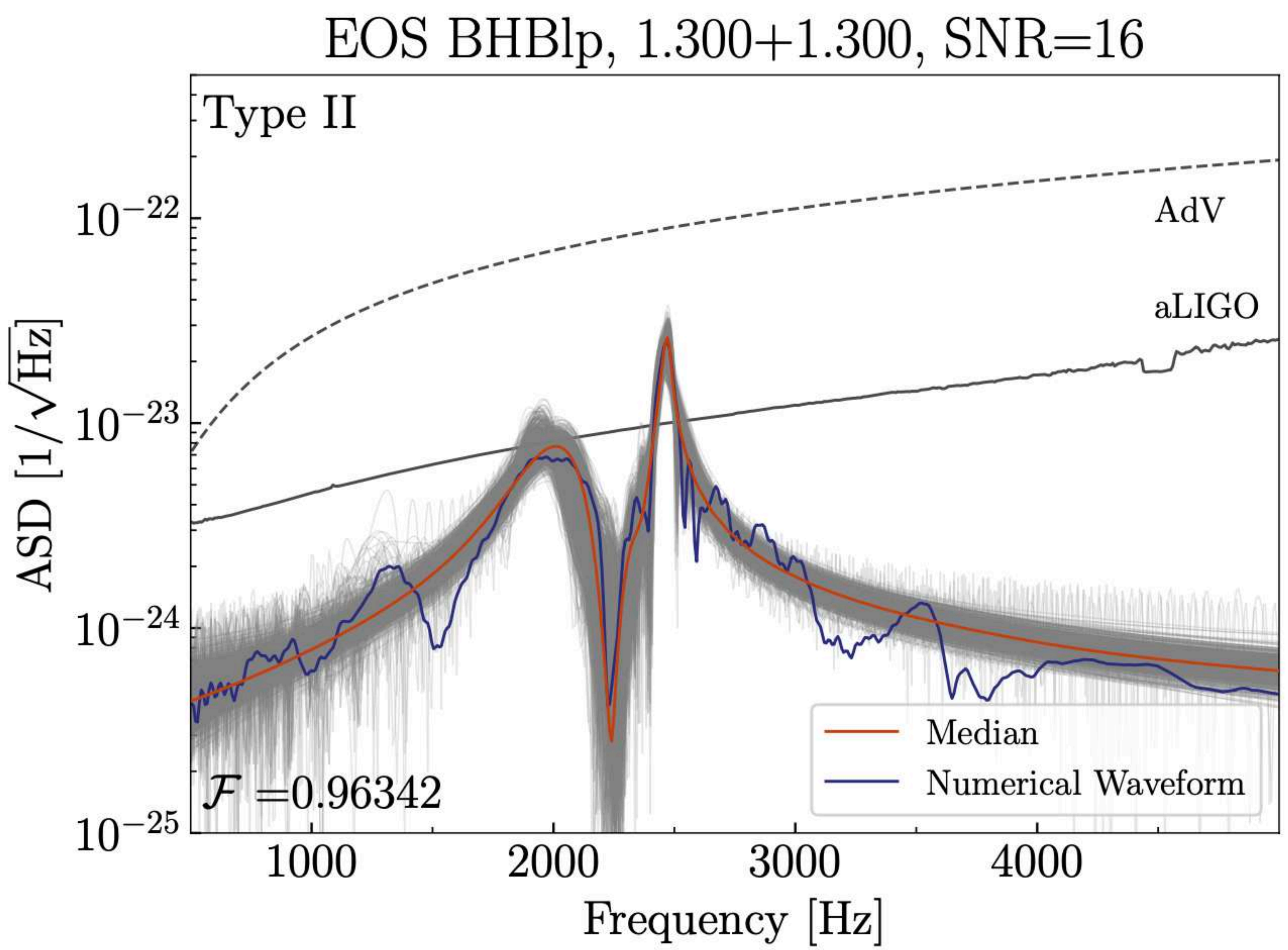


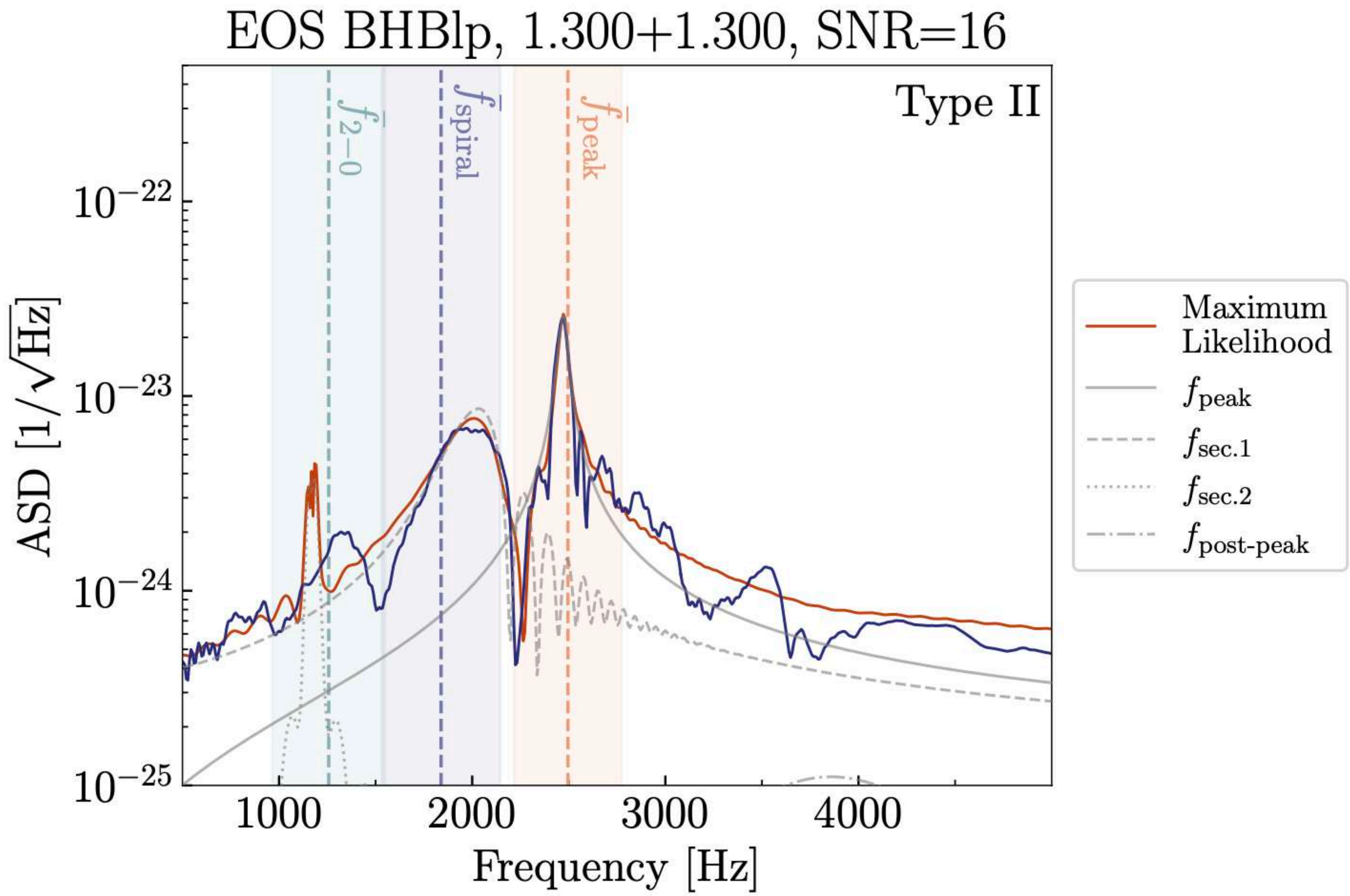
RECONSTRUCTION IN THE FREQUENCY DOMAIN



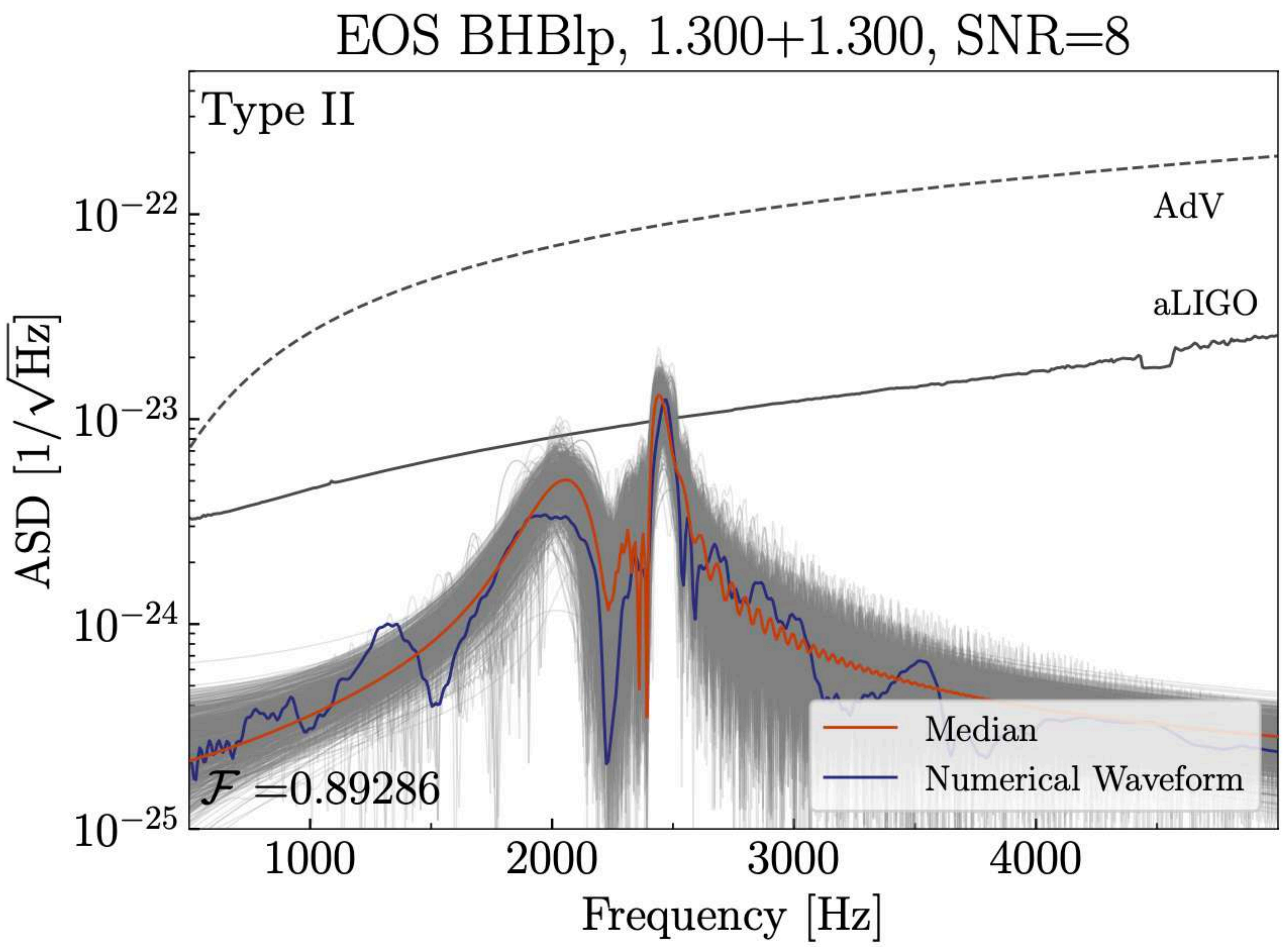


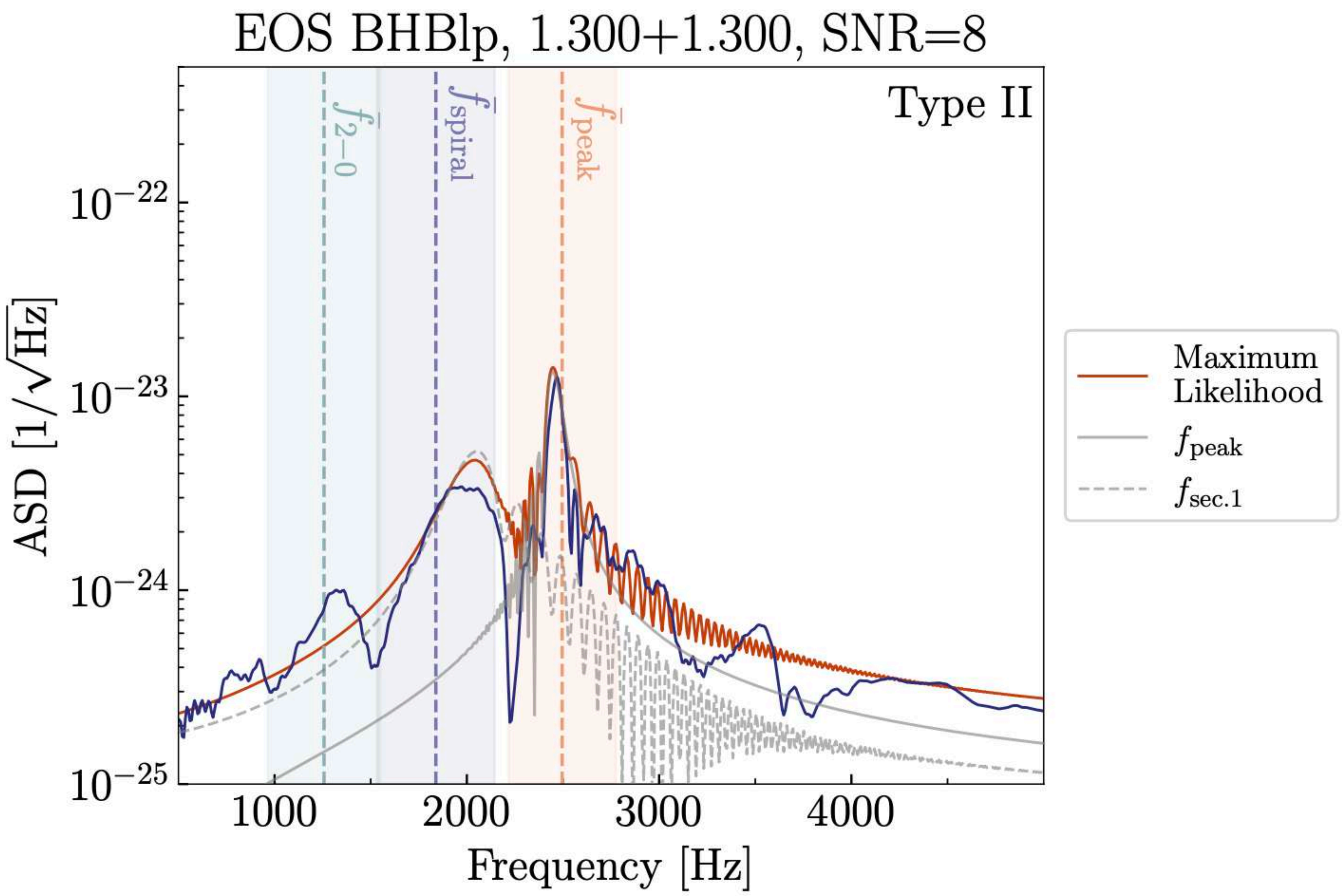
RECONSTRUCTION IN THE FREQUENCY DOMAIN





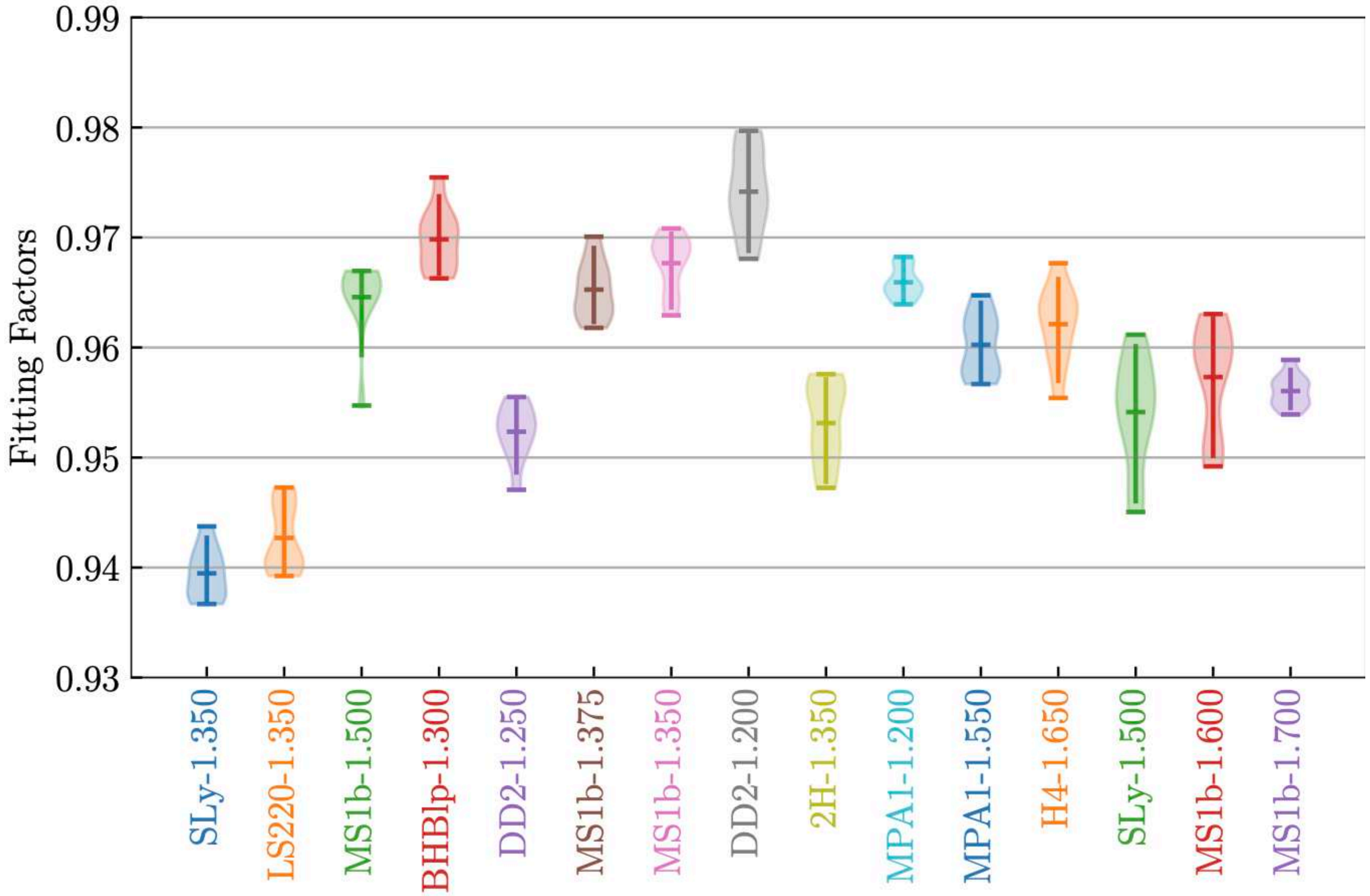
RECONSTRUCTION IN THE FREQUENCY DOMAIN





FITTING FACTORS FOR SNR=50

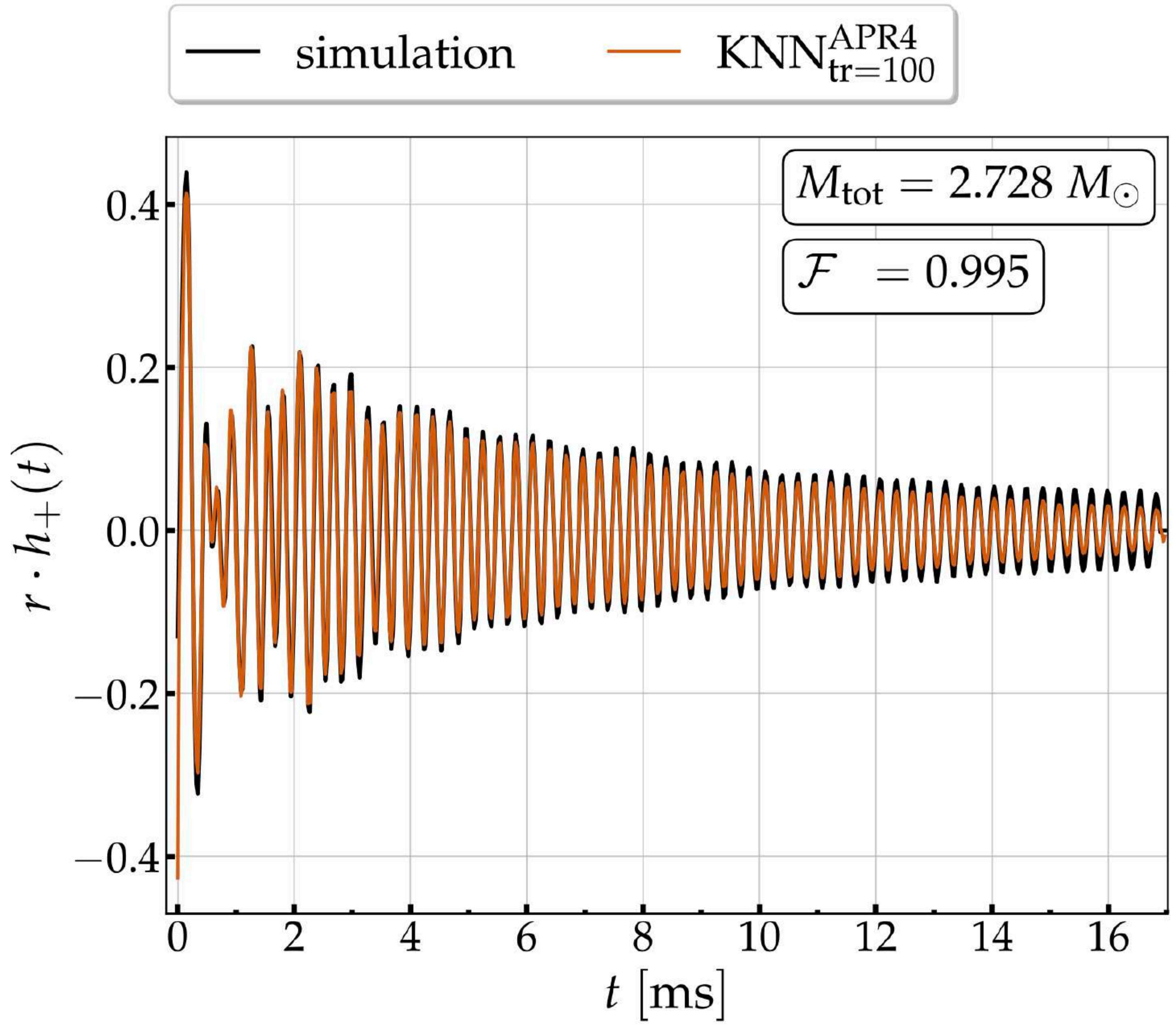
The fitting factors achieved are consistent with those in Easter et al. (2020) for a smaller set of EOS.



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:



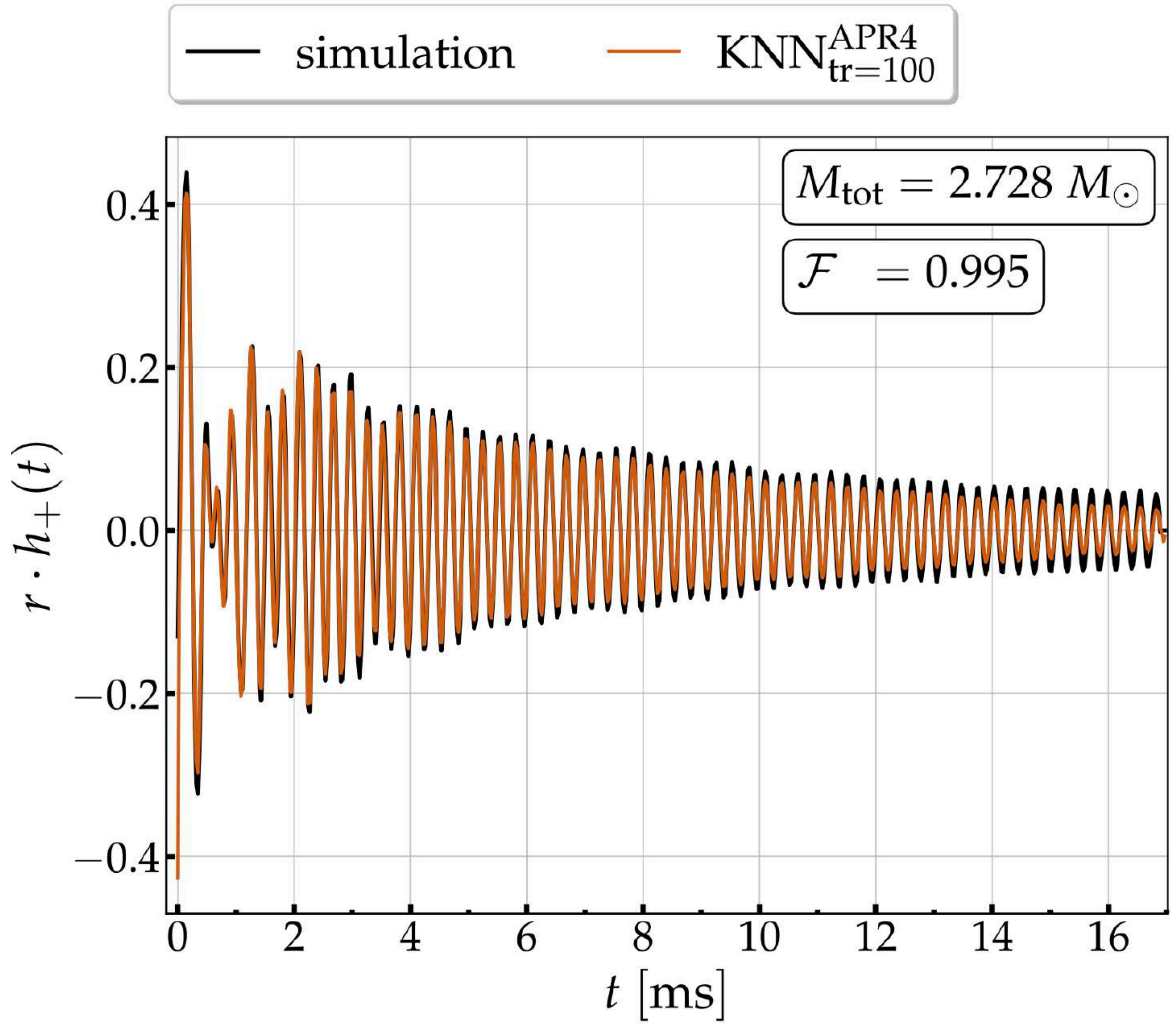
Soultanis et al. (2025)

APPLICATION OF MACHINE LEARNING TO THE POST-MERGER PHASE

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

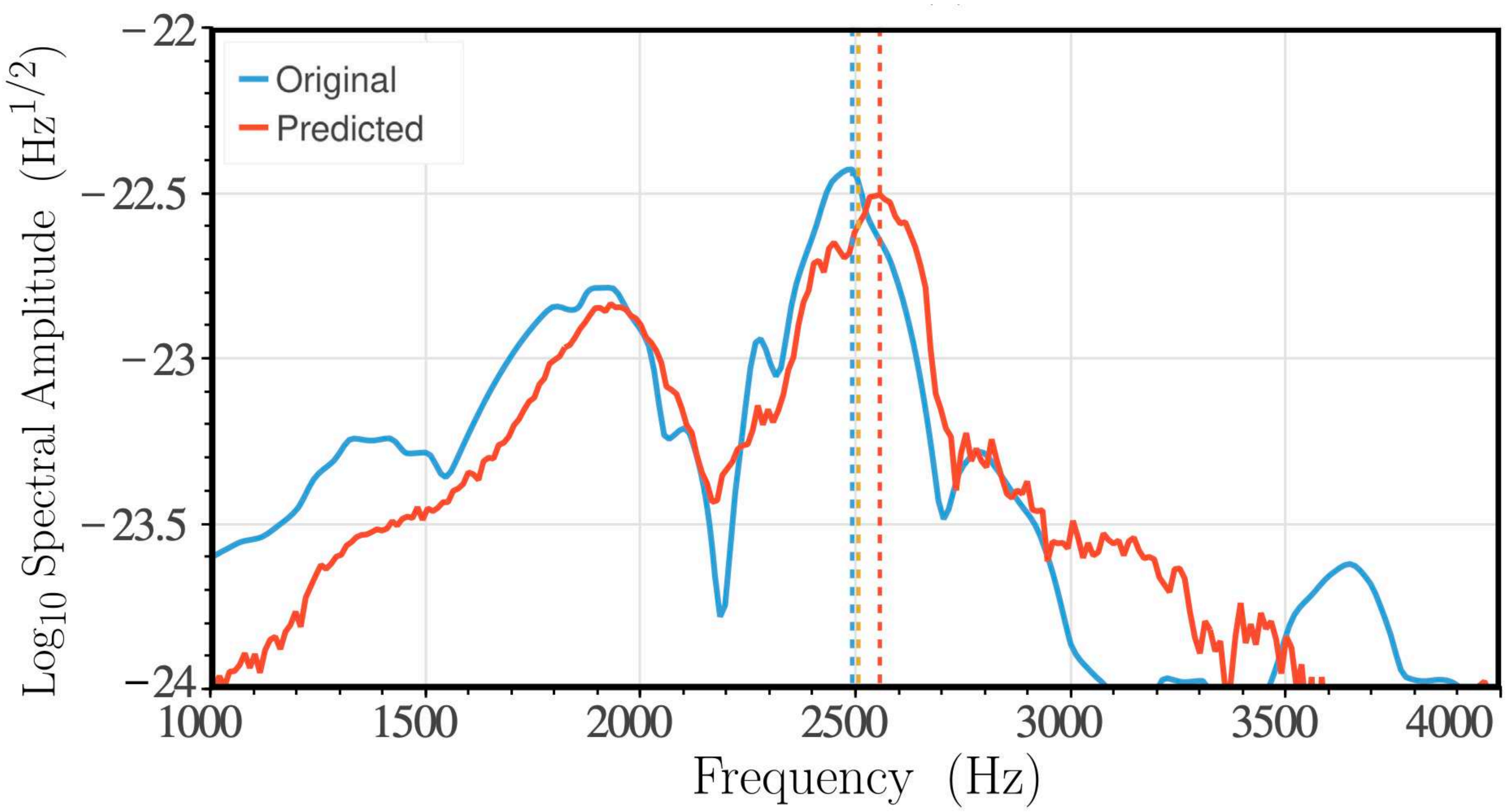
Solution: Construct surrogate model of post-merger GWs as function of e.g. M, q, EOS(Λ)

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)

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}.

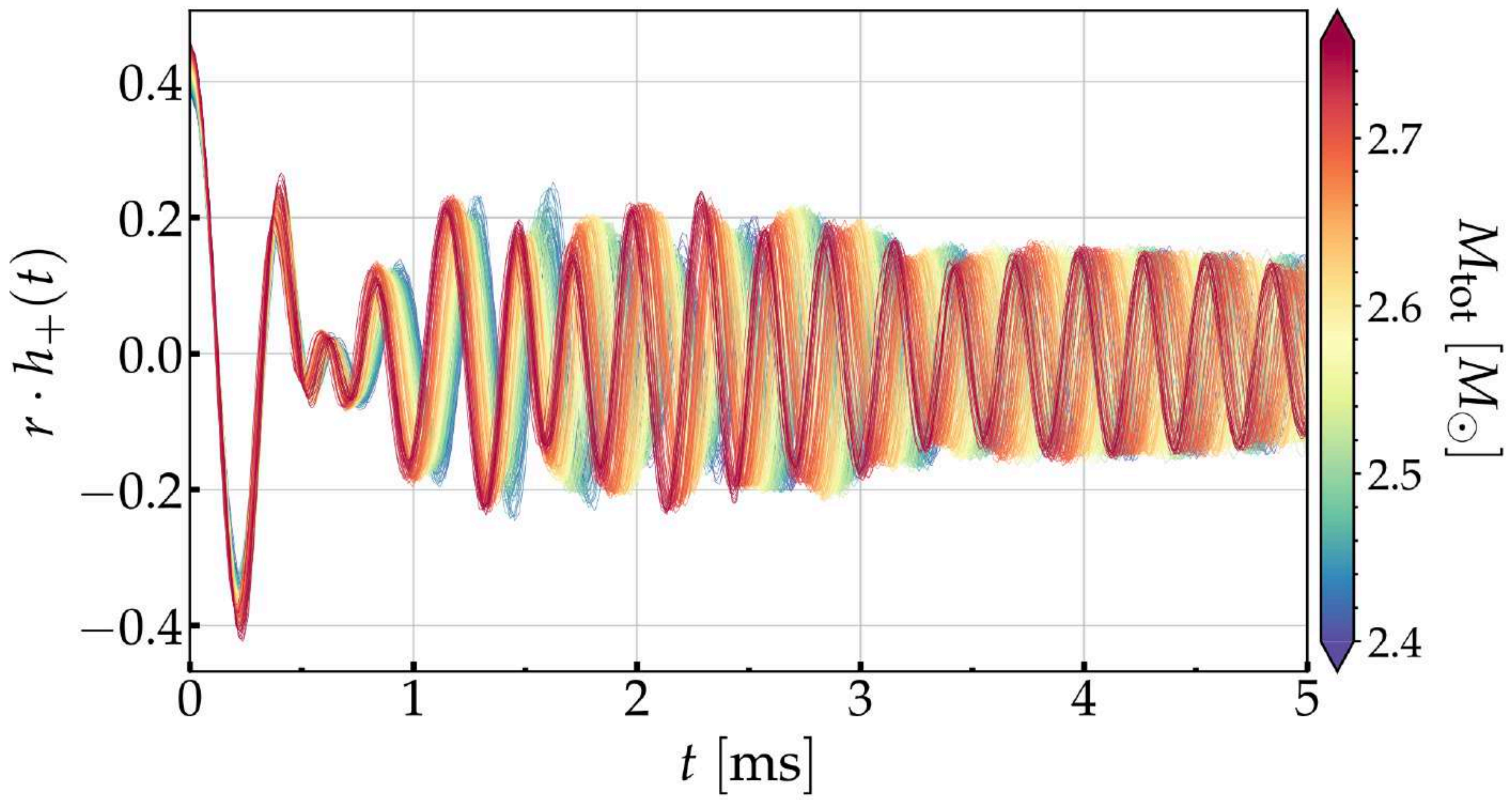
Choose specific EOS (APR4; SFHX).

Complex GW strain:

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

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

Signals are aligned at merger time:



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}.

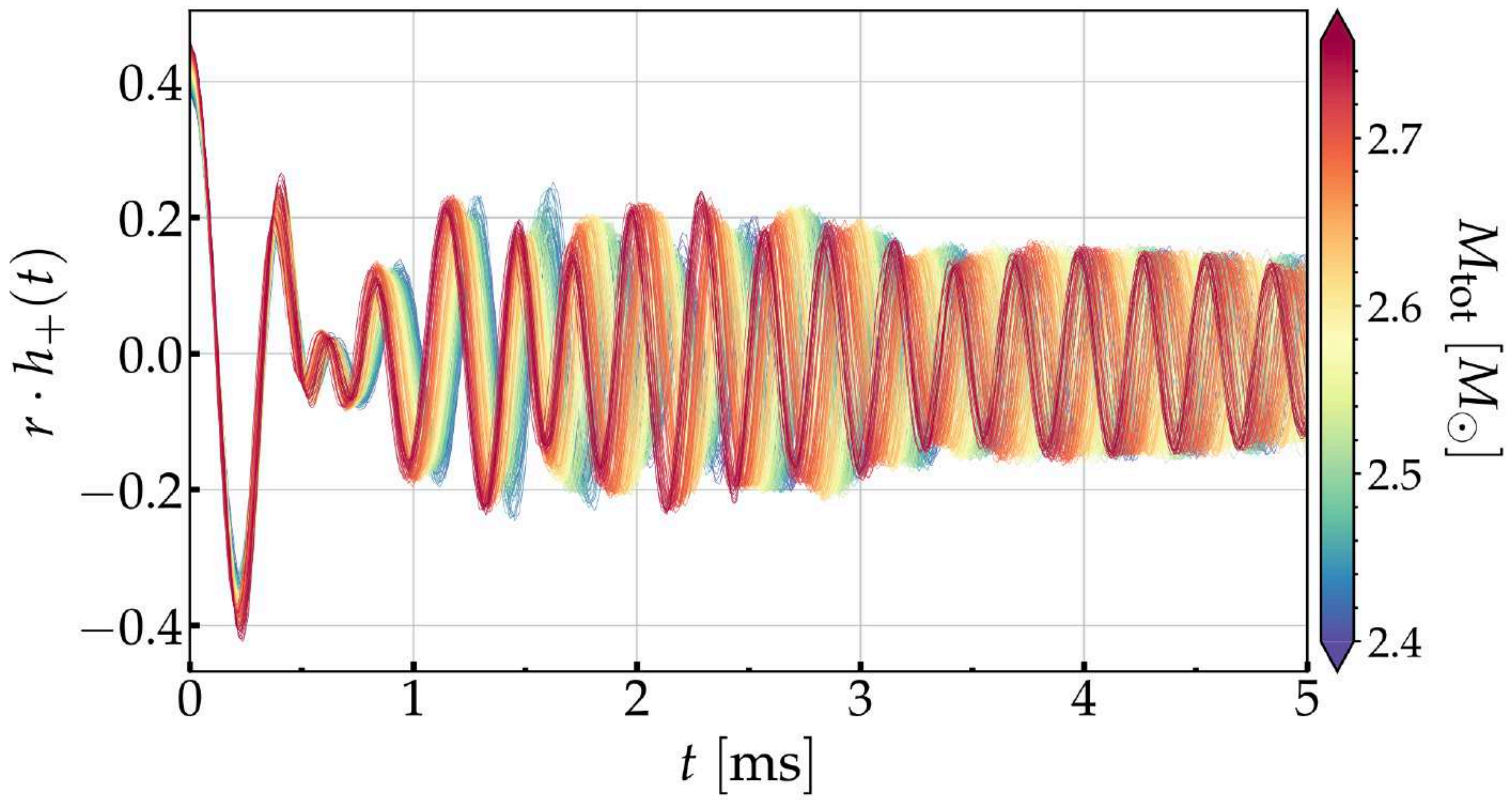
Choose specific EOS (APR4; SFHX).

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)

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:

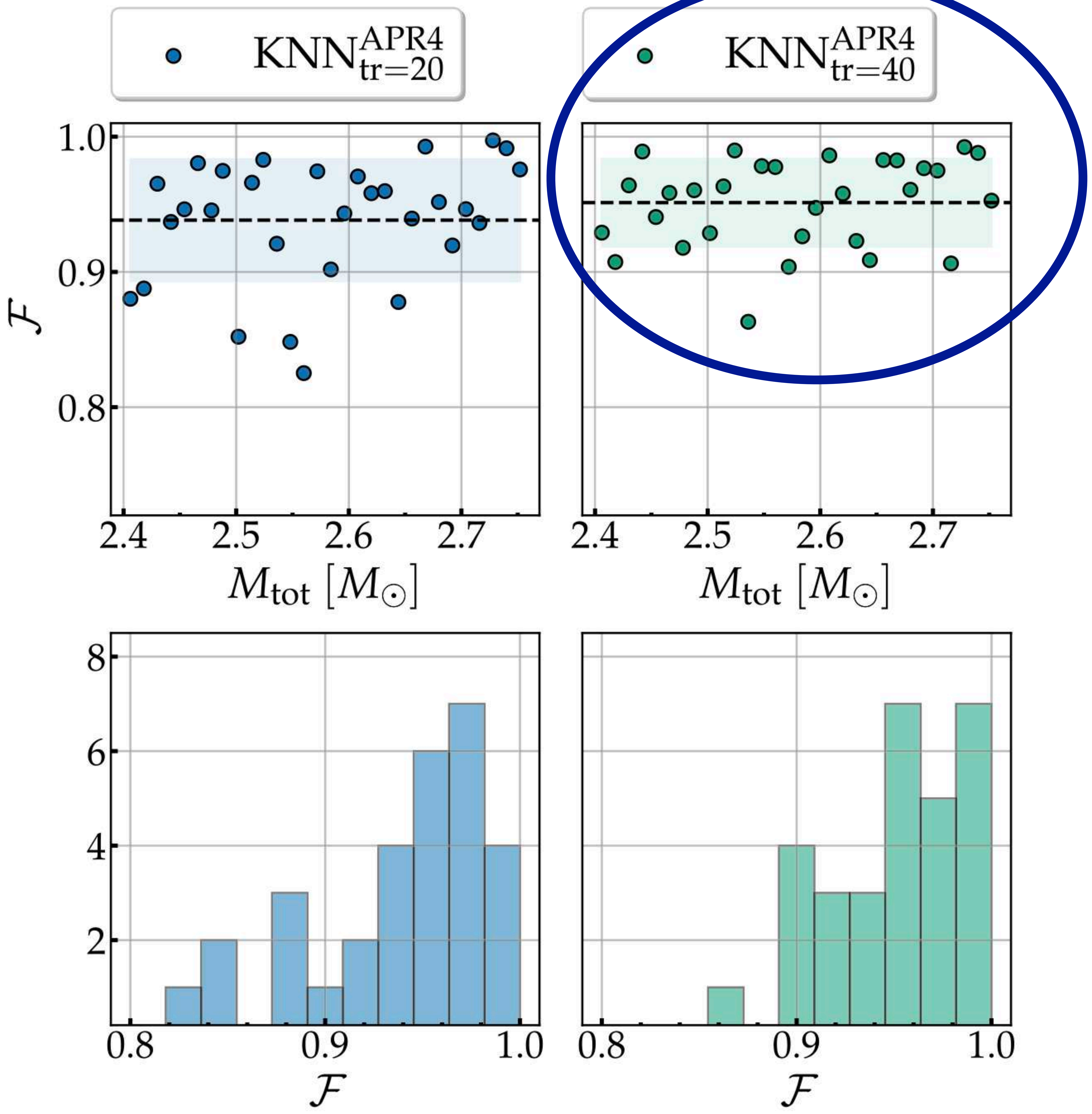
$$\langle h_1(t), h_2(t)\rangle \equiv 4\,\mathrm{Re}\int_0^\infty df\,\frac{\tilde{h}_1(f)\cdot\tilde{h}_2^*(f)}{S_h(f)}$$

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)

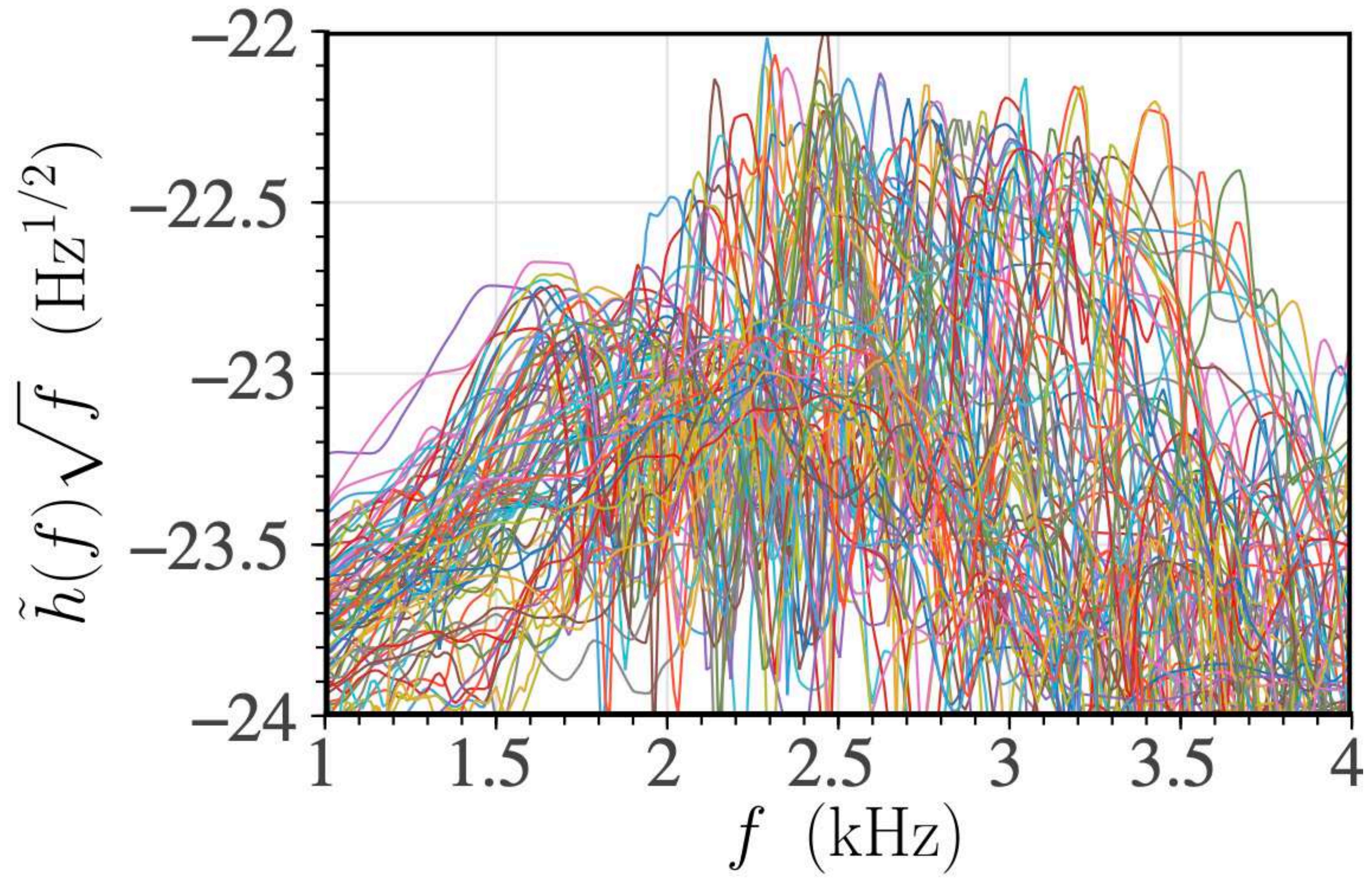
$$\mathcal{F} \equiv \max_{\phi_0, t_0}\frac{\langle h_1(t), h_2(t)\rangle}{\sqrt{\langle h_1(t), h_1(t)\rangle\,\langle h_2(t), h_2(t)\rangle}}$$



Soultanis et al. (2025)

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

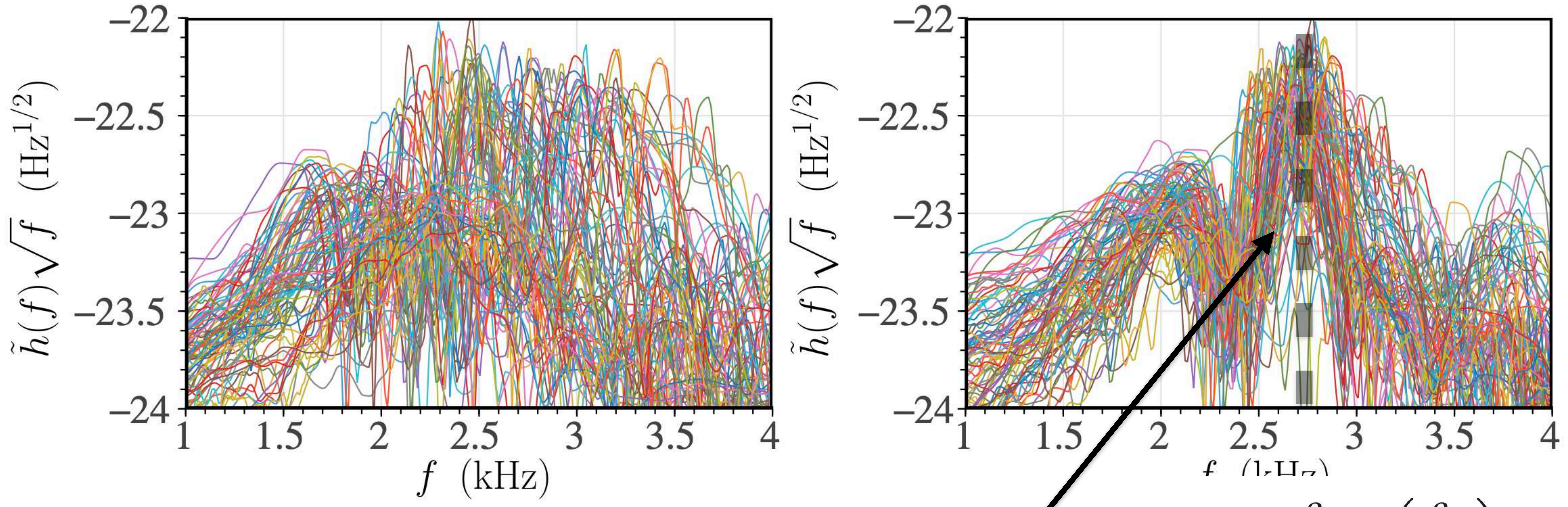
Surrogate model now depends on both mass and tidal deformability.



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: $f_{\text{peak}}(\kappa_2^\tau, M) = 4\dfrac{\beta_1}{M}\ln\left(\dfrac{\beta_0}{8\kappa_2^\tau}\right)$

Pesios, Koutalios, Kugiumtzis, Stergioulas (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).

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)

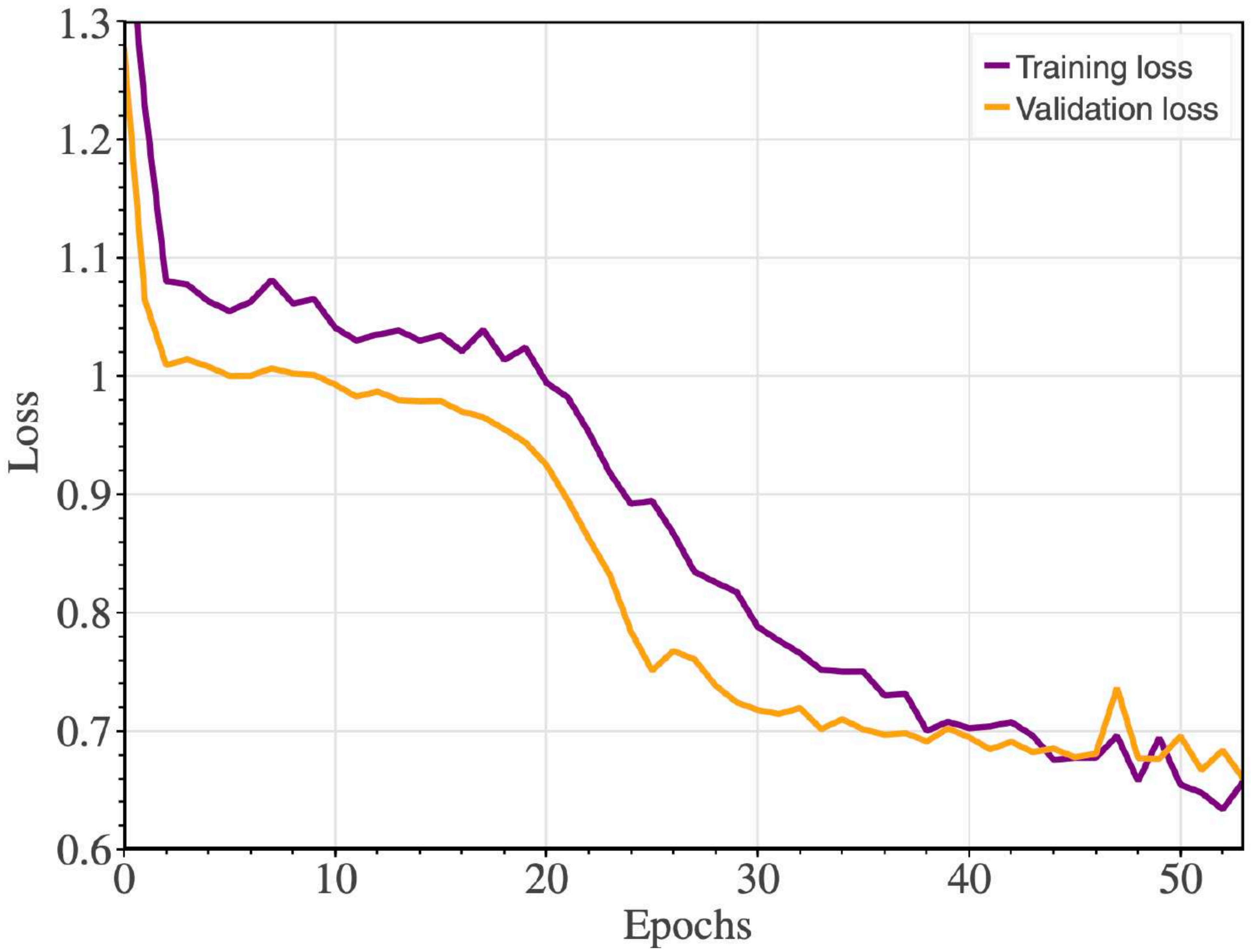
ANN SURROGATE IN THE FREQUENCY DOMAIN

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)

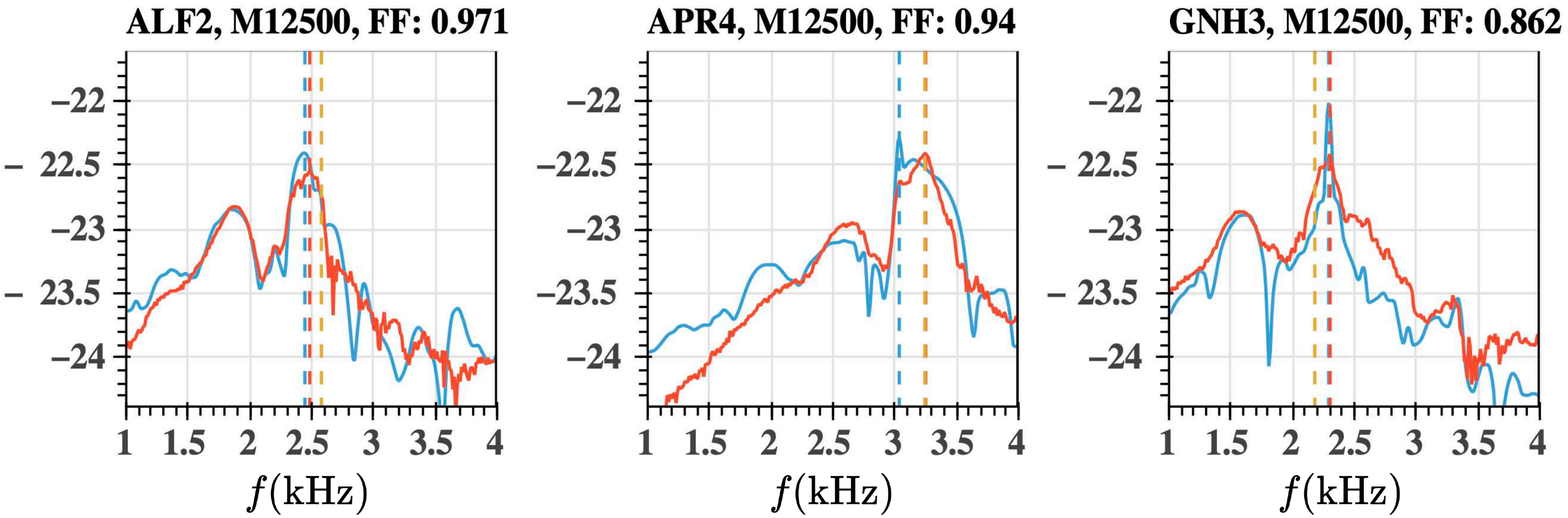
ANN SURROGATE IN THE FREQUENCY DOMAIN

Typical examples of predicted magnitudes of GW spectra

FF = O = overlap

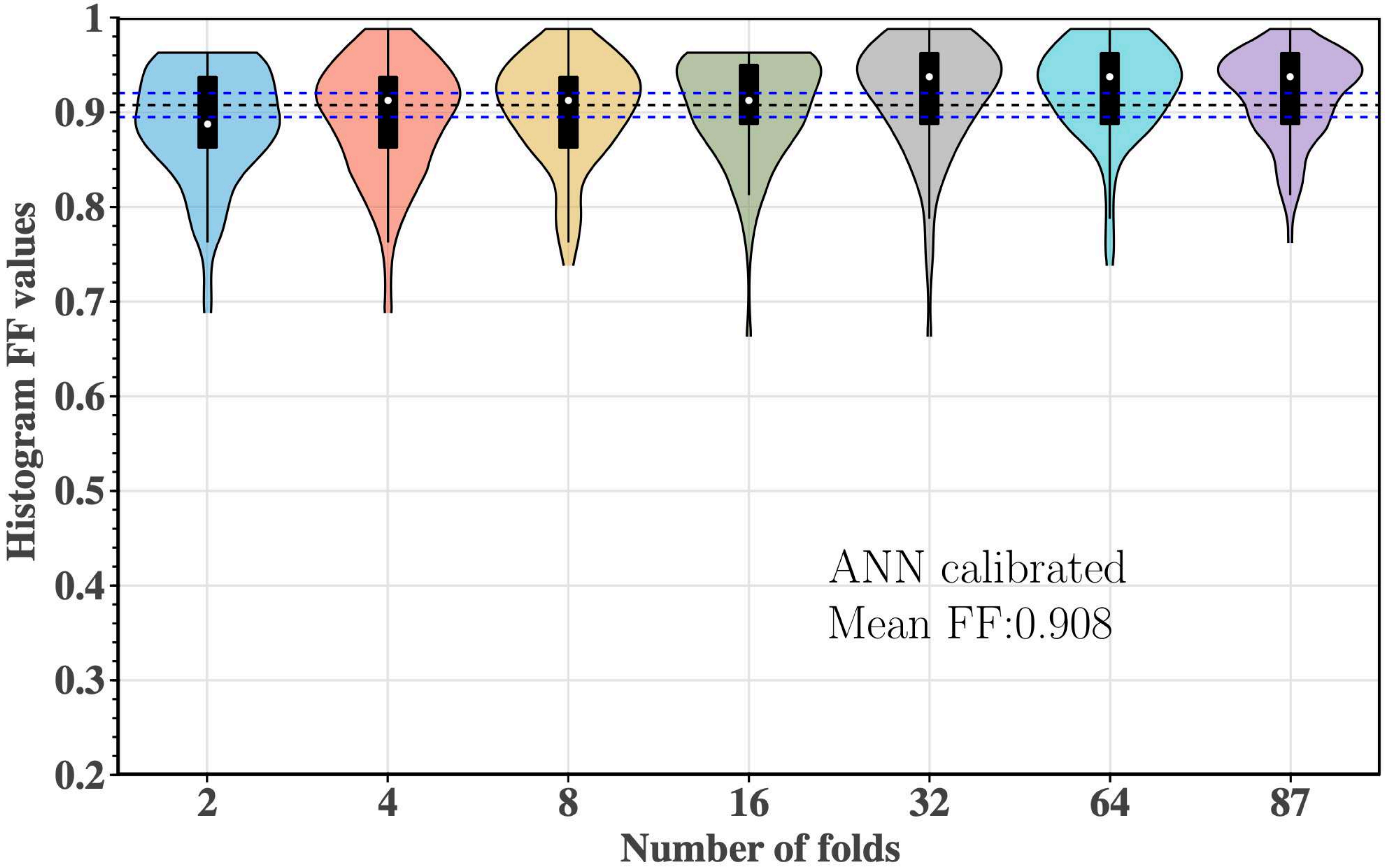
Impact of uncertainty in empirical relation offset by re-calibration of spectra.

ANN outperforms multivariate linear regression.



Pesios et al. (2024)

Cross-validation study of fitting factors distribution:



THANK YOU FOR YOUR ATTENTION

SEARCH IN O3 DATA WITH ARESGW

Total number of BBH events previously found by GWTC+OGC+IAS+pycbc_KDE inside training range = **43**

We also find **8 new candidate events** p_astro > 0.5.

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

TABLE III: Performance of AresGW in comparison to the GWTC, OGC, IAS and pycbc_KDE [15] algorithms on all 51 candidate events (43 previously published plus 8 new events detected by AresGW) that are within the effective training range.

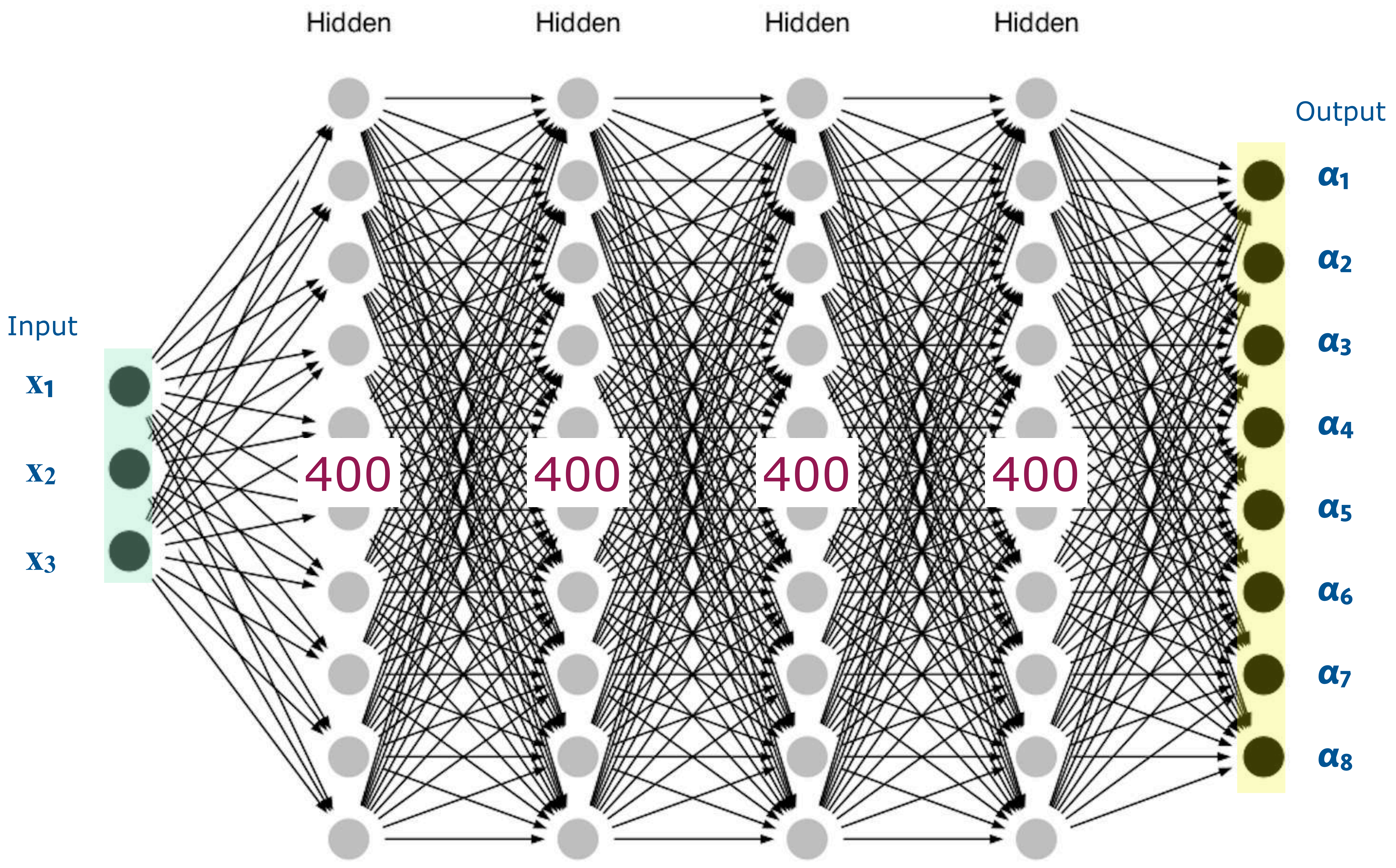
cwb	pycbc_broad	mbta	gstlal	pycbc_KDE	pycbc_bbh	IAS_a	IAS_b	OGC	AresGW
7	20	27	27	29	31	34	36	37	42

PARAMETERS OF NEW GW DETECTIONS WITH AresGW

TABLE VIII: Parameter estimation for the new AresGW candidate events.

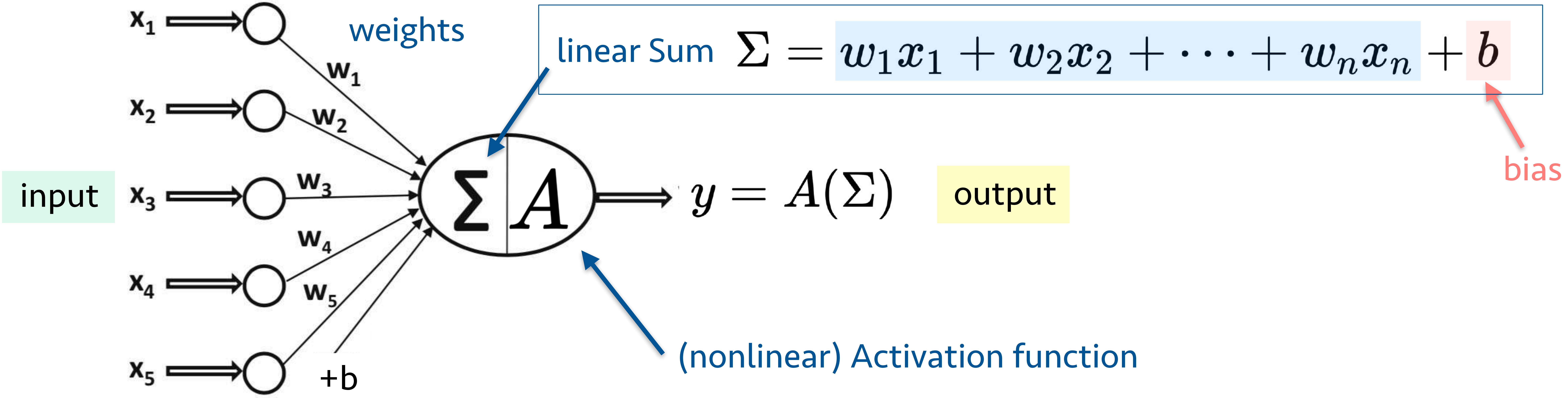
#	Event Name	$\mathcal{M}$ ($M_\odot$)	q	m_1 ($M_\odot$)	m_2 ($M_\odot$)	D_L (Mpc)	χ_{eff}	SNR (H1)	SNR (L1)	SNR $\hat{\rho}$ (network)
1	GW190511_125545	$28.95^{+9.45}_{-6.86}$	$0.72^{+0.25}_{-0.36}$	$40.7^{+16.2}_{-10.5}$	$28.2^{+11.6}_{-11.2}$	3707^{+3471}_{-2173}	$0.23^{+0.25}_{-0.29}$	2.29	7.34	7.29
2	GW190614_134749	$25.97^{+16.59}_{-6.20}$	$0.70^{+0.27}_{-0.36}$	$37.0^{+31.8}_{-10.7}$	$25.2^{+15.2}_{-9.7}$	6551^{+9562}_{-3558}	$0.05^{+0.34}_{-0.34}$	3.51	6.08	7.02
3	GW190607_083827	$30.48^{+7.21}_{-4.68}$	$0.78^{+0.19}_{-0.29}$	$40.5^{+12.0}_{-7.6}$	$31.0^{+9.1}_{-8.2}$	4928^{+2725}_{-2435}	$0.01^{+0.26}_{-0.30}$	4.04	7.29	8.33
4	GW190904_104631	$21.24^{+5.76}_{-4.40}$	$0.64^{+0.31}_{-0.33}$	$31.3^{+14.5}_{-8.5}$	$19.7^{+7.1}_{-7.2}$	5614^{+4441}_{-2864}	$0.05^{+0.30}_{-0.37}$	4.50	4.88	6.64
5	GW190523_085933	$23.82^{+10.24}_{-7.95}$	$0.49^{+0.45}_{-0.32}$	$41.7^{+19.3}_{-15.5}$	$19.4^{+14.6}_{-10.5}$	6091^{+6613}_{-3702}	$0.42^{+0.31}_{-0.45}$	3.48	5.14	6.02
6	GW200208_211609	$18.83^{+4.68}_{-3.18}$	$0.69^{+0.28}_{-0.40}$	$26.9^{+14.6}_{-6.3}$	$18.0^{+6.4}_{-6.9}$	3669^{+3413}_{-1985}	$0.01^{+0.37}_{-0.37}$	4.75	6.22	7.83
7	GW190705_164632	$27.21^{+7.34}_{-5.24}$	$0.52^{+0.41}_{-0.32}$	$44.7^{+24.8}_{-12.8}$	$23.0^{+11.7}_{-9.8}$	5692^{+4030}_{-2863}	$0.29^{+0.26}_{-0.34}$	4.42	6.88	8.11
8	GW190426_082124	$17.93^{+4.12}_{-3.42}$	$0.45^{+0.45}_{-0.28}$	$31.5^{+22.5}_{-11.3}$	$13.8^{+6.9}_{-5.2}$	3213^{+4555}_{-1573}	$-0.01^{+0.39}_{-0.50}$	5.15	4.46	6.41

ARTIFICIAL NEURAL NETWORKS

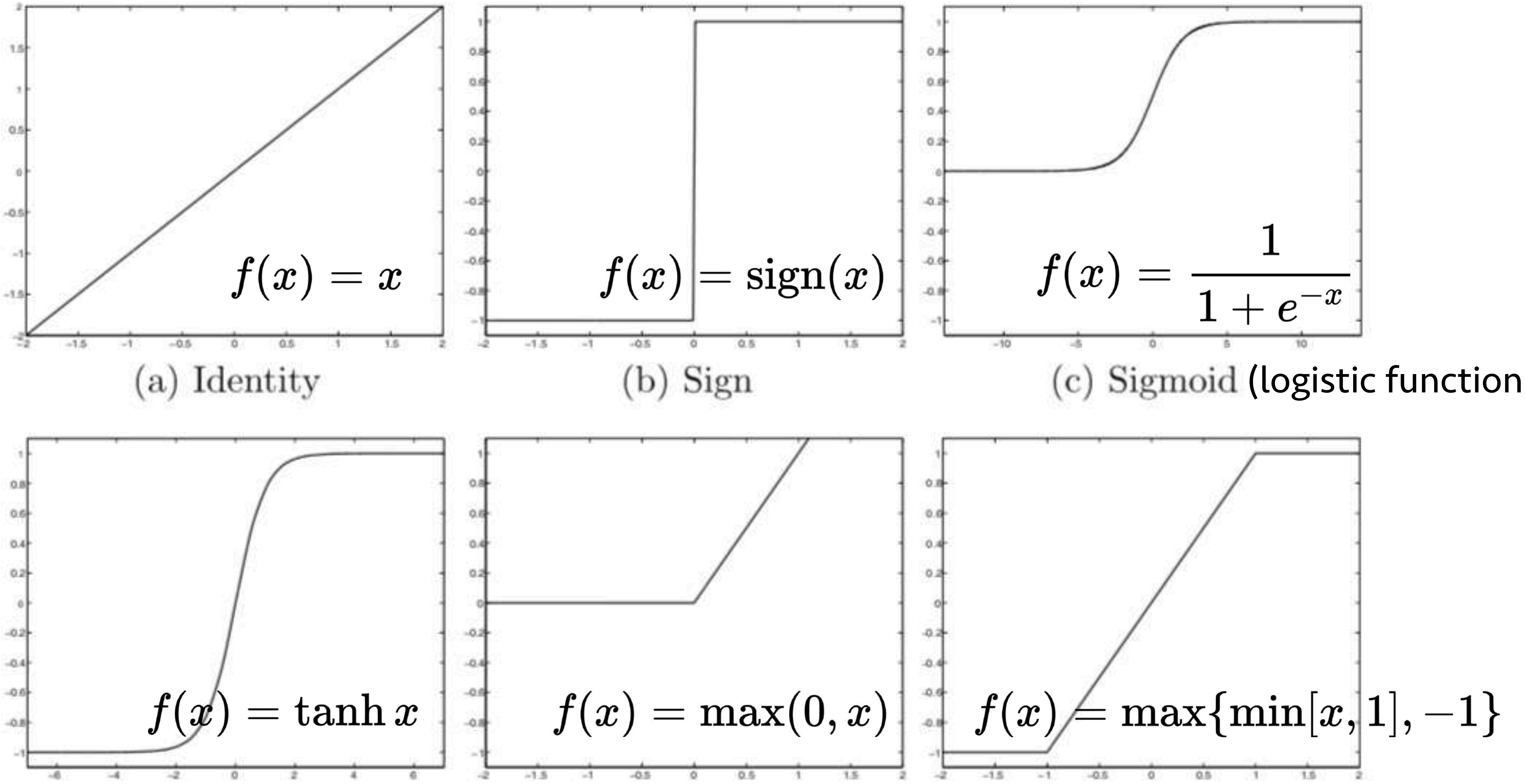


ARTIFICIAL NEURON

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

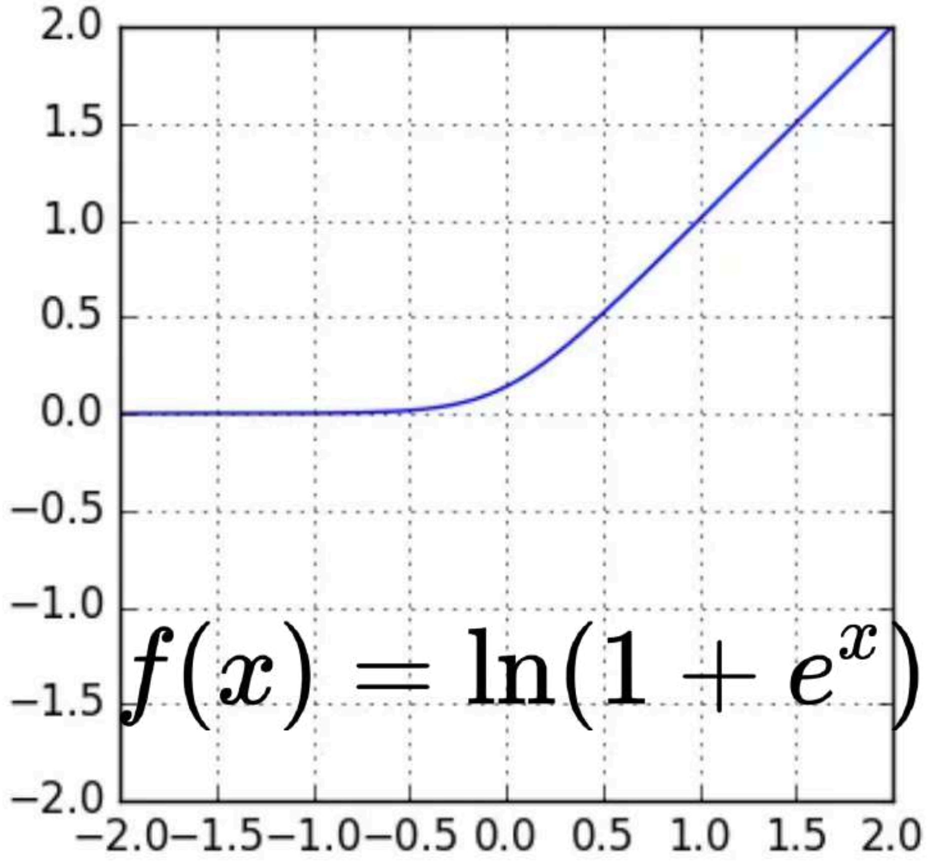


- Different common activation functions:



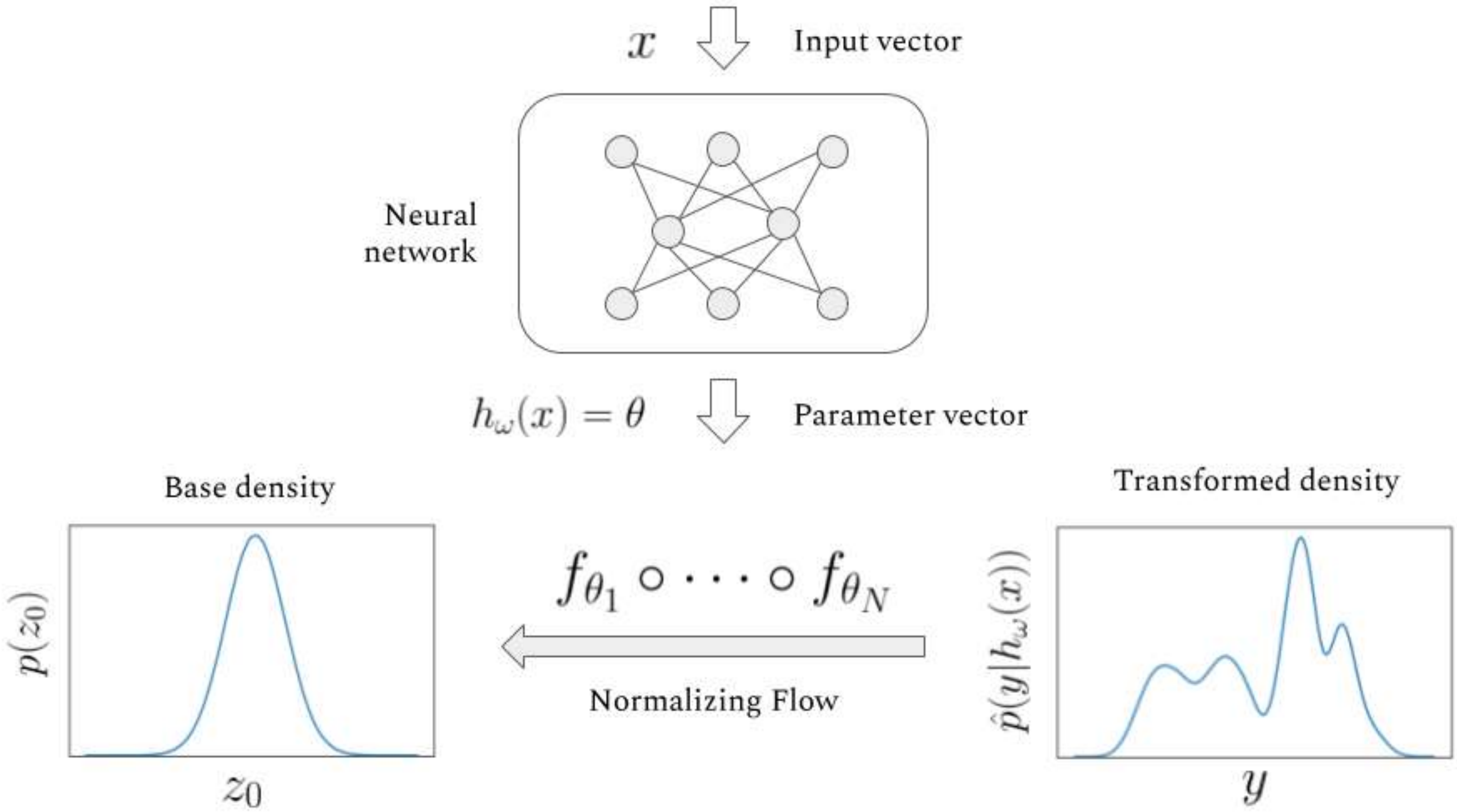
(a) Identity (b) Sign (c) Sigmoid (logistic function)

(d) Tanh (e) ReLU (Rectified Linear Unit) (f) Hard Tanh



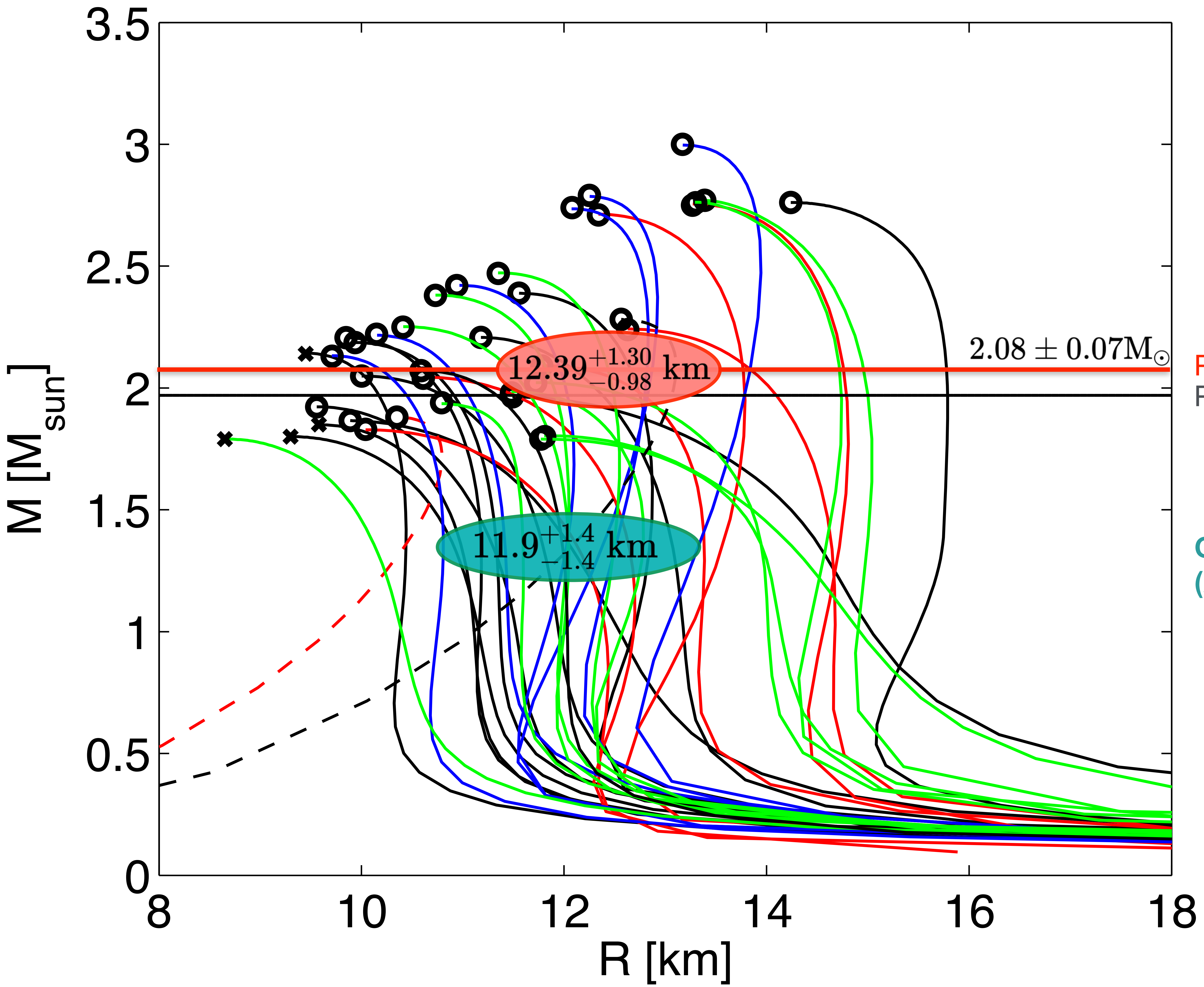
Softplus

NORMALIZING FLOWS



BACKUP SLIDES

CURRENT CONSTRAINTS ON MASS-RADIUS RELATION



Raajmakers et al. (2021)

PSR J0740+6620 (NANOGrav/CHIME) +NICER

PSR J0348+0432

GW170817 (LIGO/Virgo)
(Λ < 800)

ANN REGRESSION IN FREQUENCY DOMAIN

EOS	Waveforms	References	Component masses (in $M_\odot$)
ALF2	10	Rezzolla and Takami [45] and CoRe v2 [27]	1.2, 1.225, 1.25, 1.275, 1.3, 1.325, 1.35, 1.3505, 1.375, 1.3755
APR4	7	Rezzolla and Takami [45]	1.2, 1.225, 1.25, 1.275, 1.3, 1.325, 1.35
BHBlp	4	CoRe v2 [27]	1.25, 1.3, 1.35, 1.4
BLh	4	CoRe v2 [27]	1.3, 1.3325, 1.364, 1.4
DD2	7	CoRe v2 [27]	1.2, 1.25, 1.3, 1.35, 1.364, 1.4, 1.5
ENG	1	CoRe v2 [27]	1.3495
GNH3	7	Rezzolla and Takami [45]	1.2, 1.225, 1.25, 1.275, 1.3, 1.325, 1.35
H4	13	Rezzolla and Takami [45] and CoRe v2 [27]	1.2, 1.225, 1.25, 1.275, 1.3, 1.325, 1.3495 1.35, 1.3505, 1.3715, 1.3725, 1.3735, 1.3795
LS220	4	CoRe v2 [27]	1.2, 1.35, 1.364, 1.4
MPA1	8	Soultanis *et al.* [97]	1.2, 1.25, 1.3, 1.35, 1.4, 1.45, 1.5, 1.55
MS1	2	CoRe v2 [27]	1.3495, 1.351
MS1b	8	CoRe v2 [27]	1.35, 1.3505, 1.375, 1.3805, 1.381, 1.5, 1.6, 1.7
SFHo	2	CoRe v2 [27]	1.35, 1.364
SLy	10	Rezzolla and Takami [45] and CoRe v2 [27]	1.2, 1.225, 1.25, 1.275, 1.3, 1.325, 1.35, 1.351, 1.3575, 1.364

BAYESIAN PARAMETER ESTIMATION

Injected waveform

$$d(t) = h(t) + n(t)$$

Gaussian likelihood function

$$\mathcal{L}(d \mid \boldsymbol{\theta}) \propto \exp[-\langle d(t) - h(\boldsymbol{\theta}, t), d(t) - h(\boldsymbol{\theta}, t)\rangle]$$

where the inner product is

$$\langle h, g\rangle = 4\mathfrak{Re}\int_0^\infty \frac{\tilde{h}(f)\tilde{g}^*(f)}{S(f)}df$$

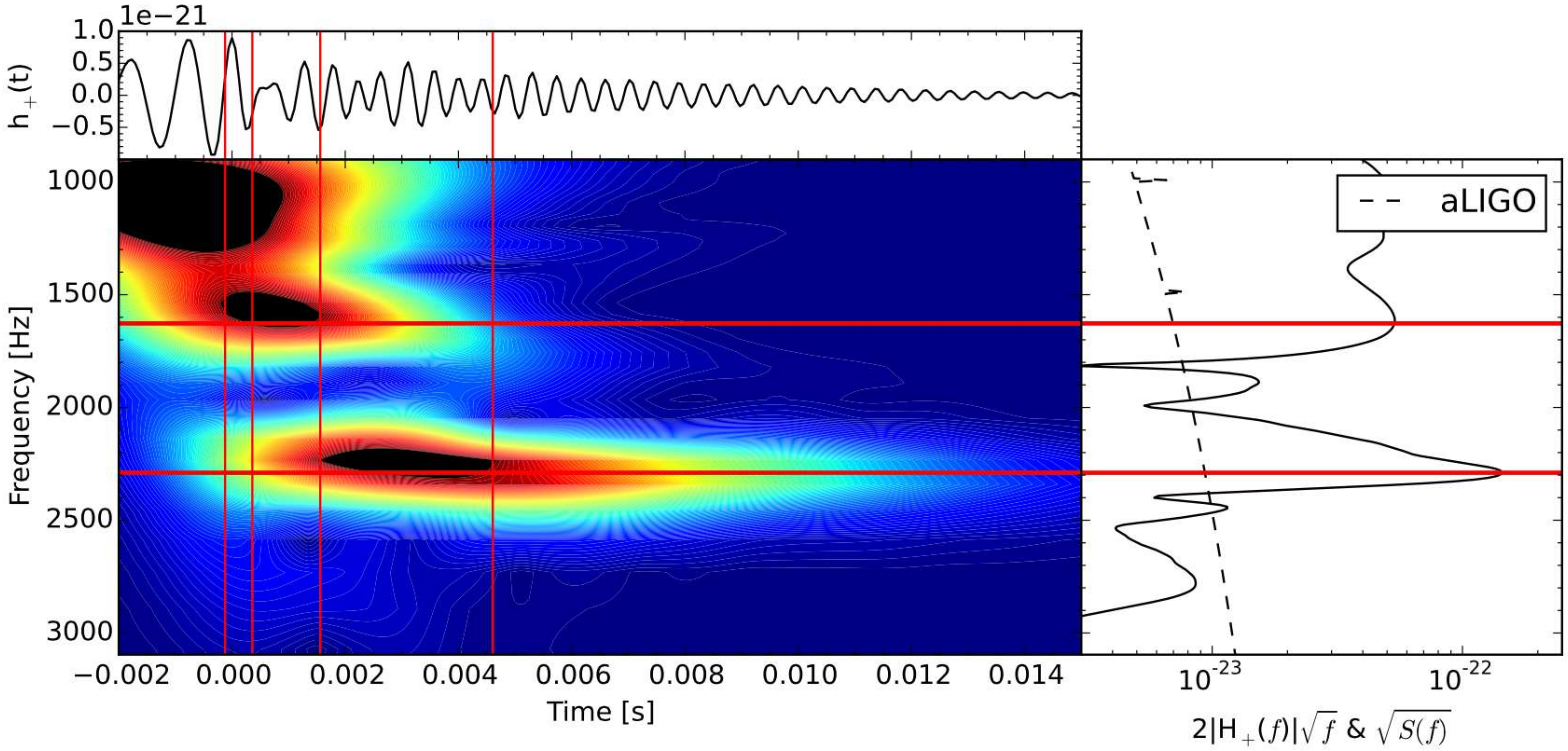
Fitting factor

$$\mathcal{F}(d(t), h(\boldsymbol{\theta}, t)) \equiv \frac{\langle d(t) \mid h(\boldsymbol{\theta}, t)\rangle}{\sqrt{\langle d(t) \mid d(t)\rangle\langle h(\boldsymbol{\theta}, t) \mid h(\boldsymbol{\theta}, t)\rangle}}$$

Optimal SNR

$$\rho_{\mathrm{opt},i} = \langle h, h\rangle^{1/2} \qquad \rho_{\mathrm{opt}} = \left(\sum_{i\in\mathrm{HLV}} \rho^2_{\mathrm{opt},i}\right)^{1/2}$$

POST-MERGER SPECTOGRAMS

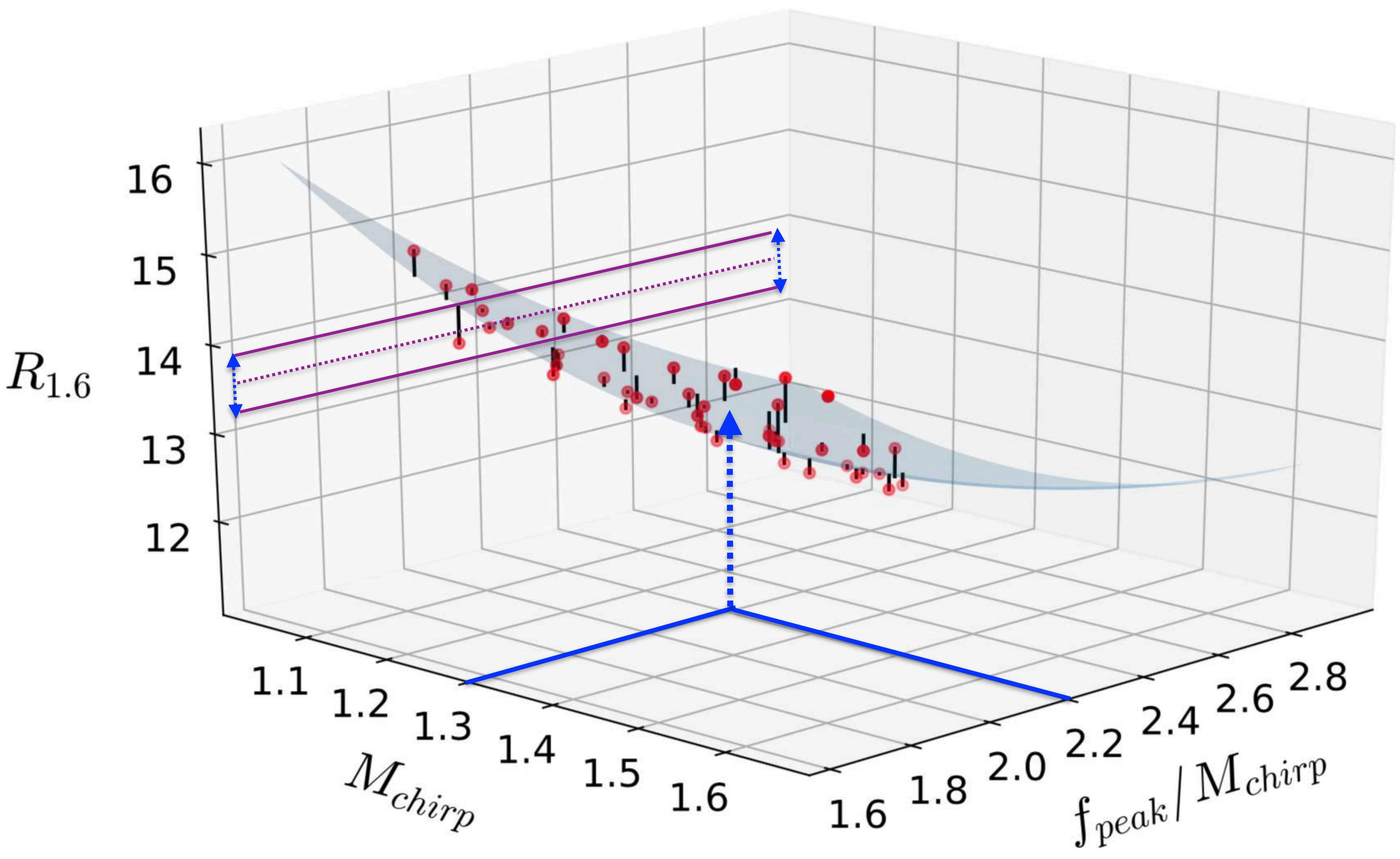


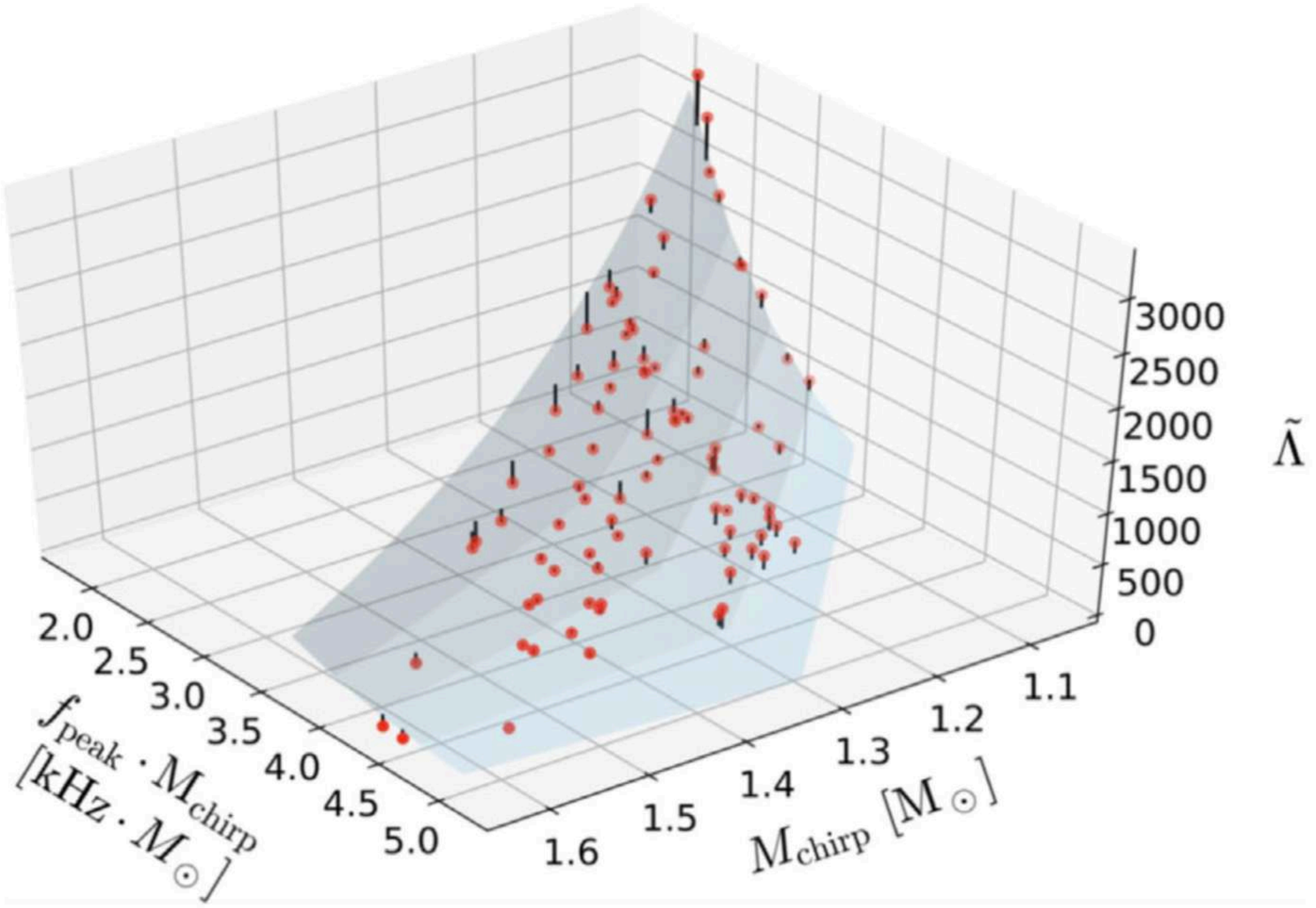
Clark, Bauswein, Stergioulas, Shoemaker (2016)

EMPIRICAL RELATIONS FOR GW ASTEROSEISMOLOGY OF BNS MERGERS

$$R_{1.6} = 43.796 - 19.984 M_{\rm chirp} - 12.921 f_{\rm peak}/M_{\rm chirp} + 4.674 M^2_{\rm chirp} + 3.371 f_{\rm peak} + 1.26 (f_{\rm peak}/M_{\rm chirp})^2$$

$$\tilde{\Lambda} = -1344 + 108.9 M_{\rm chirp} f_{\rm peak} + 17208 f_{\rm peak}^{-2}$$

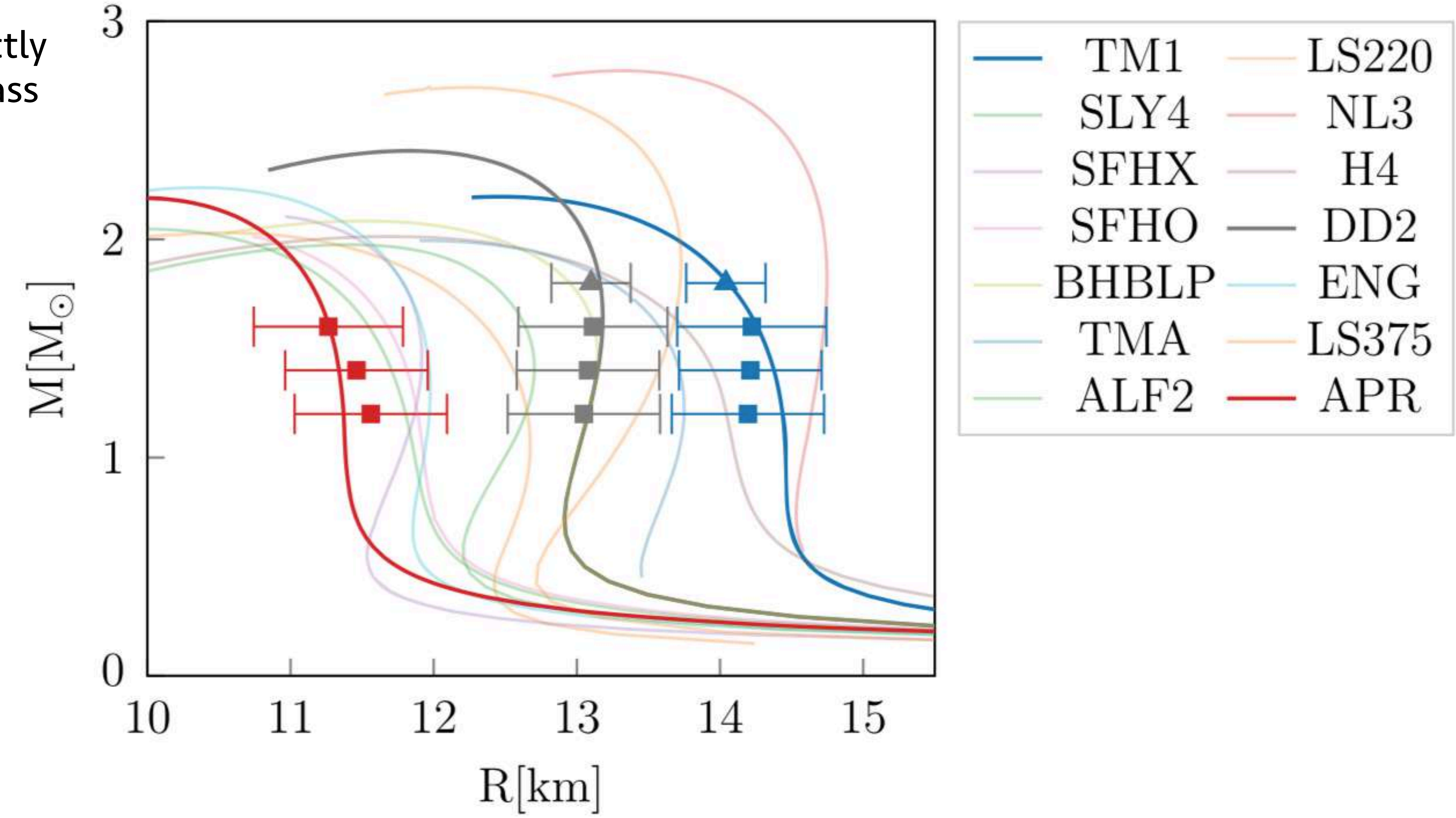




Vretinaris, Stergioulas & Bauswein (2020)

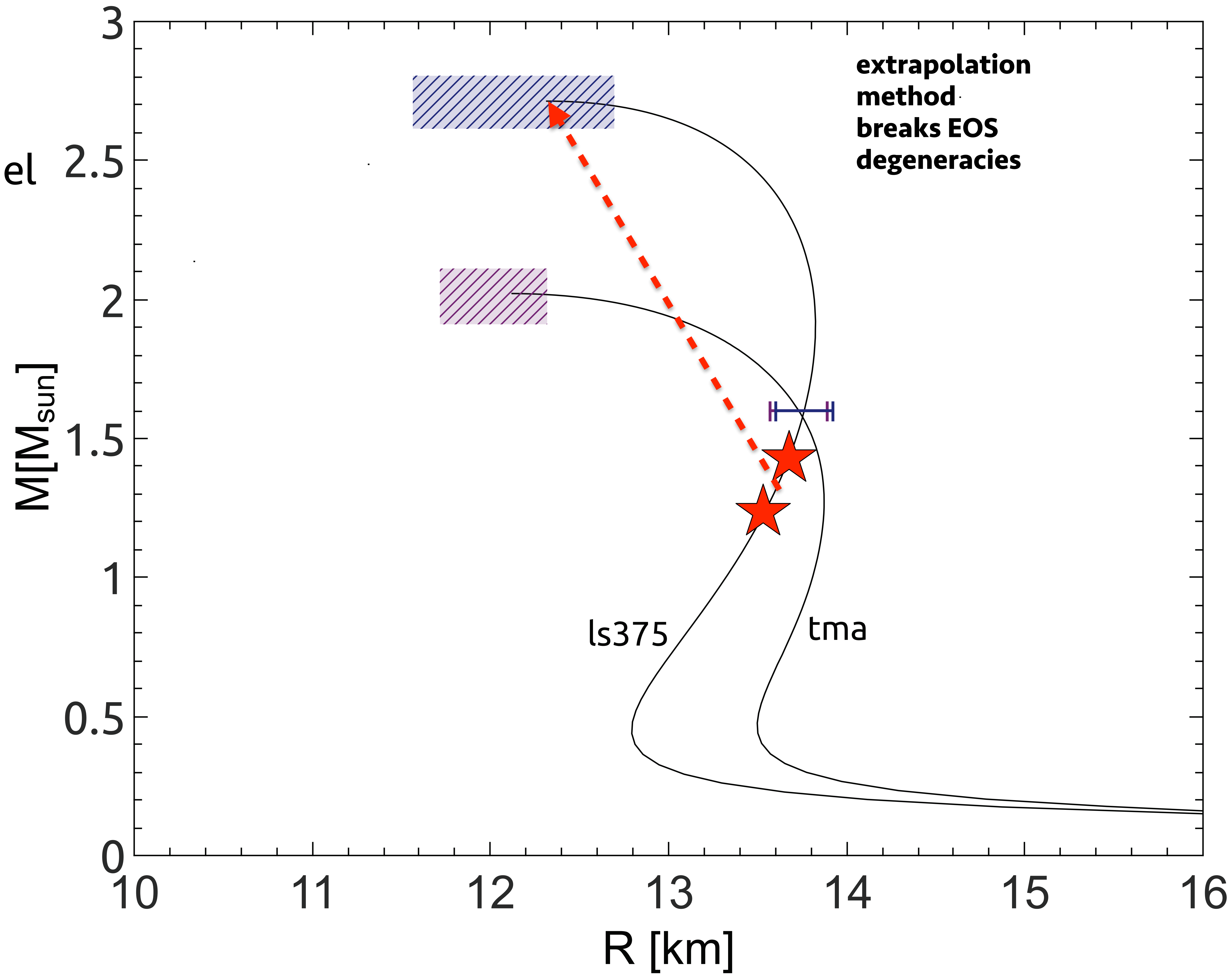
EOS CONSTRAINTS THROUGH POST-MERGER REMNANTS

It will be possible to directly extract the radius in a mass range $1.2 - 1.8M_{\odot}$



Vretinaris, Stergioulas & Bauswein (2020)

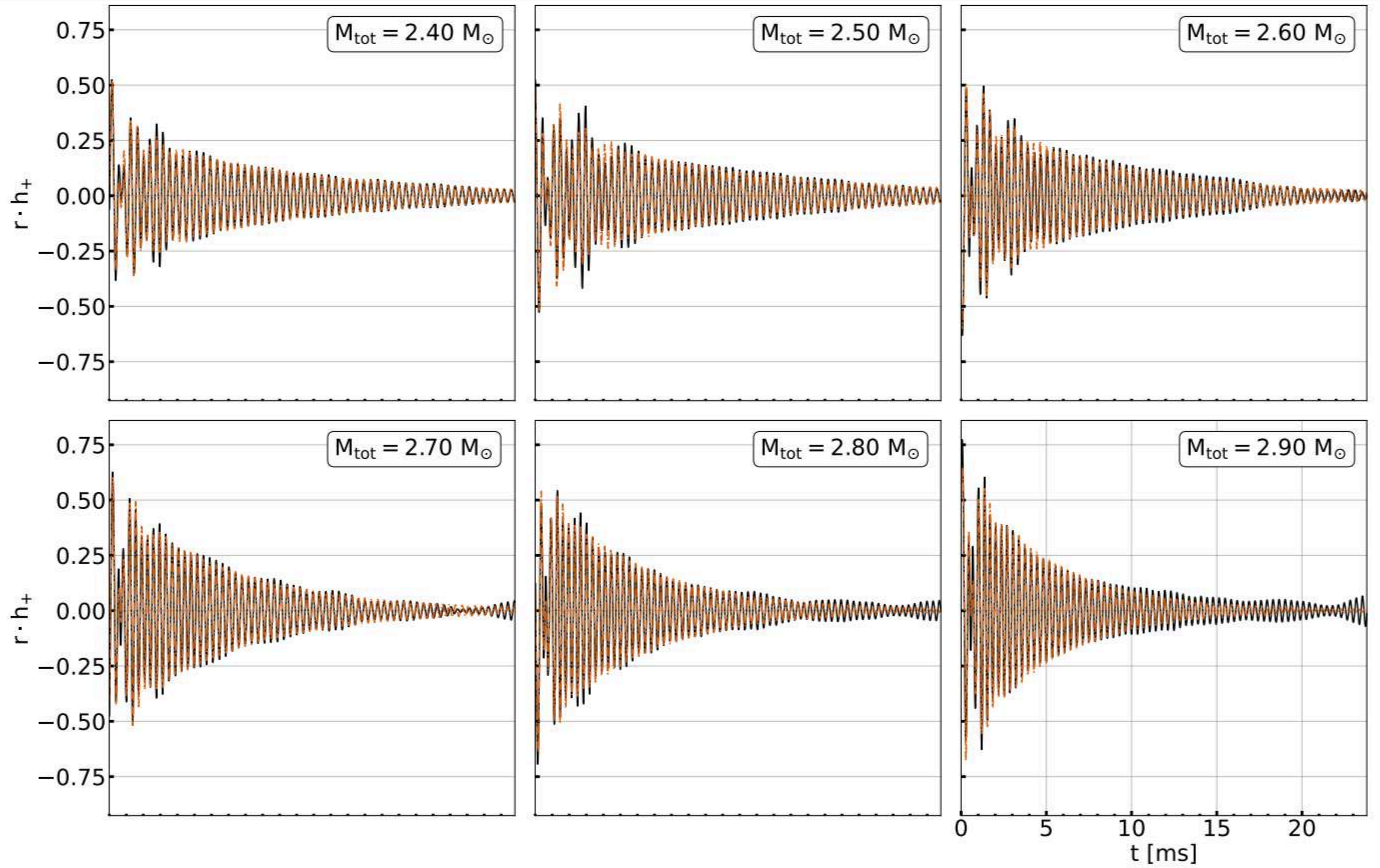
Extrapolation method allows for constraints on maximum mass model using only low-mass detections.

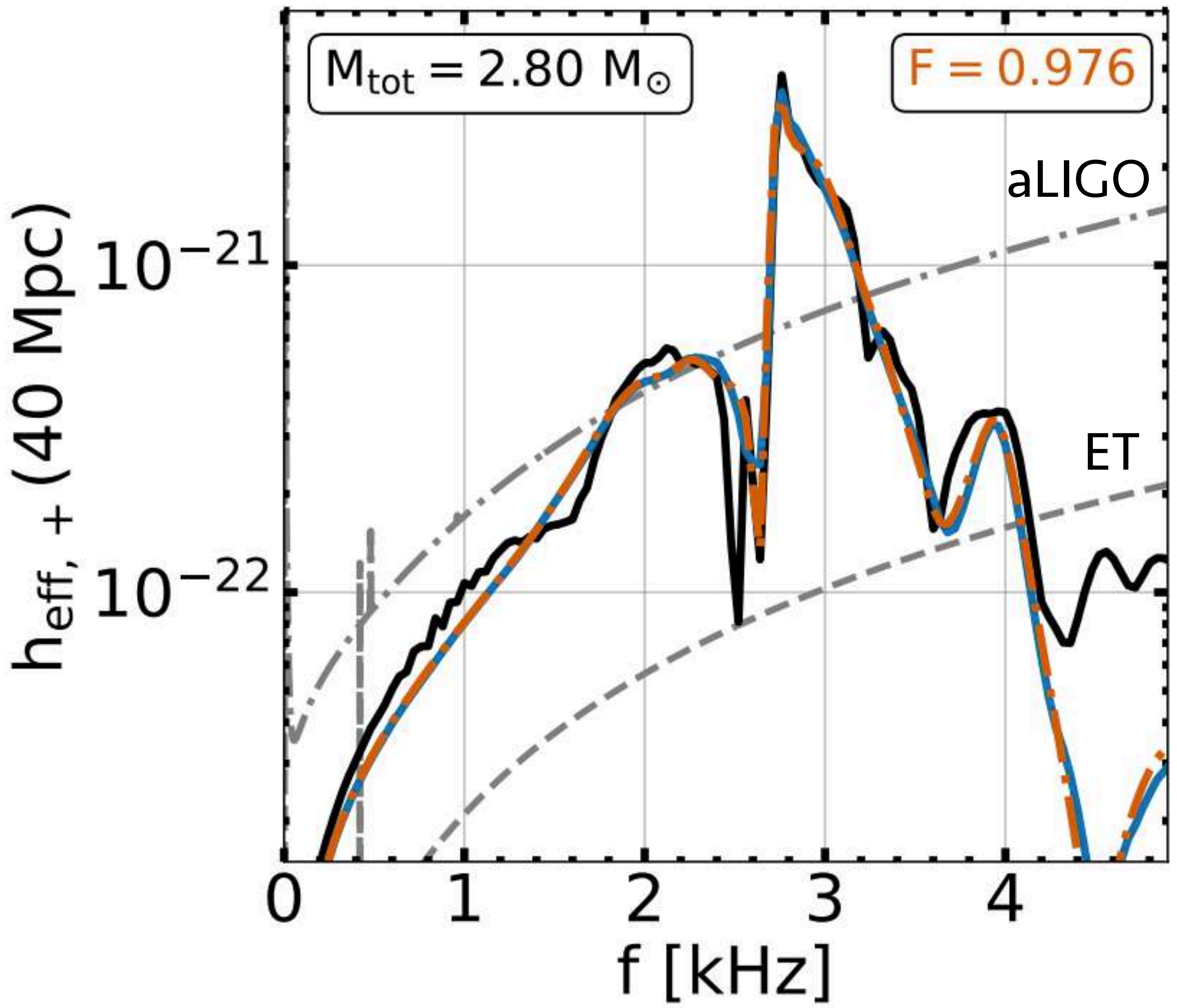


Bauswein, Stergioulas (2015)

NEW ANALYTIC WAVEFORM TEMPLATE

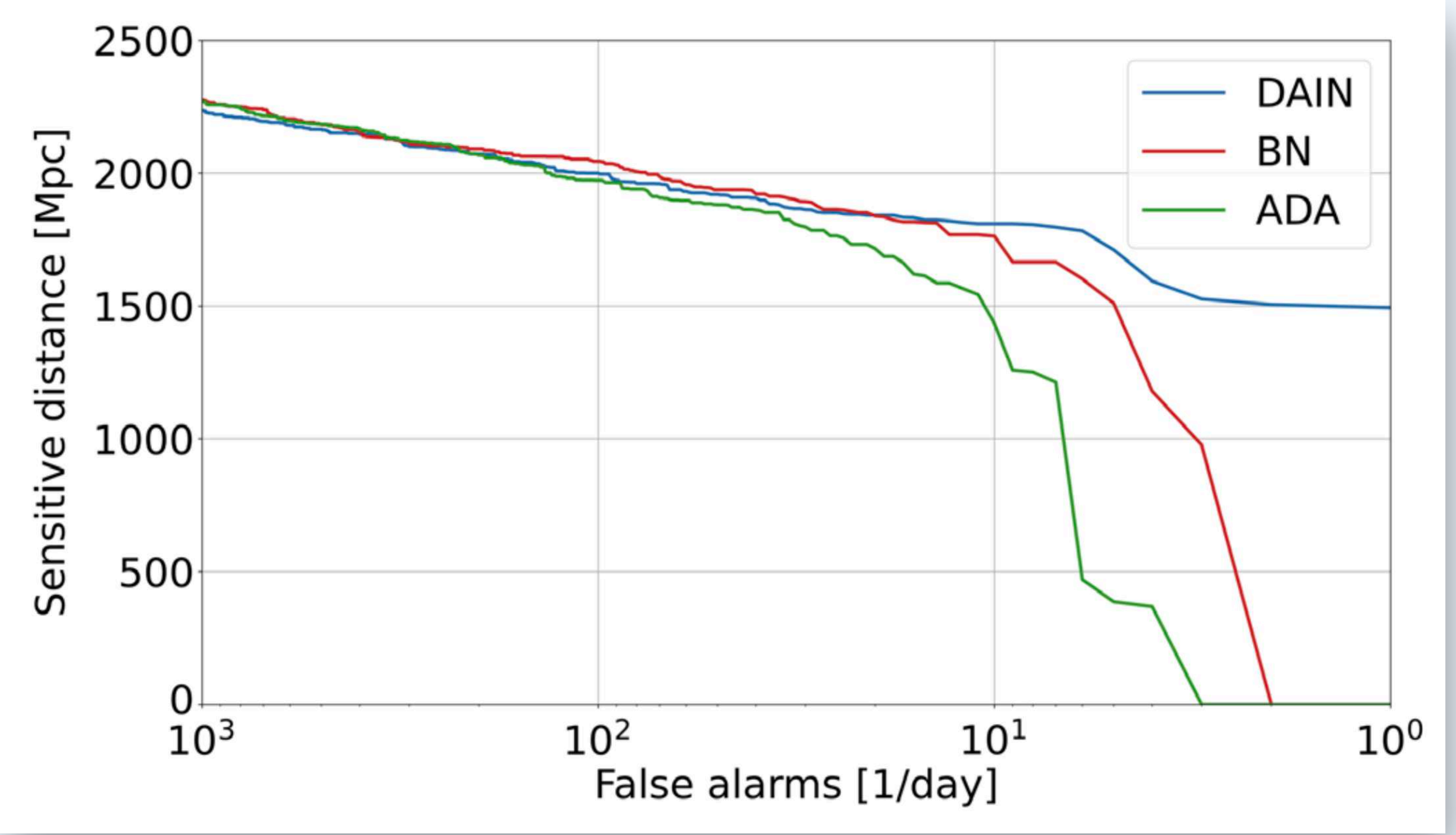
Relies only on physical parameters





Soultanis, Bauswein, Stergioulas (2022)

Significant Improvement in Sensitive Distance using DAIN

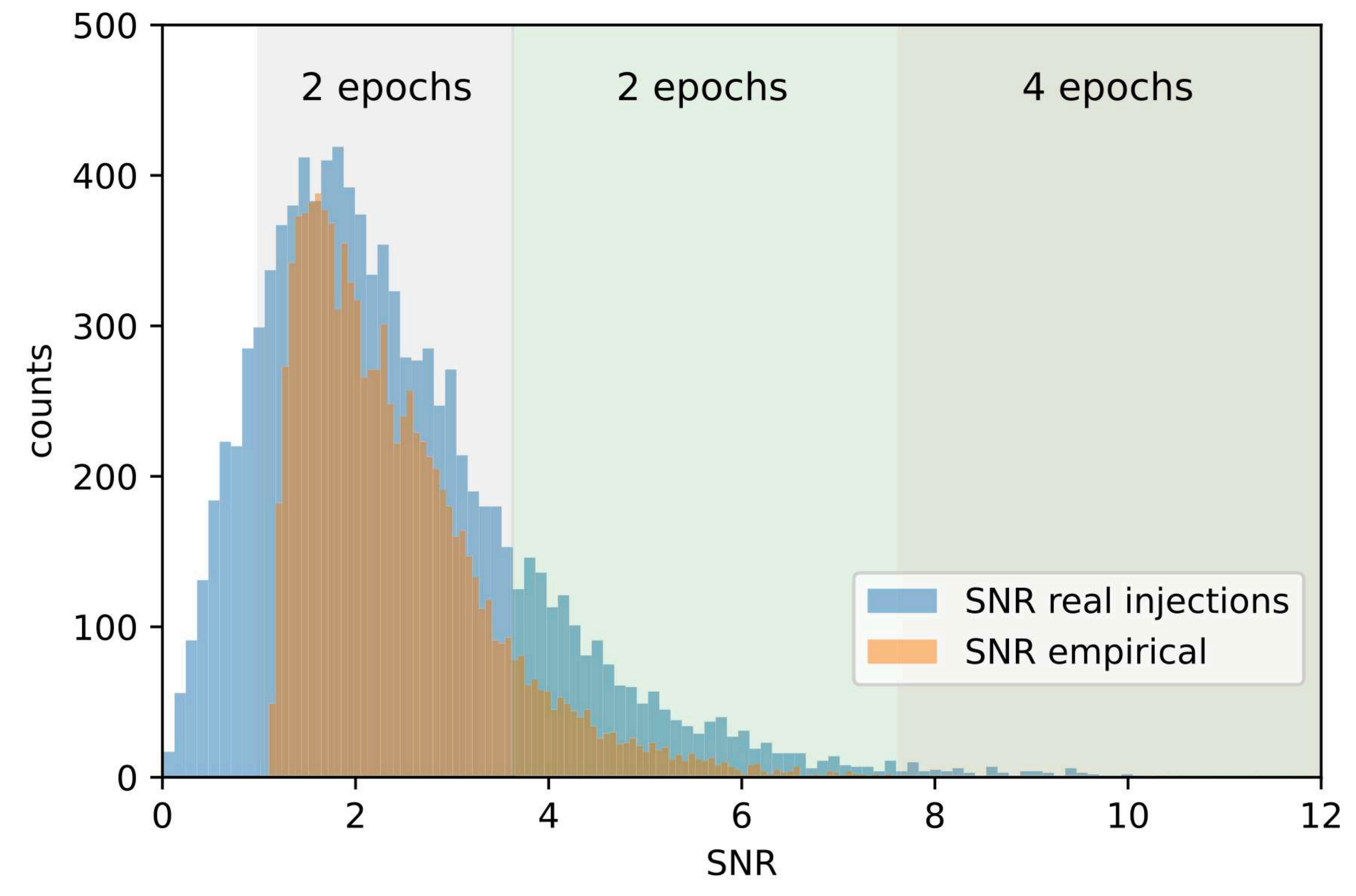


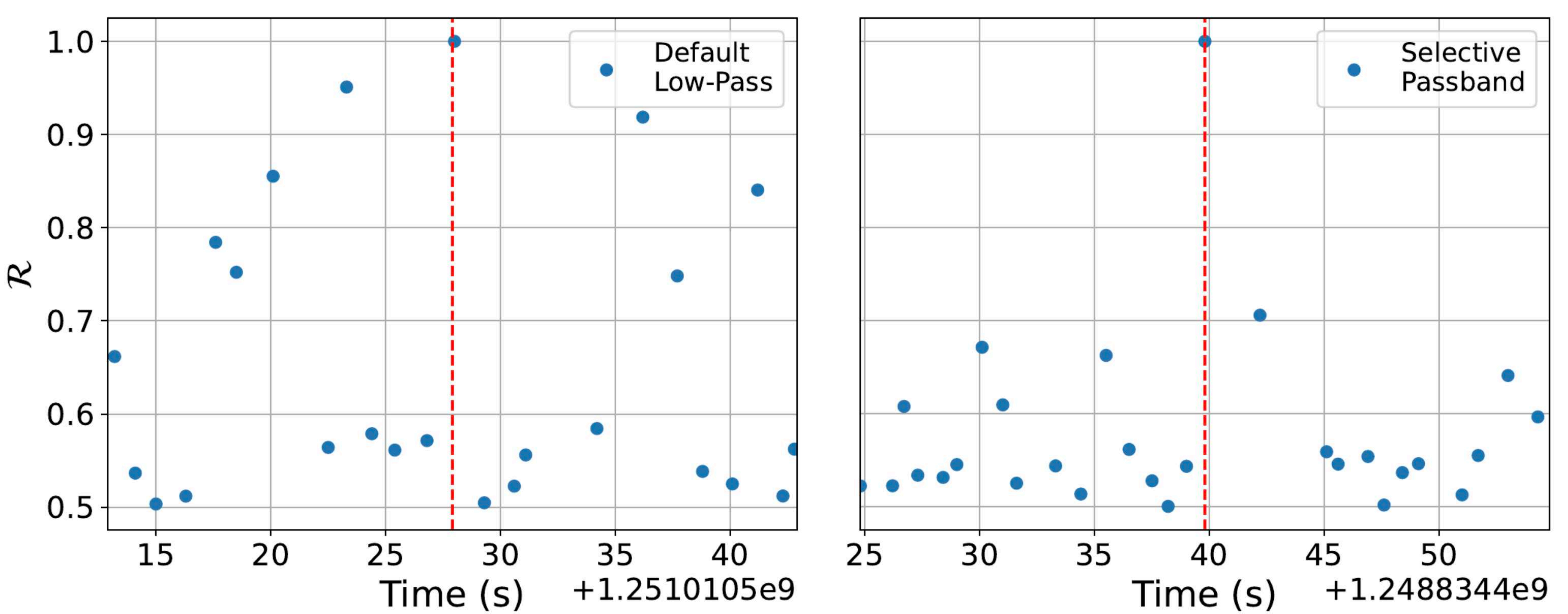
CURRICULUM LEARNING

Curriculum learning based on signal-to-noise ratio (**SNR**)

Epochs	Min(SNR)	Max(SNR)
4	7.63	100
2	3.63	100
2	1	3.63
2	1	7.63
Remaining	0	∞

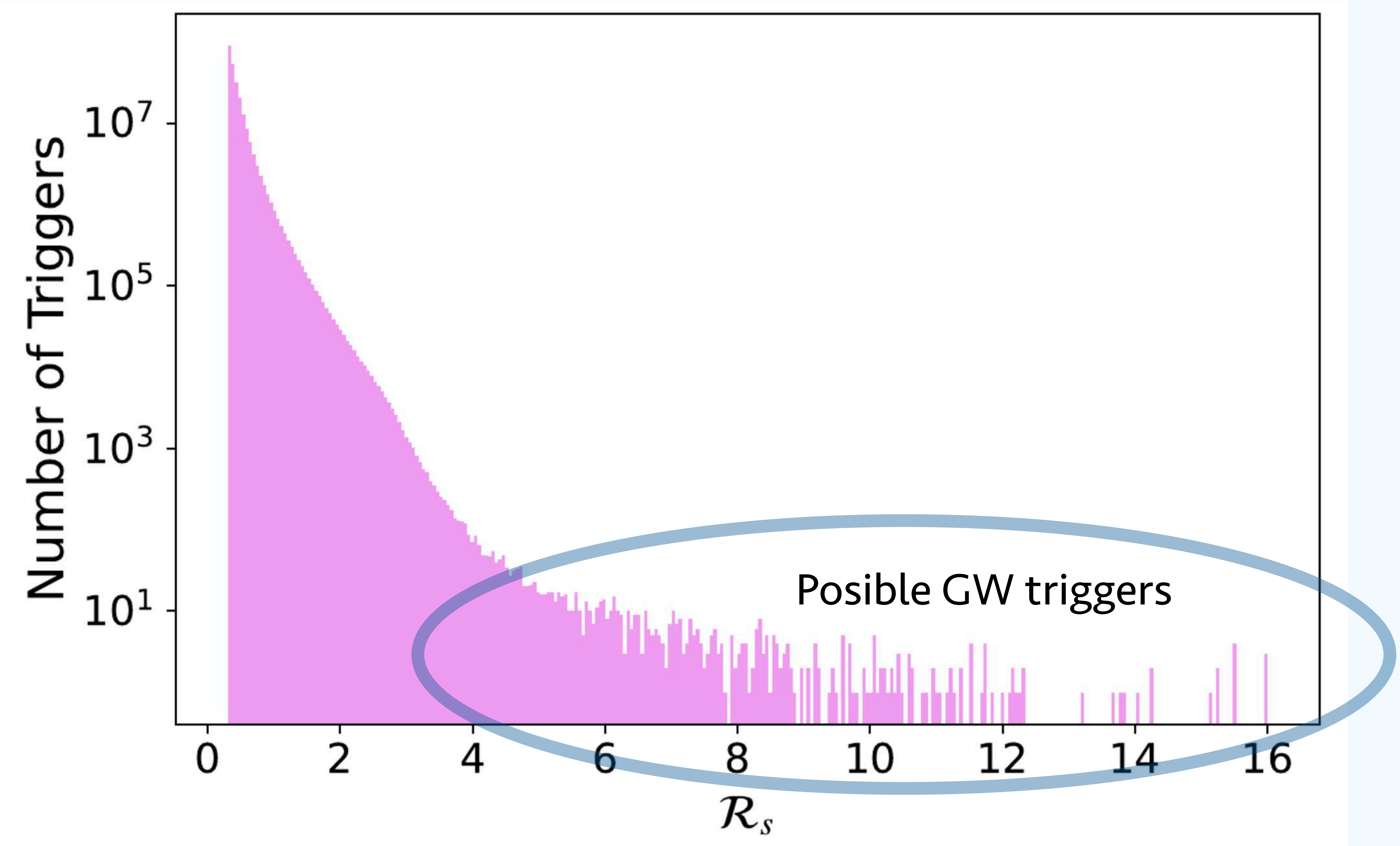
The networks learns the louder signals first and then refines its weights to accommodate weaker signals





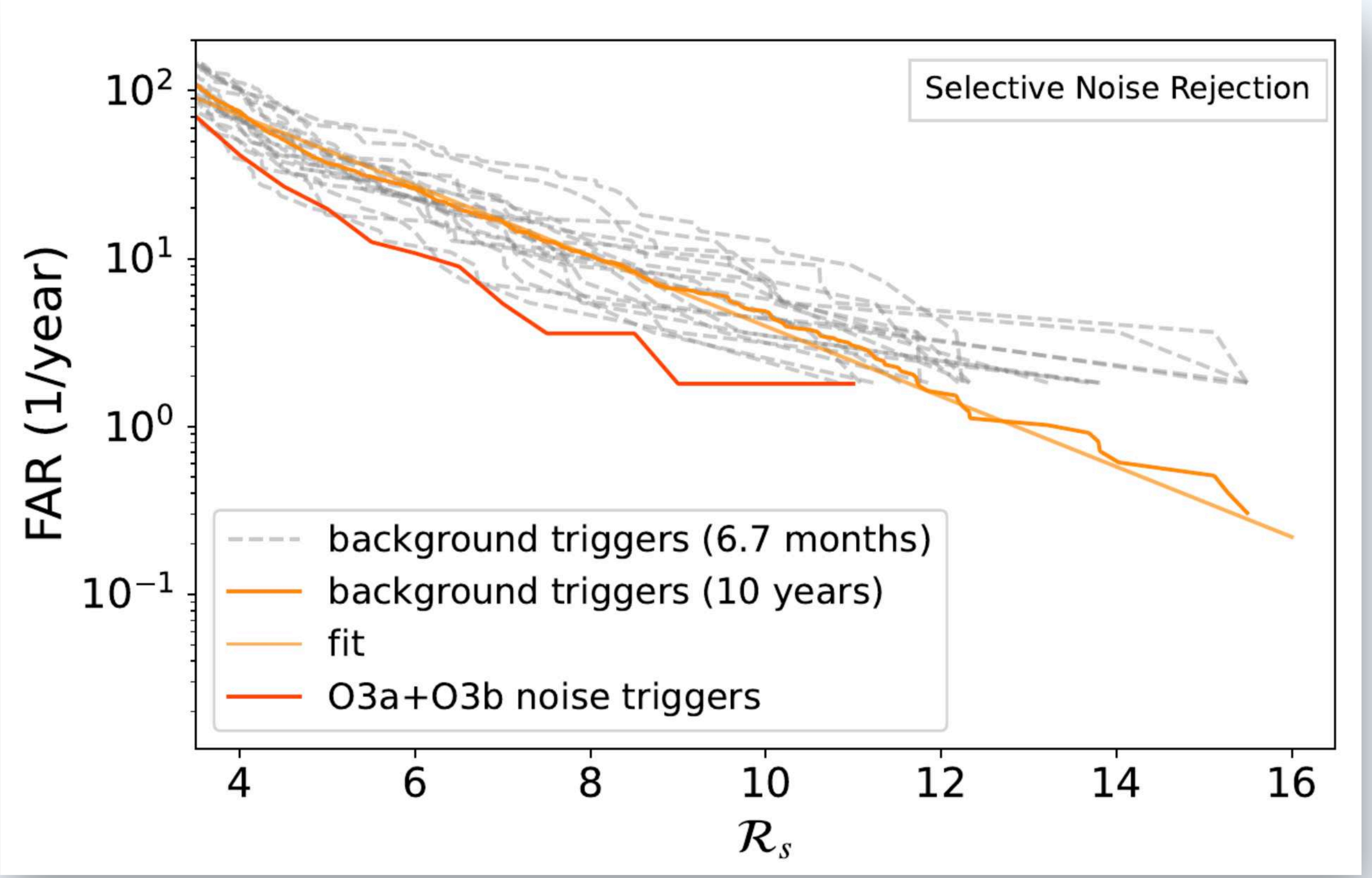
1.0
0.9
0.8
0.7
0.6
0.5
$\mathcal{R}$
Default Low-Pass
15 20 25 30 35 40
Time (s) +1.2510105e9
Selective Passband
25 30 35 40 45 50
Time (s) +1.2488344e9

RANKING STATISTIC DISTRIBUTION

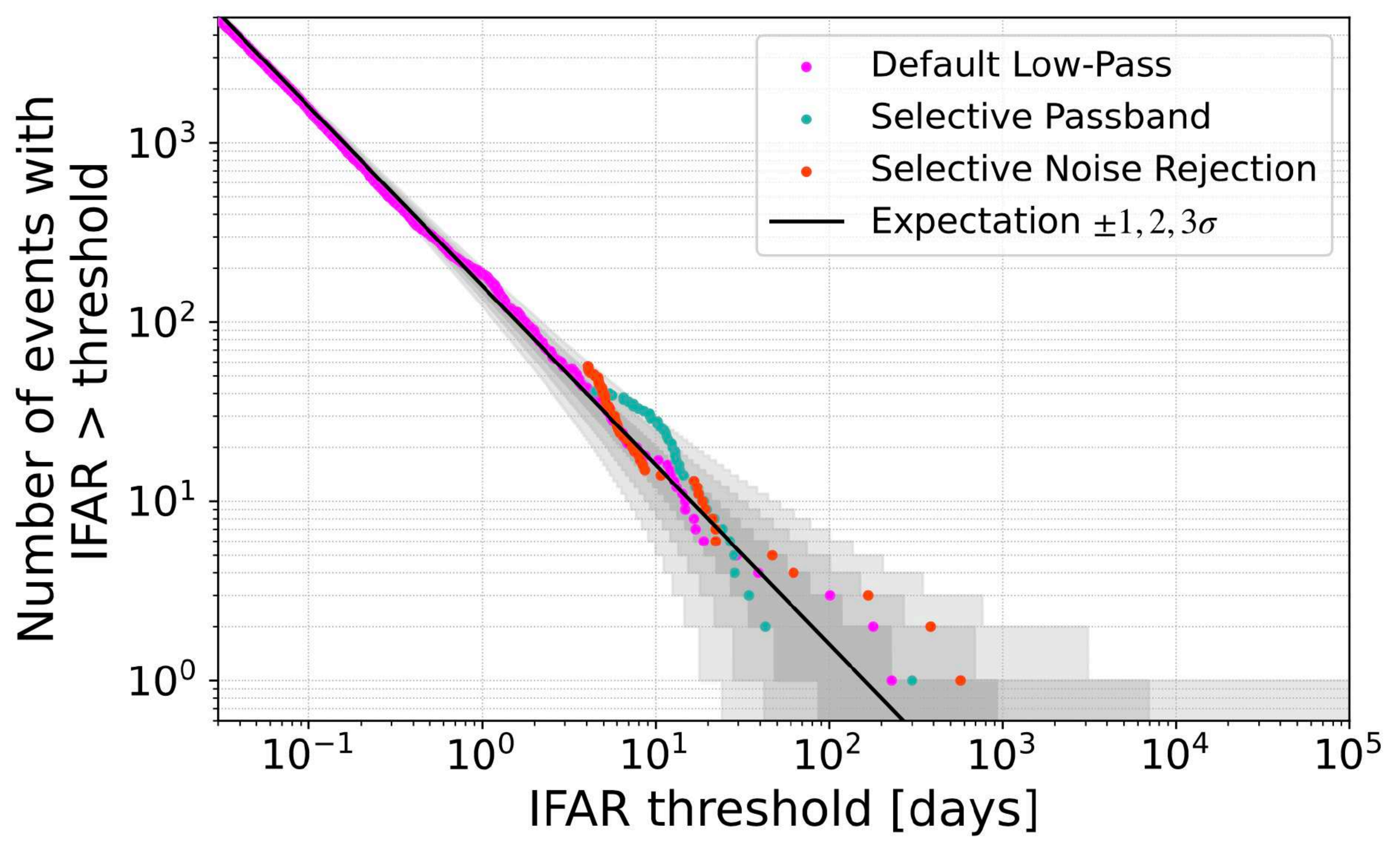


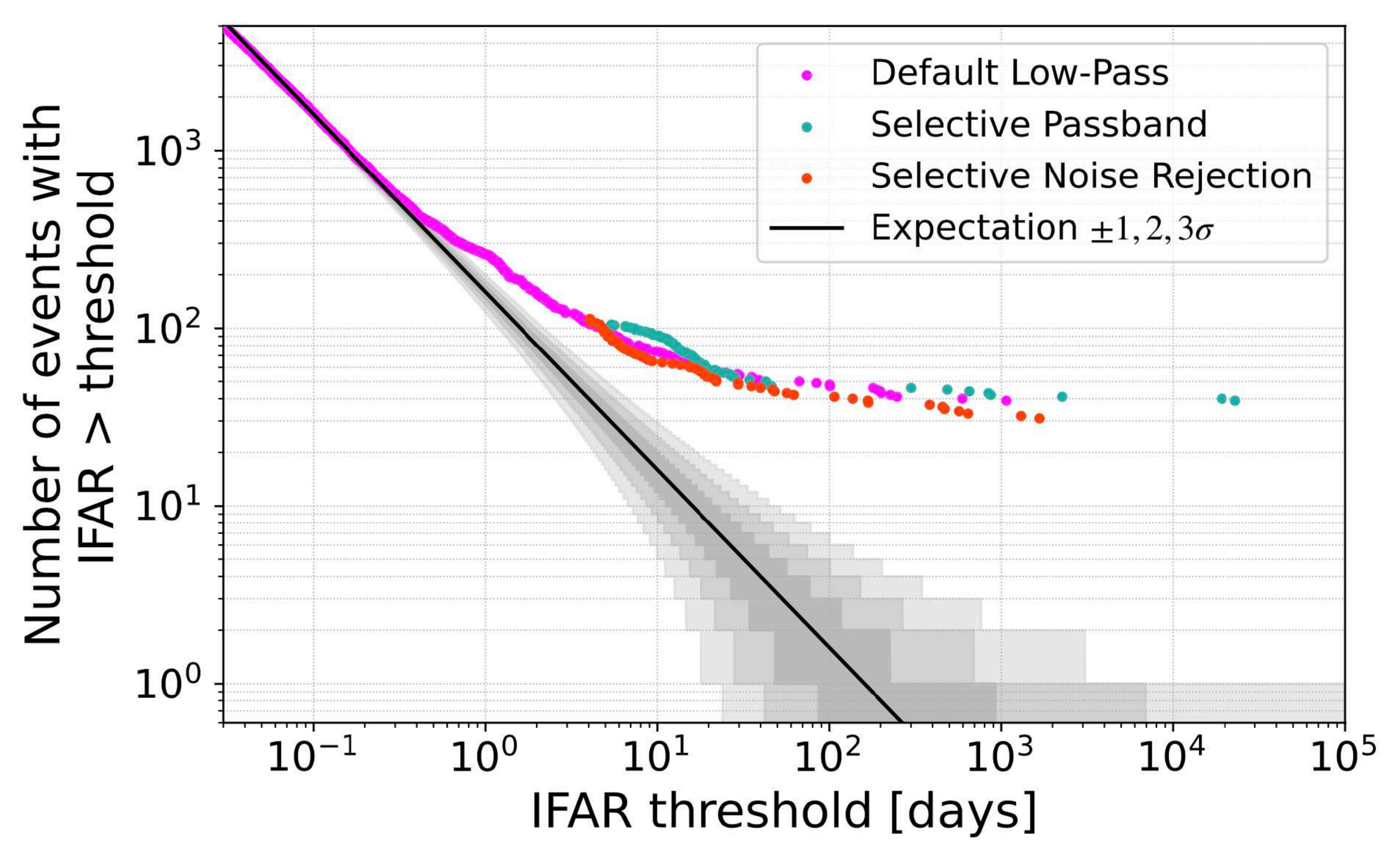
FALSE-ALARM RATE

FAR: We construct a 10-year background using time-slides of O3 noise

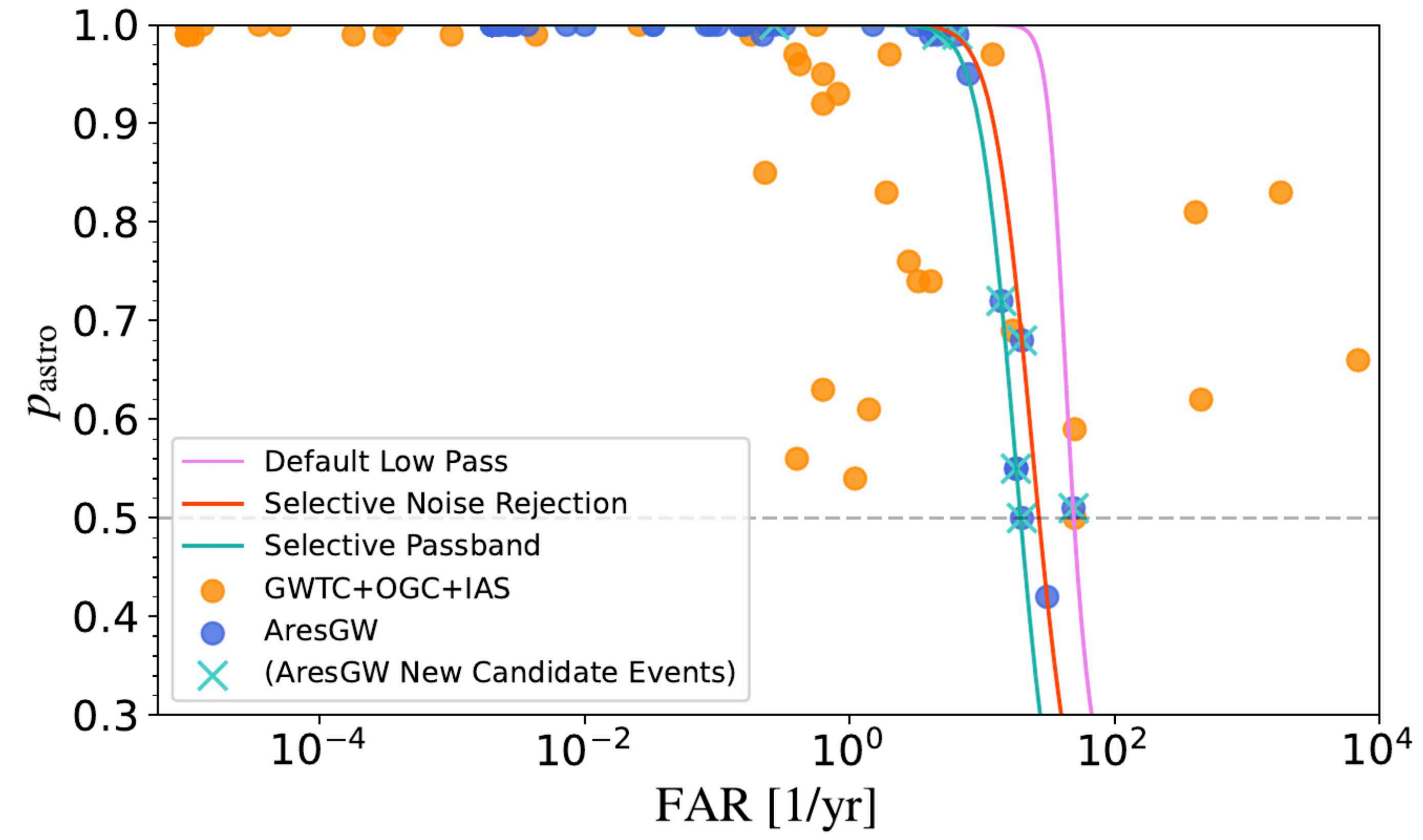


AresGW **performance on O3 data**





ASTROPHYSICAL PROBABILITY



PARAMETERS OF NEW GW DETECTIONS WITH AresGW

TABLE VIII: Parameter estimation for the new AresGW candidate events.

#	Event Name	$\mathcal{M}$ ($M_\odot$)	q	m_1 ($M_\odot$)	m_2 ($M_\odot$)	D_L (Mpc)	χ_{eff}	SNR (H1)	SNR (L1)	SNR $\hat{\rho}$ (network)
1	GW190511_125545	$28.95^{+9.45}_{-6.86}$	$0.72^{+0.25}_{-0.36}$	$40.7^{+16.2}_{-10.5}$	$28.2^{+11.6}_{-11.2}$	3707^{+3471}_{-2173}	$0.23^{+0.25}_{-0.29}$	2.29	7.34	7.29
2	GW190614_134749	$25.97^{+16.59}_{-6.20}$	$0.70^{+0.27}_{-0.36}$	$37.0^{+31.8}_{-10.7}$	$25.2^{+15.2}_{-9.7}$	6551^{+9562}_{-3558}	$0.05^{+0.34}_{-0.34}$	3.51	6.08	7.02
3	GW190607_083827	$30.48^{+7.21}_{-4.68}$	$0.78^{+0.19}_{-0.29}$	$40.5^{+12.0}_{-7.6}$	$31.0^{+9.1}_{-8.2}$	4928^{+2725}_{-2435}	$0.01^{+0.26}_{-0.30}$	4.04	7.29	8.33
4	GW190904_104631	$21.24^{+5.76}_{-4.40}$	$0.64^{+0.31}_{-0.33}$	$31.3^{+14.5}_{-8.5}$	$19.7^{+7.1}_{-7.2}$	5614^{+4441}_{-2864}	$0.05^{+0.30}_{-0.37}$	4.50	4.88	6.64
5	GW190523_085933	$23.82^{+10.24}_{-7.95}$	$0.49^{+0.45}_{-0.32}$	$41.7^{+19.3}_{-15.5}$	$19.4^{+14.6}_{-10.5}$	6091^{+6613}_{-3702}	$0.42^{+0.31}_{-0.45}$	3.48	5.14	6.02
6	GW200208_211609	$18.83^{+4.68}_{-3.18}$	$0.69^{+0.28}_{-0.40}$	$26.9^{+14.6}_{-6.3}$	$18.0^{+6.4}_{-6.9}$	3669^{+3413}_{-1985}	$0.01^{+0.37}_{-0.37}$	4.75	6.22	7.83
7	GW190705_164632	$27.21^{+7.34}_{-5.24}$	$0.52^{+0.41}_{-0.32}$	$44.7^{+24.8}_{-12.8}$	$23.0^{+11.7}_{-9.8}$	5692^{+4030}_{-2863}	$0.29^{+0.26}_{-0.34}$	4.42	6.88	8.11
8	GW190426_082124	$17.93^{+4.12}_{-3.42}$	$0.45^{+0.45}_{-0.28}$	$31.5^{+22.5}_{-11.3}$	$13.8^{+6.9}_{-5.2}$	3213^{+4555}_{-1573}	$-0.01^{+0.39}_{-0.50}$	5.15	4.46	6.41

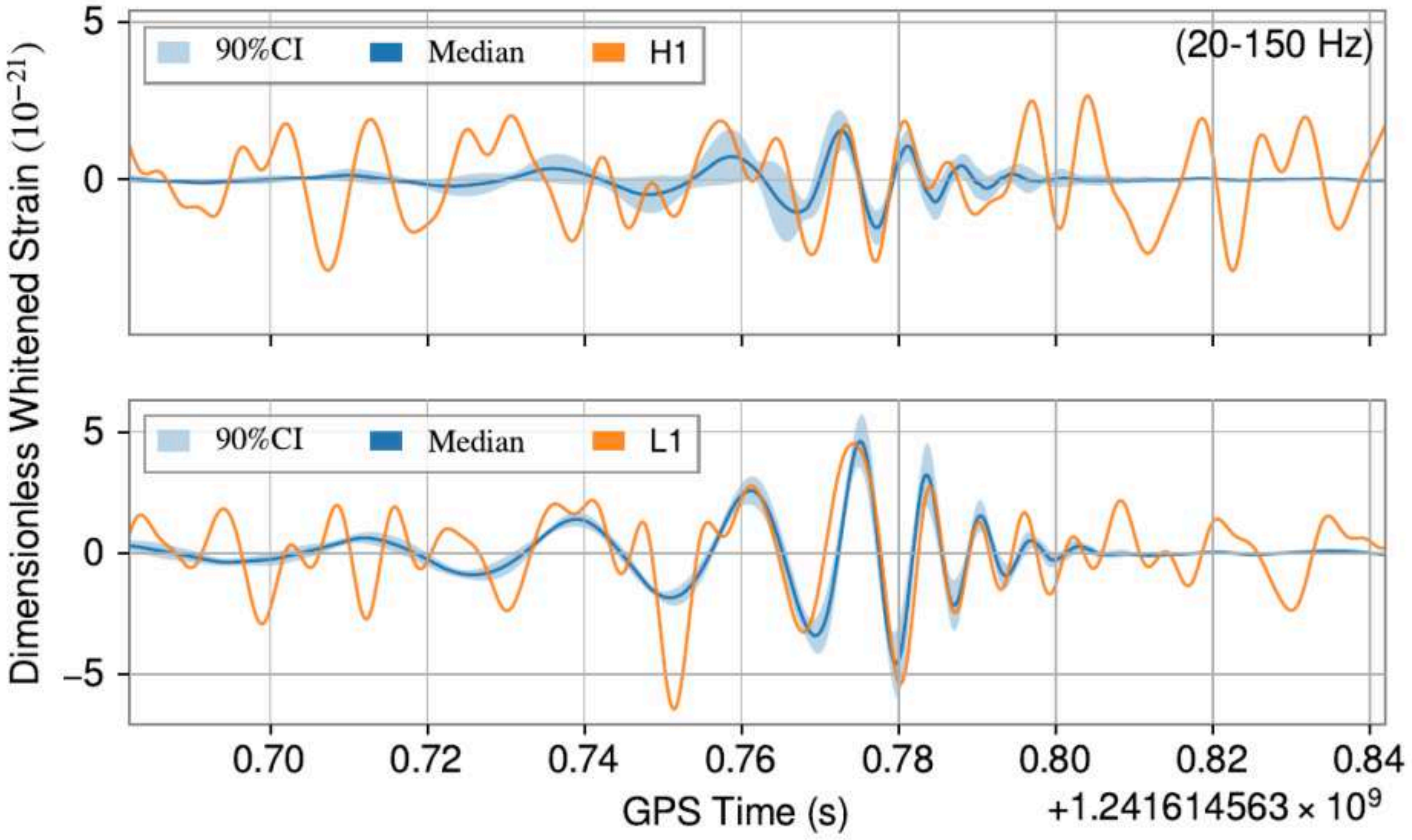


FIG. 35: Whitened, bandpassed strain data and reconstructed waveform for the new event GW190511_125545 identified by AresGW.

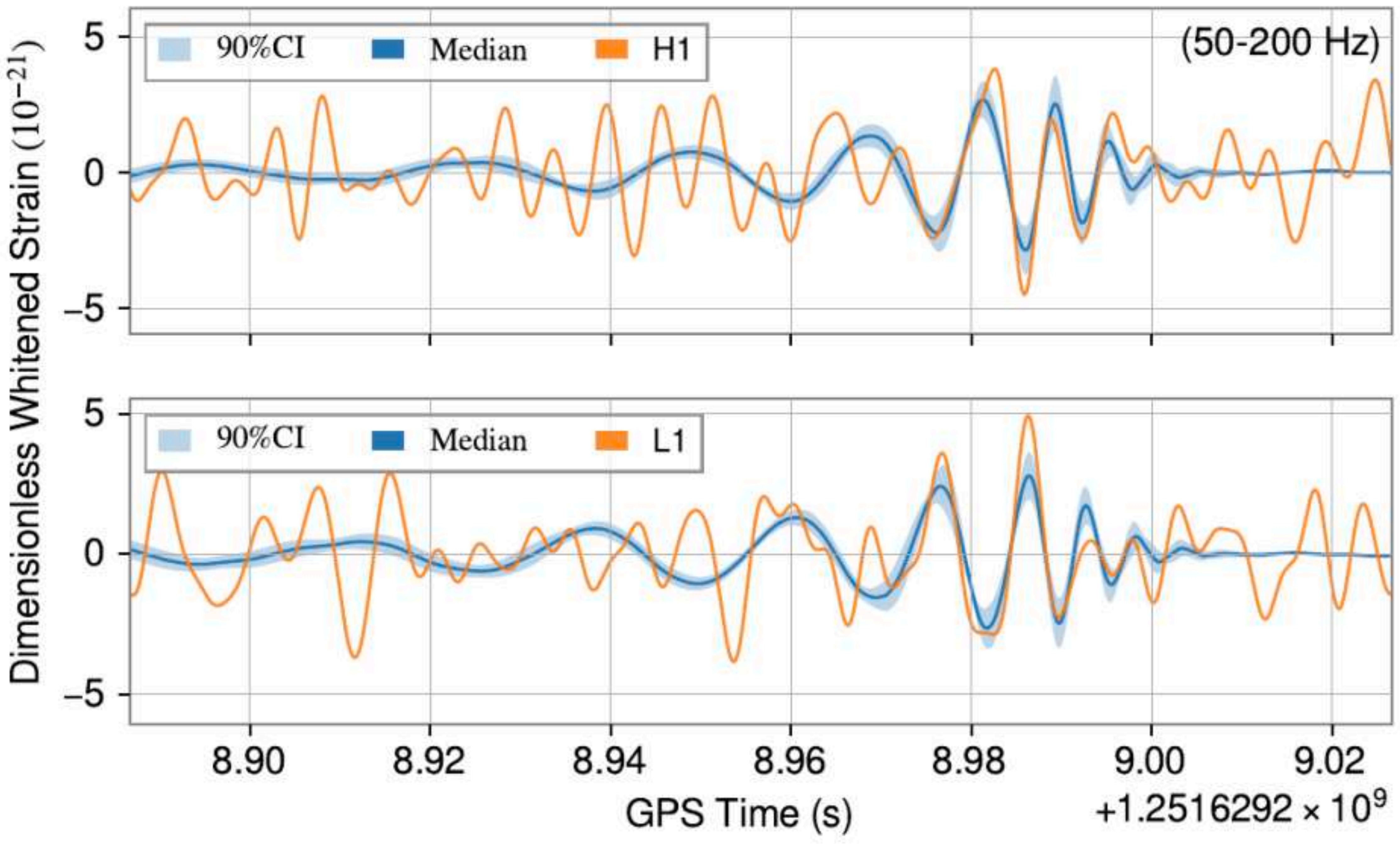


FIG. 38: Same as Fig. 35, but for the new event GW190904_104631.

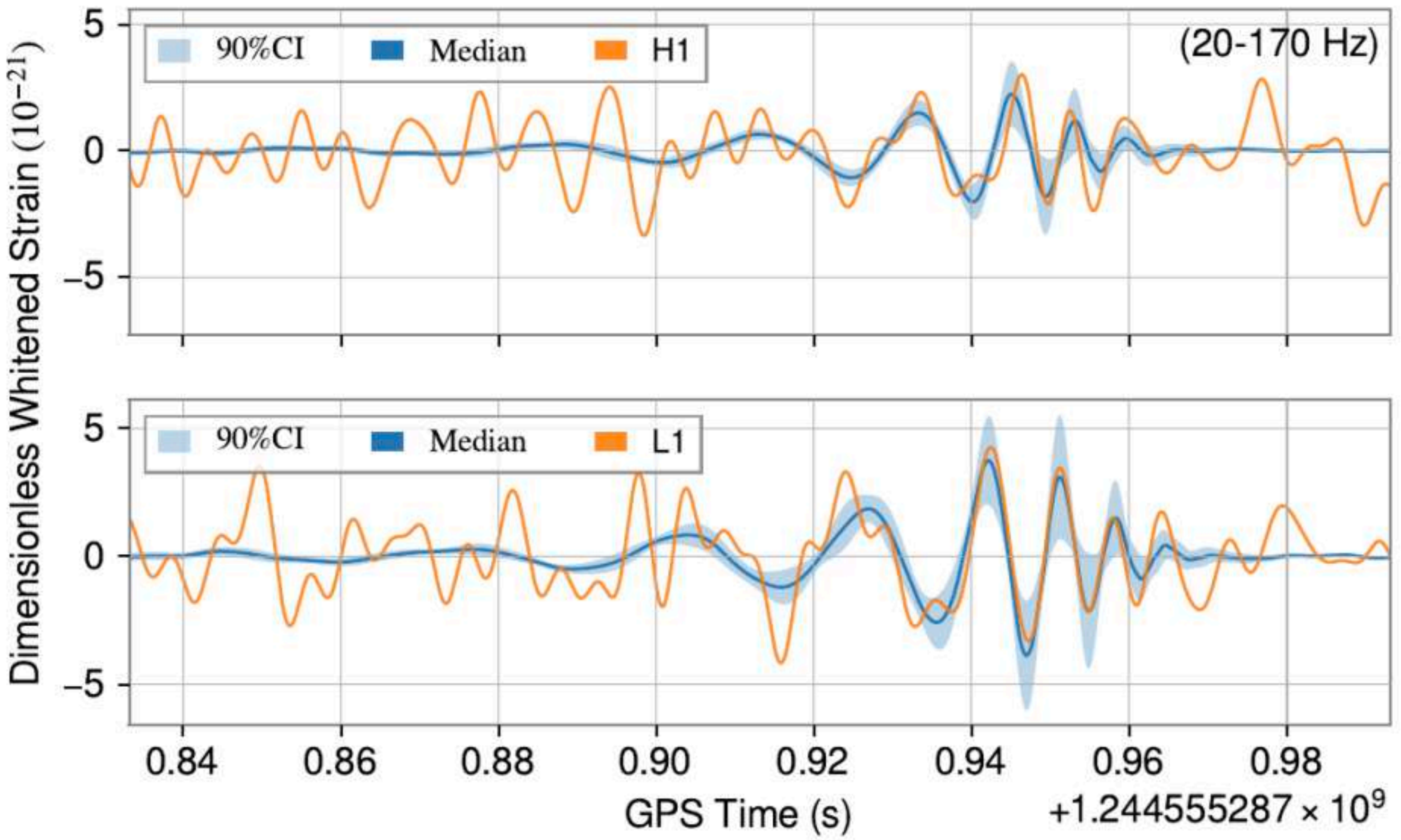


FIG. 36: Same as Fig. 35, but for the new event GW190614_134749.

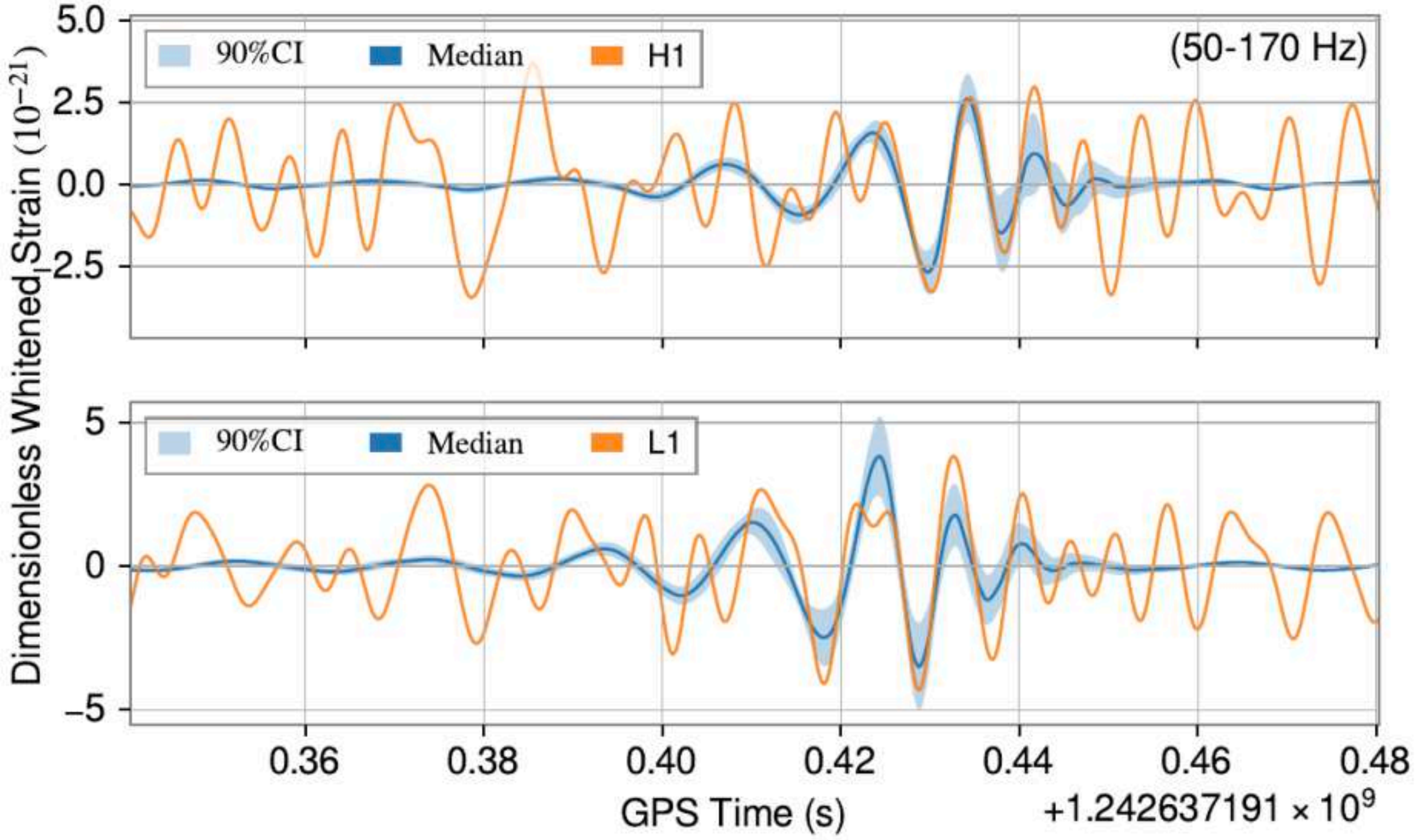


FIG. 39: Same as Fig. 35, but for the new event GW190523_085933.

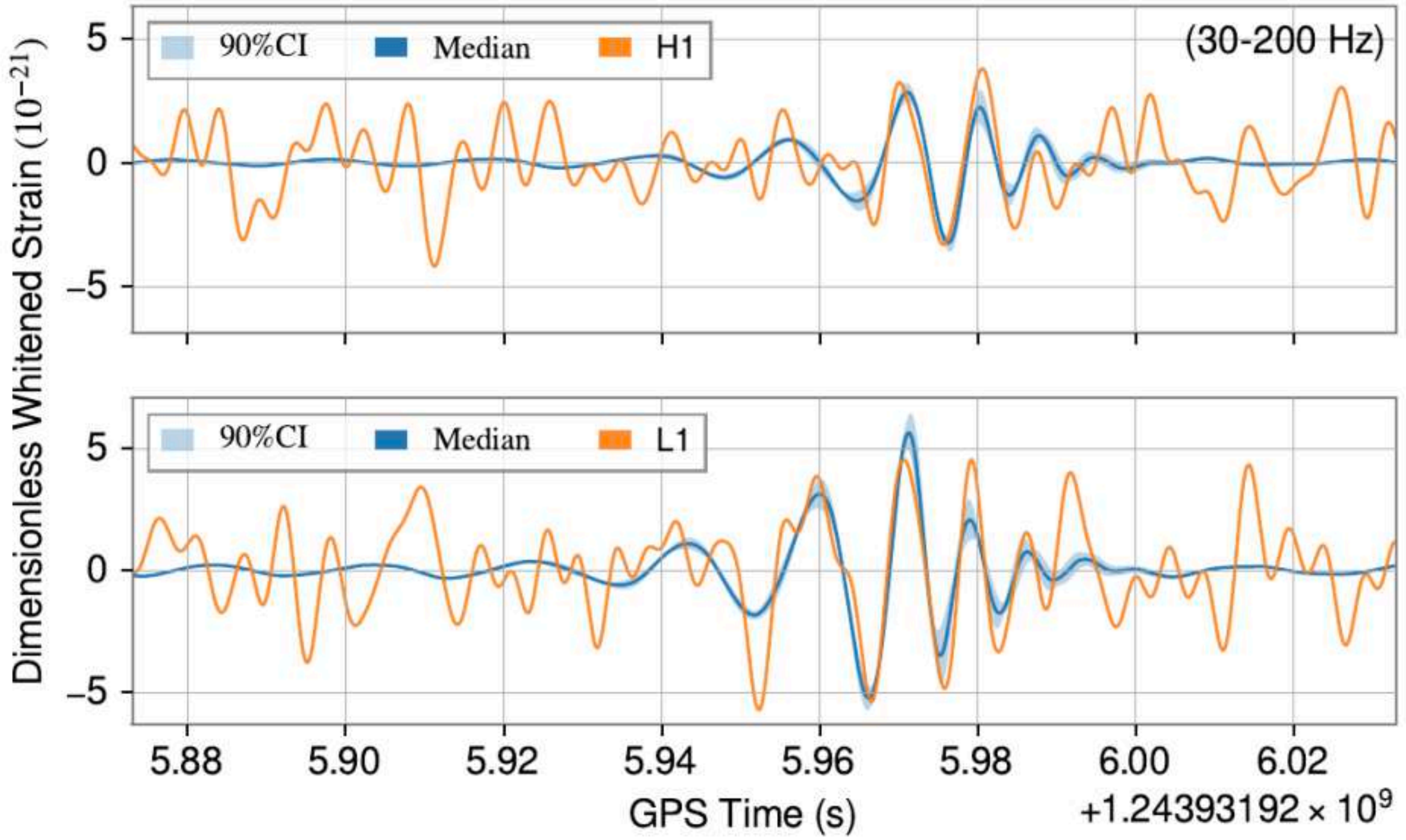


FIG. 37: Same as Fig. 35, but for the new event GW190607_083827.

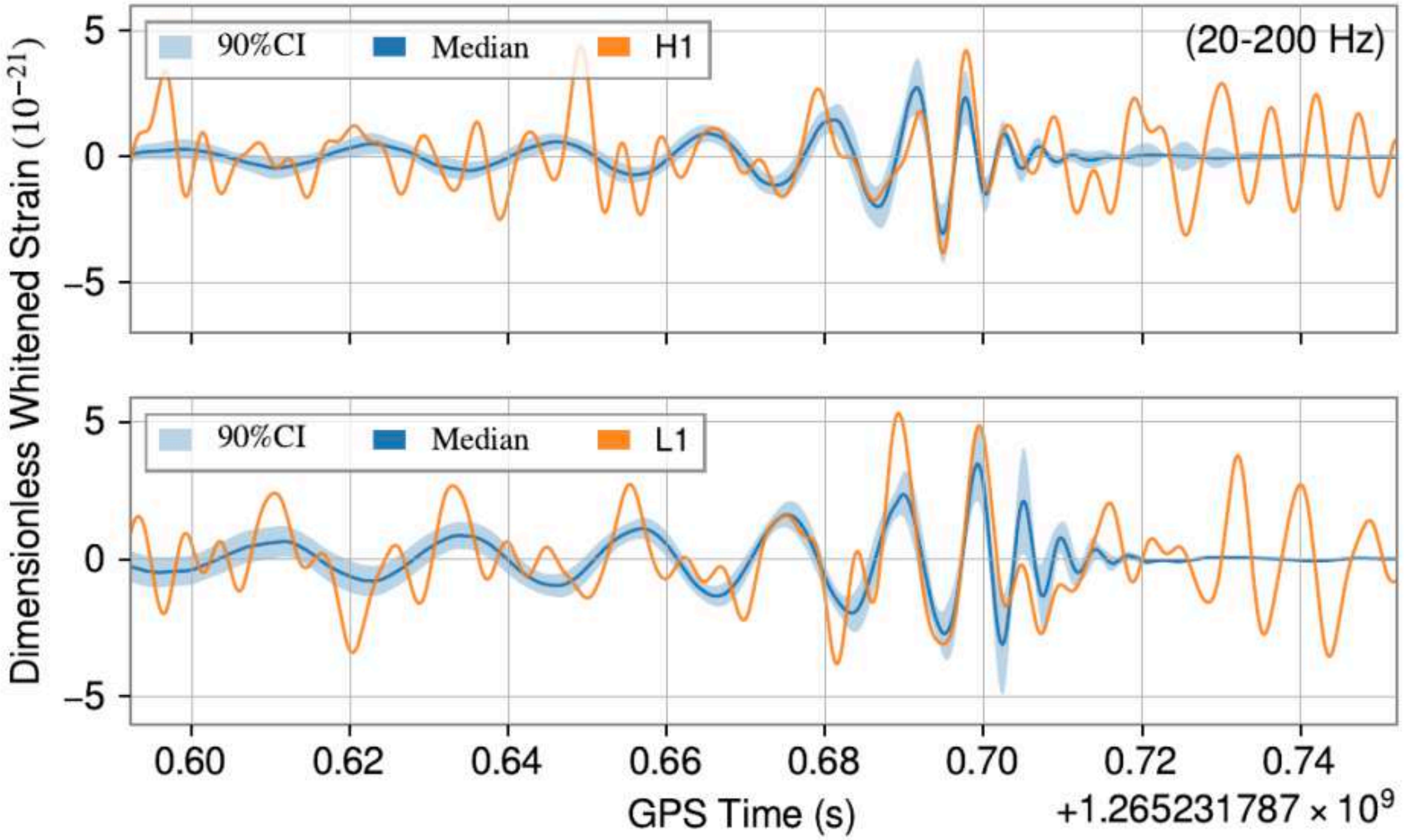


FIG. 40: Same as Fig. 35, but for the new event GW200208_211609.

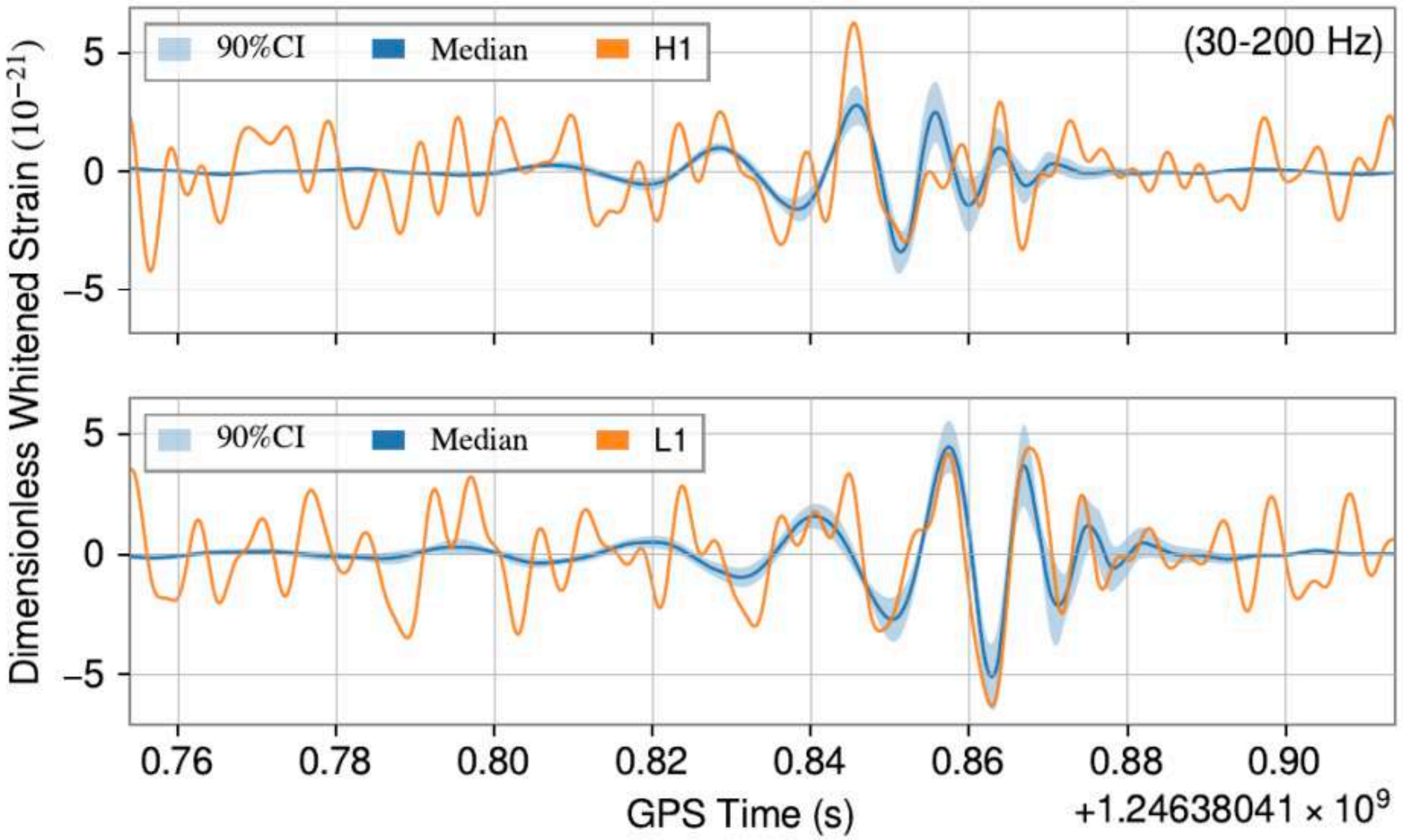


FIG. 41: Same as Fig. 35, but for the new event GW190705_164632.

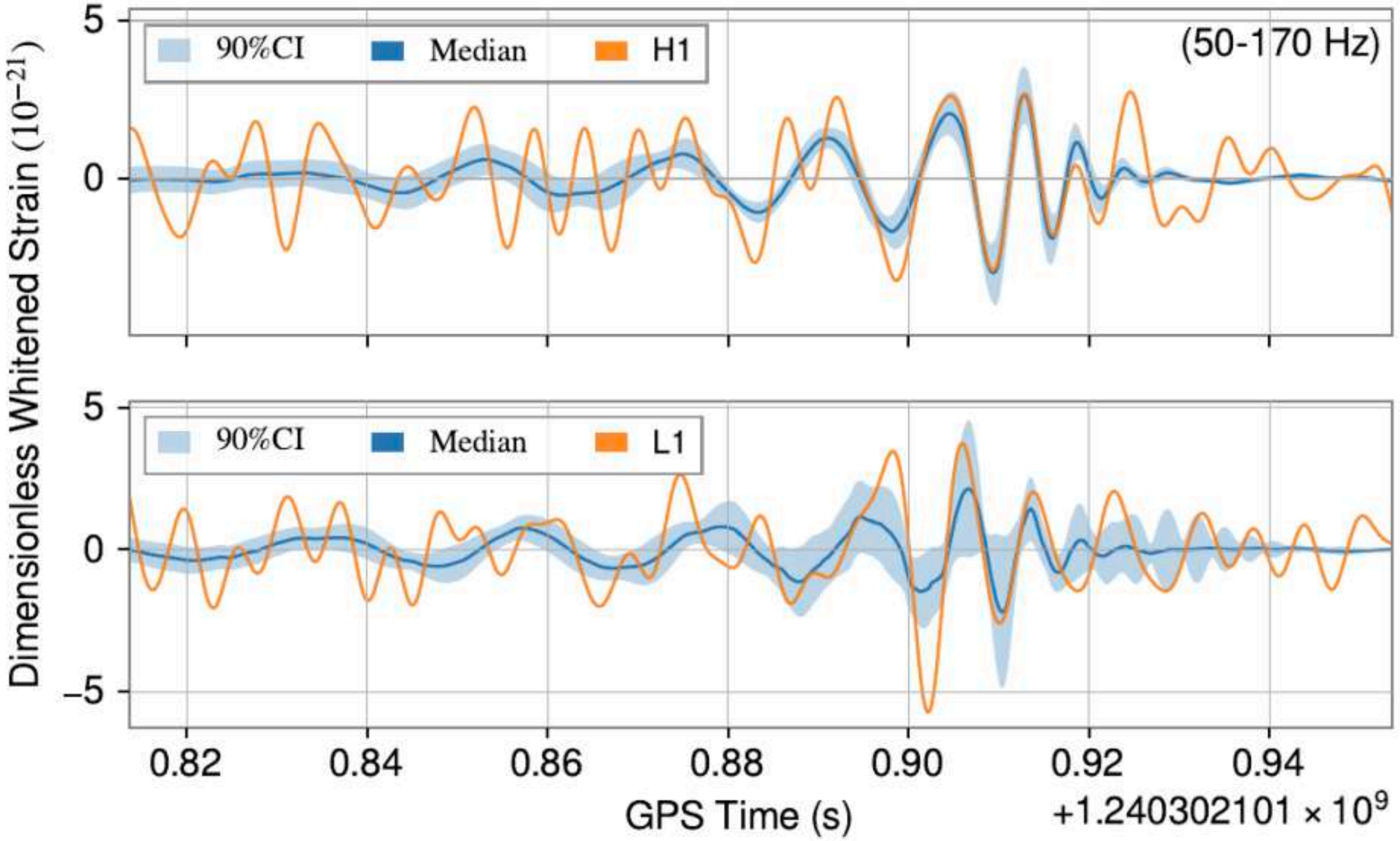
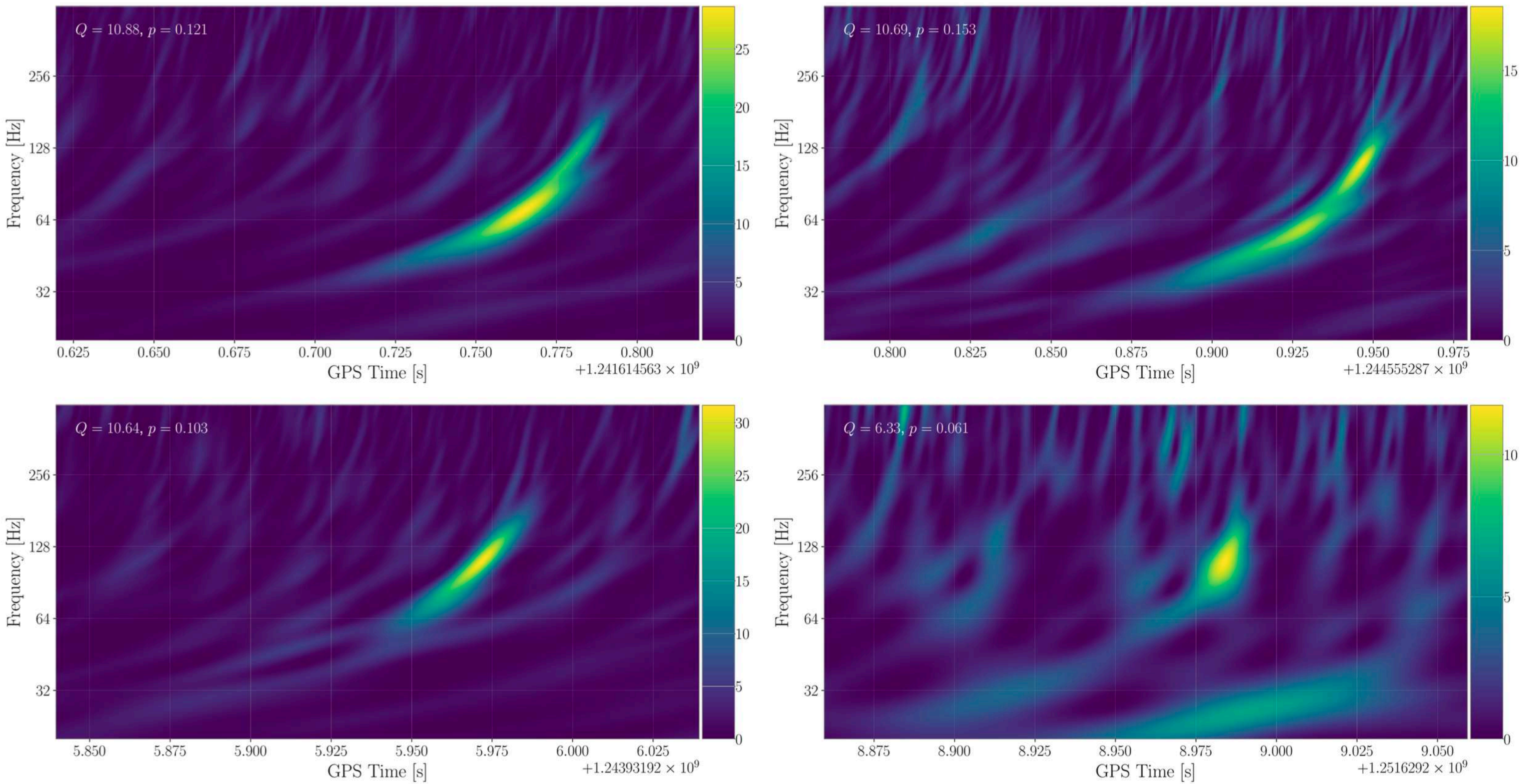
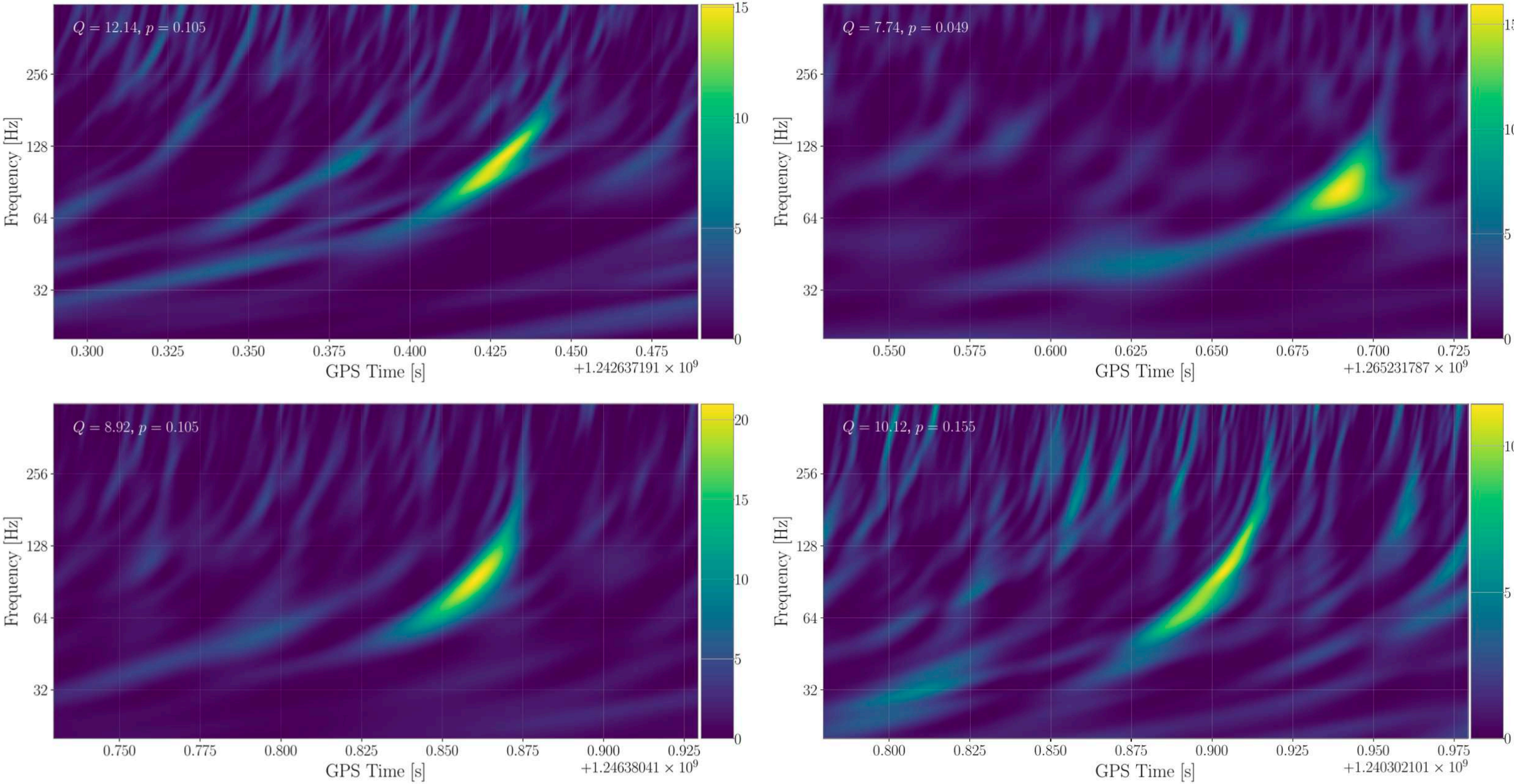


FIG. 42: Same as Fig. 35, but for the new event GW190426_082124.

ARESGW NEW CANDIDATE EVENTS



ARESGW NEW CANDIDATE EVENTS



DATASETS AND PARAMETER SPACE

Dataset 1

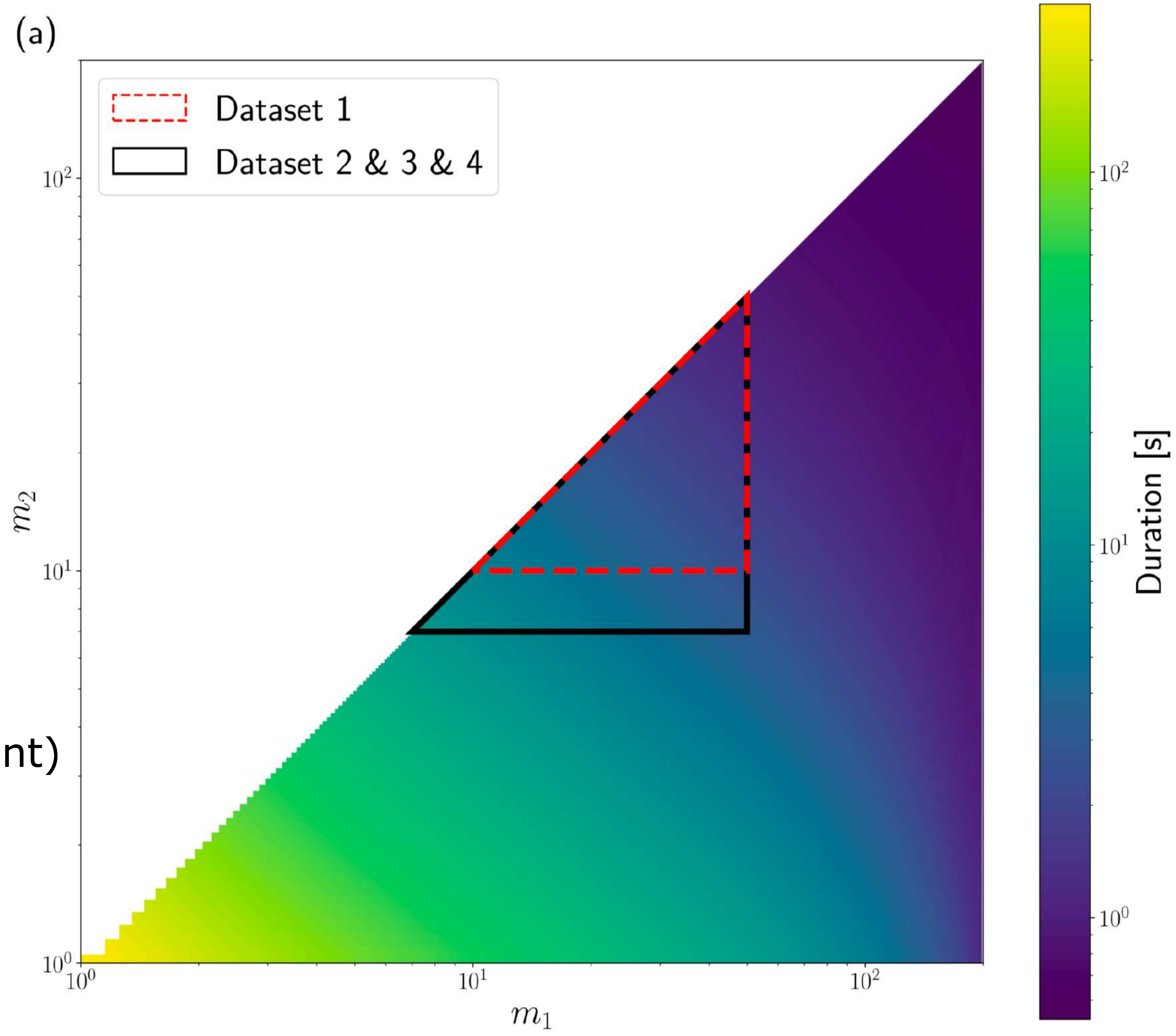
- Purely Gaussian noise (with aLIGOZeroDetHighPower PSD)
- no spin, no higher-order modes
- 10-50 Msun component masses (uniformly)

Dataset 2

- Gaussian noise (but with O3a PSD)
- aligned spin (-0.99,0.99), no higher-order modes
- 7-50 Msun (uniformly)

Dataset 3

- Gaussian noise (but with O3a PSD, random for each 2h segment)
- non-aligned spin (with isotropic distribution in magnitude)
- higher-order modes up to (4,-4) as in IMRPhenomXPHM
- 7-50 Msun (uniformly)



Dataset 4

- as dataset 3, but with real noise sampled from O3a parts (excl. any GWTC-2 detections)

Evaluation:

1. Each algorithm is evaluated on the background dataset and returns a ranking statistic for each (false) positive event.
2. FAR at given ranking statistic *R* = (# of background events with ranking statistic > R) / duration *T*

$$\mathrm{FAR}(\mathcal{R}) = \frac{N_{\mathrm{FP},\mathcal{R}}}{T}$$

1. For uniform (*in volume*) injections up to a distance d_max, the sensitive volume is approximately

$$V(\mathrm{FAR}) \approx V(d_{\max})\frac{N_{I,\mathrm{FAR}}}{N_I}$$

where N_I,FAR = found injections at given FAR and N_I = total # of found injections (within +/- Δt each).

However, the injections are not uniform in volume, but are sampled over the *chirp distance*, instead:

$$d_c = d\left(\frac{\mathcal{M}_{c,0}}{\mathcal{M}_c}\right)^{5/6}$$

with $\mathcal{M}_{c,0} = 1.4/2^{1/5}\mathrm{M}_\odot$. To account for that:

$$V(\mathrm{FAR}) \approx \frac{V(d_{\max})}{N_I}\sum_{i=1}^{N_{I,\mathrm{FAR}}}\left(\frac{\mathcal{M}_{c,i}}{\mathcal{M}_{c,\max}}\right)^{5/2}$$