LEVERAGING MACHINE LEARNING FOR NEUTRON STAR PHYSICS
From Gravitational Waveform Modeling to Probing High-Density Nuclear Matter

NIKOLAOS STERGIOULAS

DEPARTMENT OF PHYSICS

ARISTOTLE UNIVERSITY OF THESSALONIKI

Utrecht, June 26, 2025

INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)

In operation

Under construction

LHO

LLO

GEO

Virgo

KAGRA

LIGO India

INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)

In operation

Under construction

LHO

LLO

inspiral

GEO

Virgo

KAGRA

LIGO India

INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)

In operation

Under construction

LHO

LLO

LIGO India

KAGRA

A

B

inspiral

post-merger

INTERNATIONAL GRAVITATIONAL-WAVE OBSERVATORY NETWORK (IGWN)

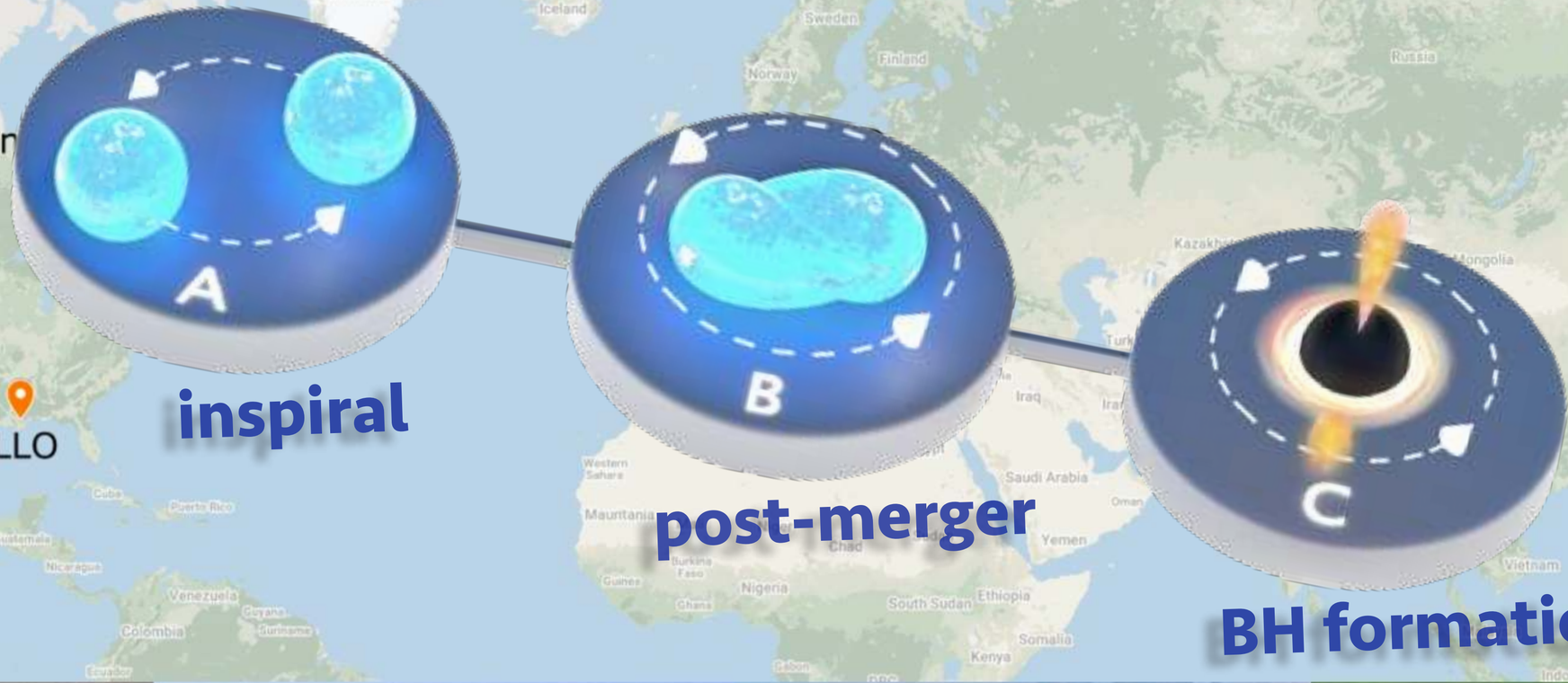
In operation

Under construction

LHO

LLO

KAGRA



inspiral

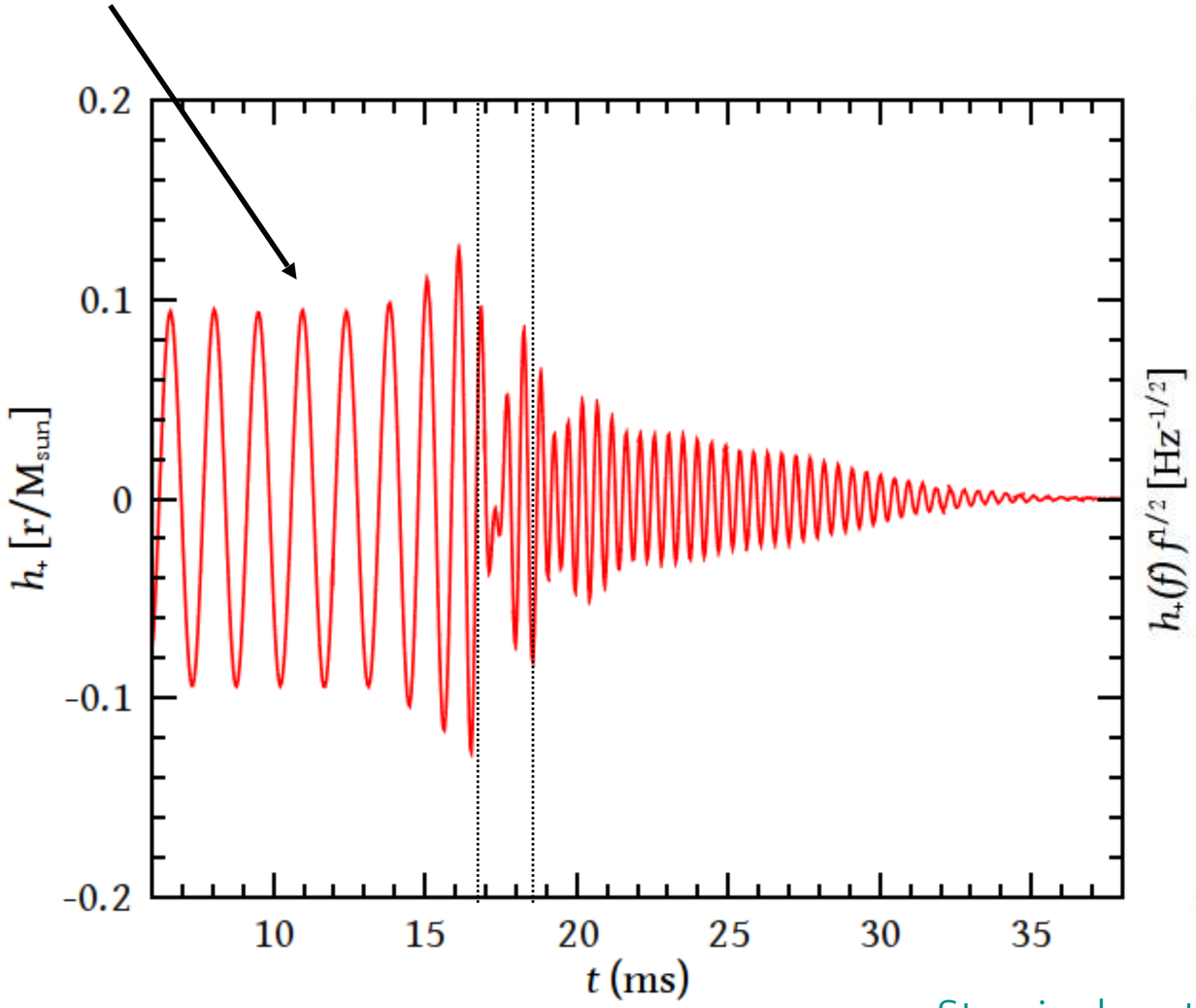
post-merger

BH formation

POST-MERGER PHASE IN BNS MERGERS

Time domain: three distinct phases of GW the signal:

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

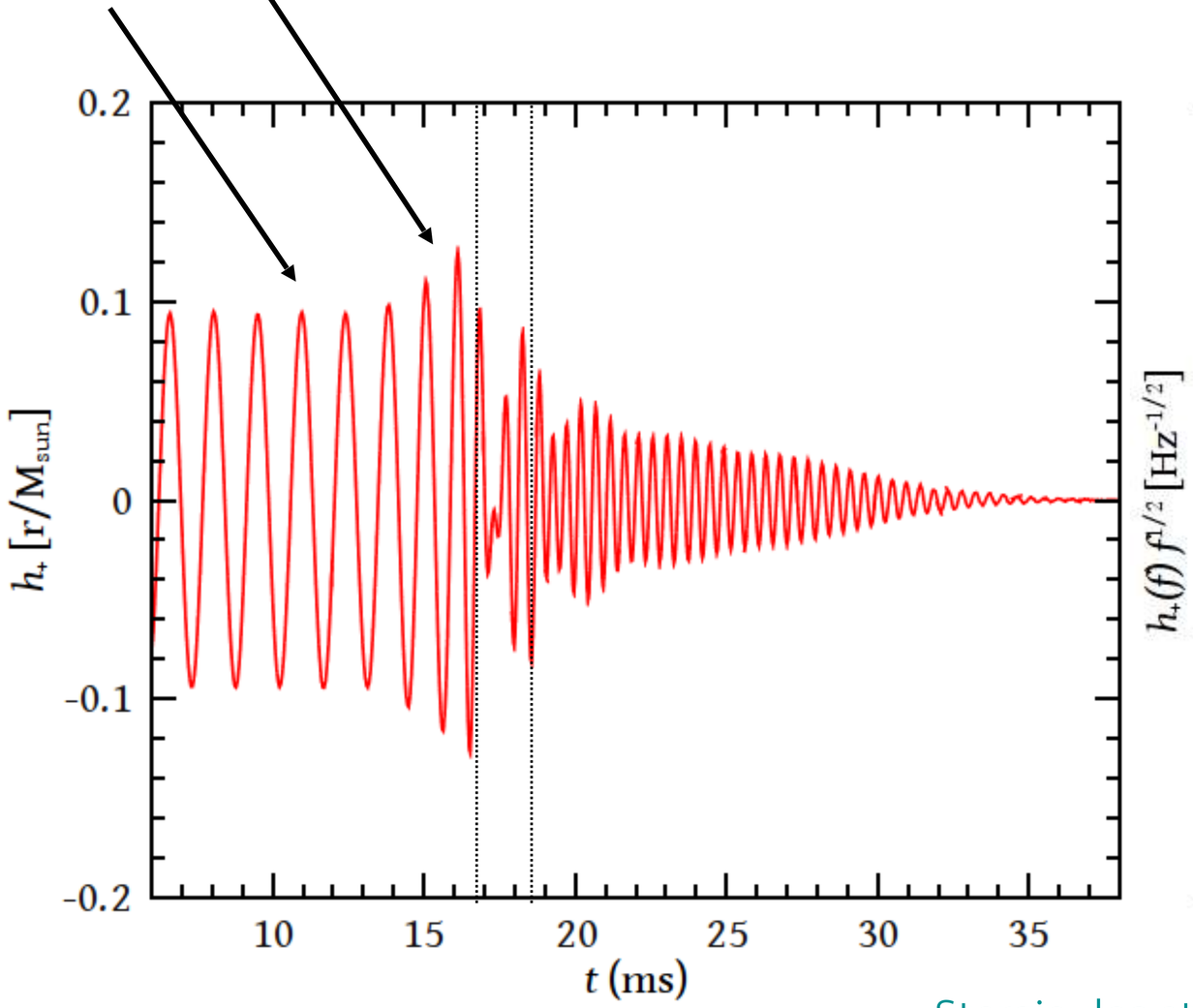


Stergioulas et al. (2011)

POST-MERGER PHASE IN BNS MERGERS

Time domain: three distinct phases of GW the signal:

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

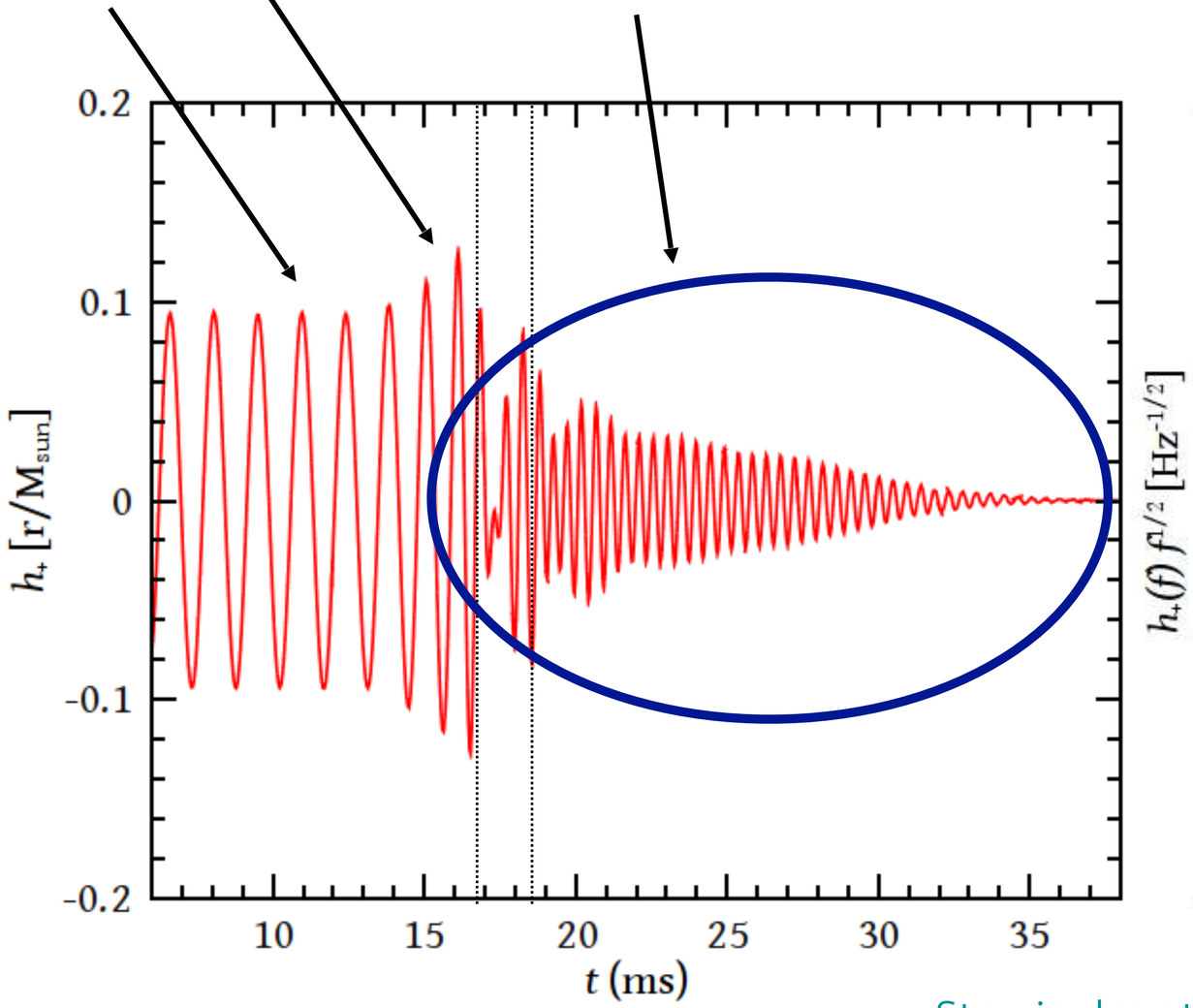


Stergioulas et al. (2011)

POST-MERGER PHASE IN BNS MERGERS

Time domain: three distinct phases of GW the signal:

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

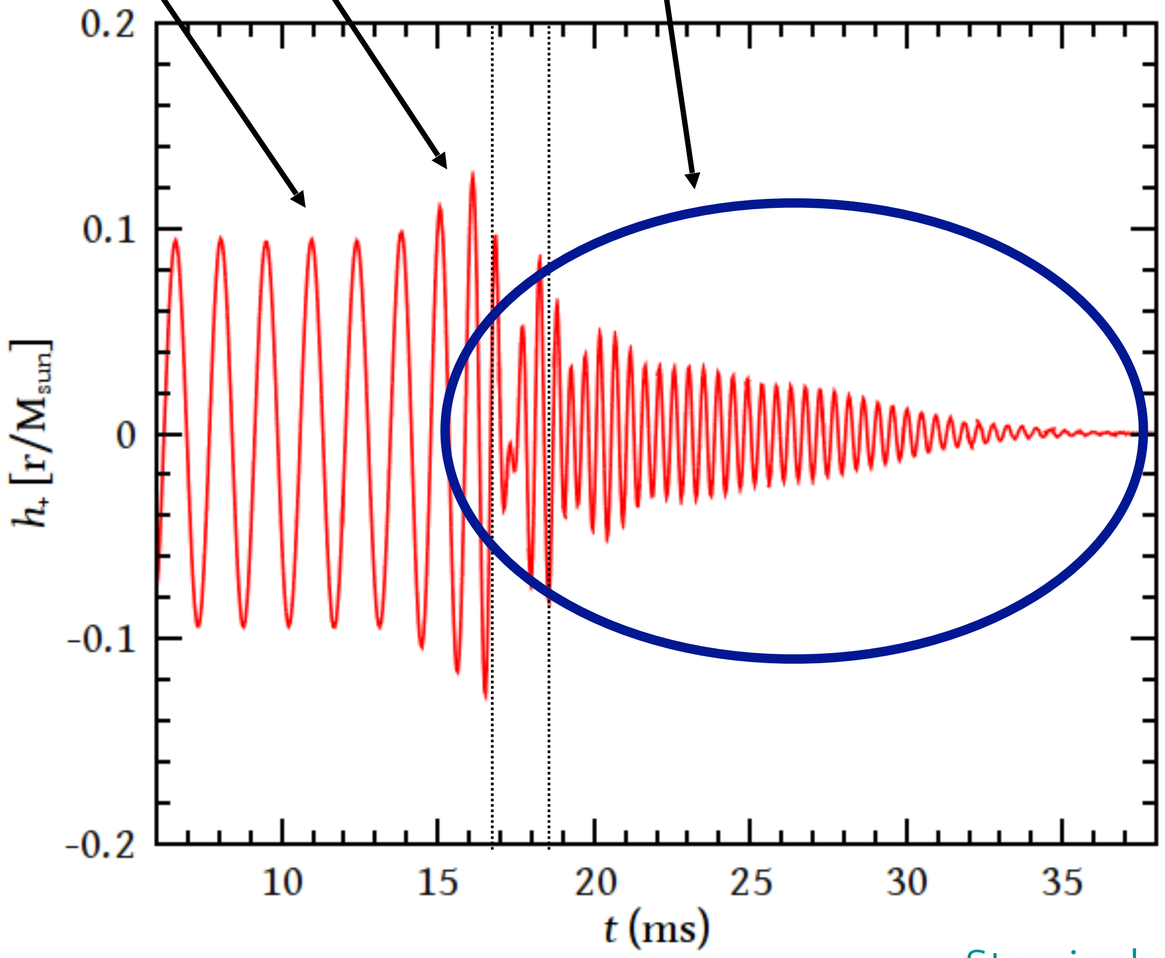


Stergioulas et al. (2011)

POST-MERGER PHASE IN BNS MERGERS

Time domain: three distinct phases of GW the signal:

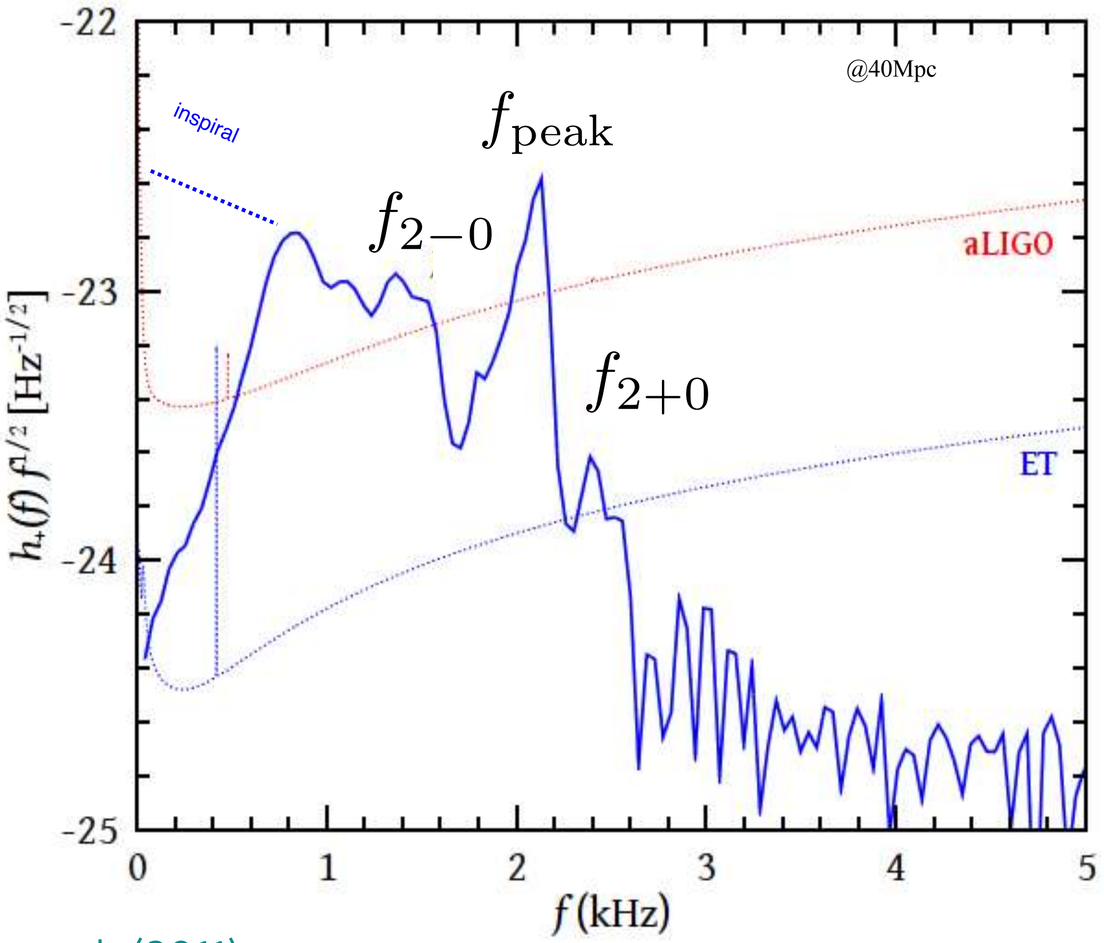
inspiral, *merger* and *post-merger oscillations*.



Frequency domain:

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

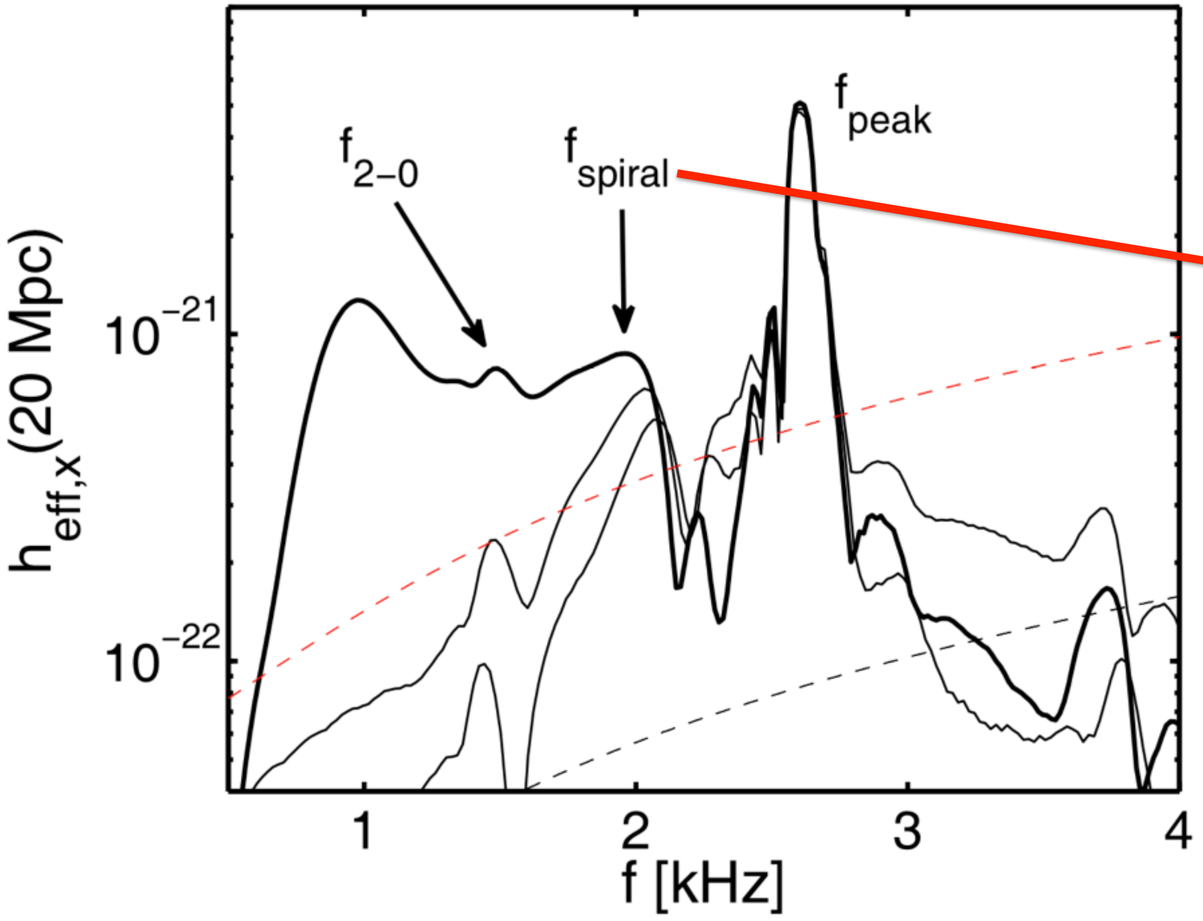
f_{2-0}, f_{2+0} : *nonlinear combination tones*

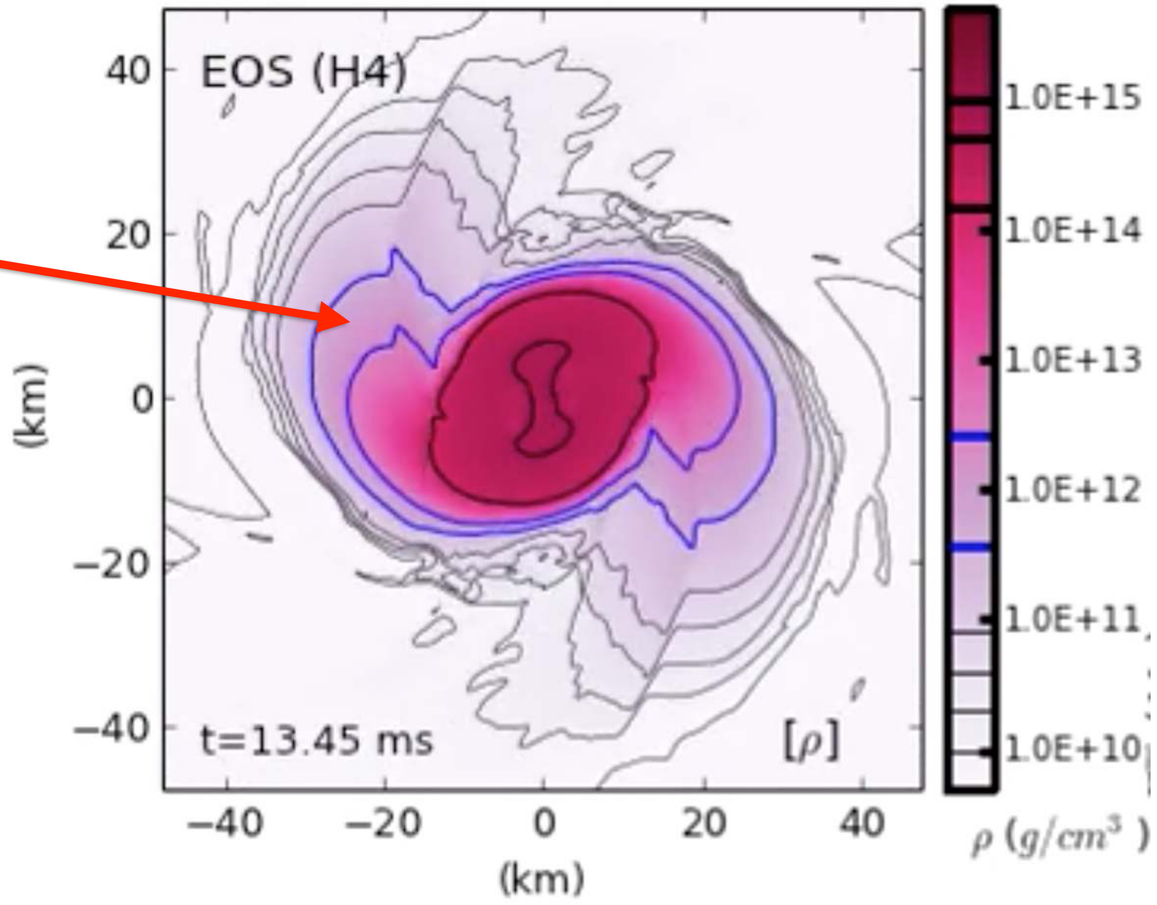


Stergioulas et al. (2011)

POST-MERGER PHASE IN BNS MERGERS

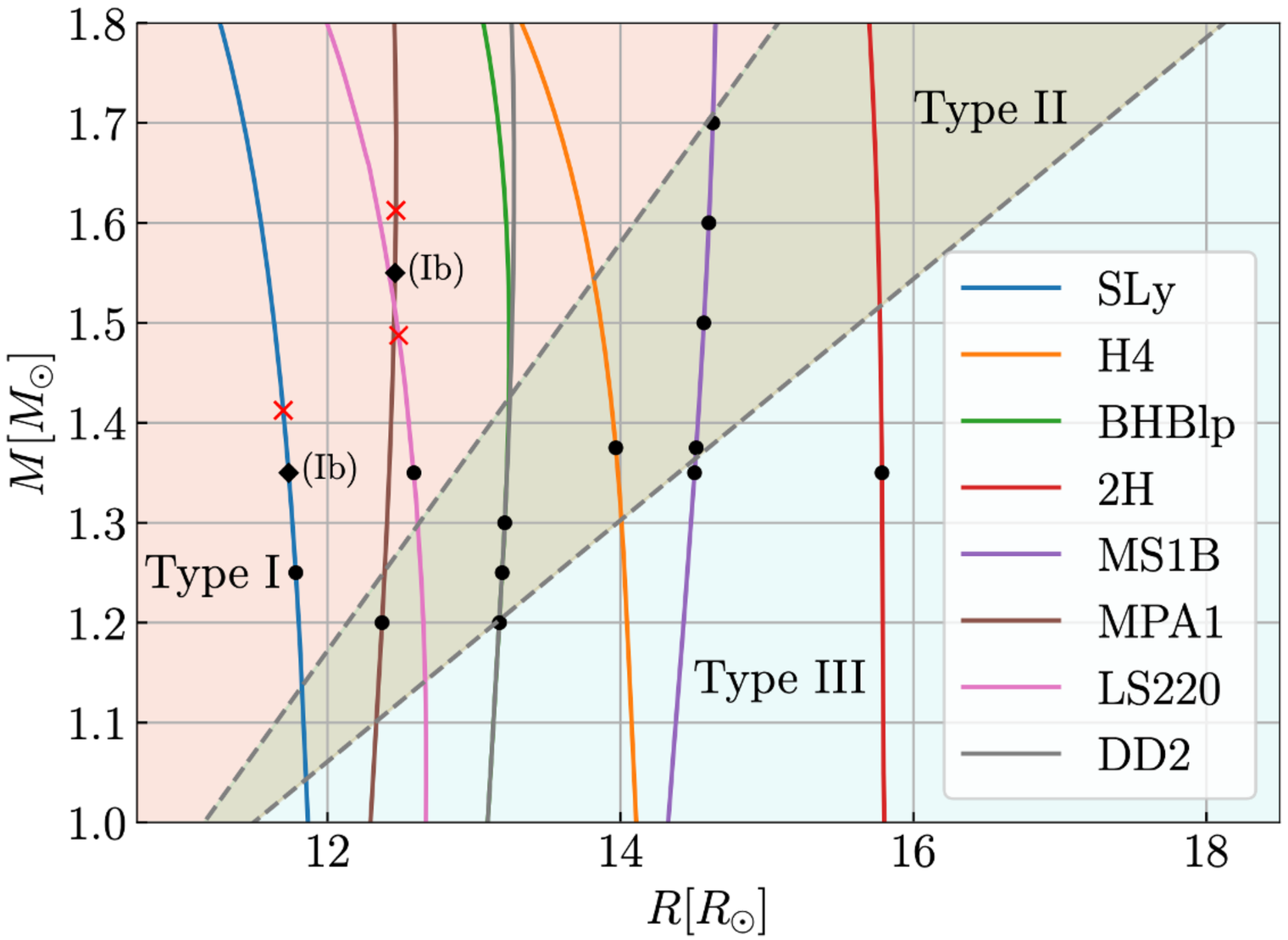
Orbiting spiral arms also lead to a distinct frequency $f_{\rm spiral}$





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:

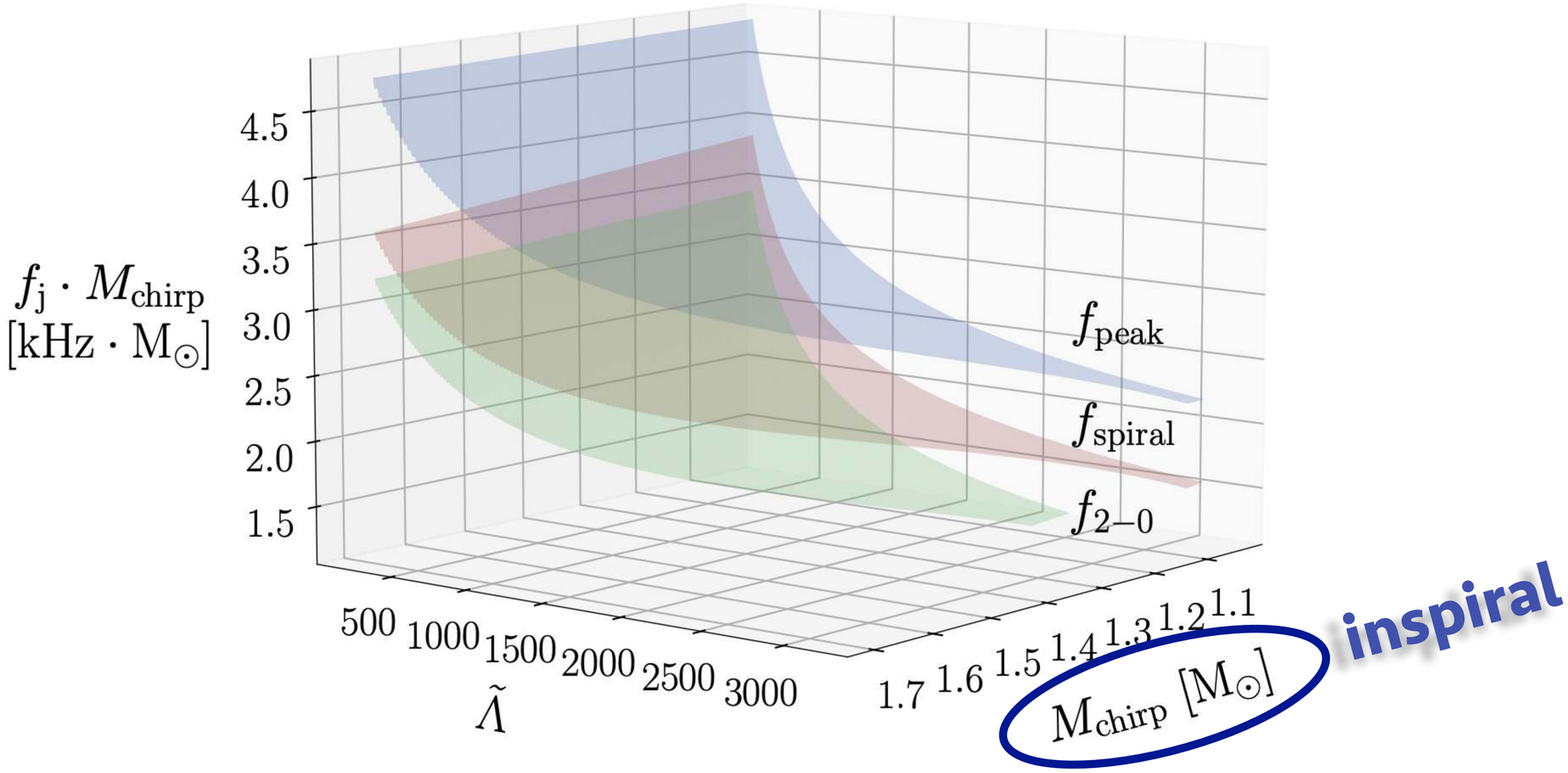
(close to M_{thres})

EMPIRICAL RELATIONS OF POST-MERGER FREQUENCIES

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

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

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



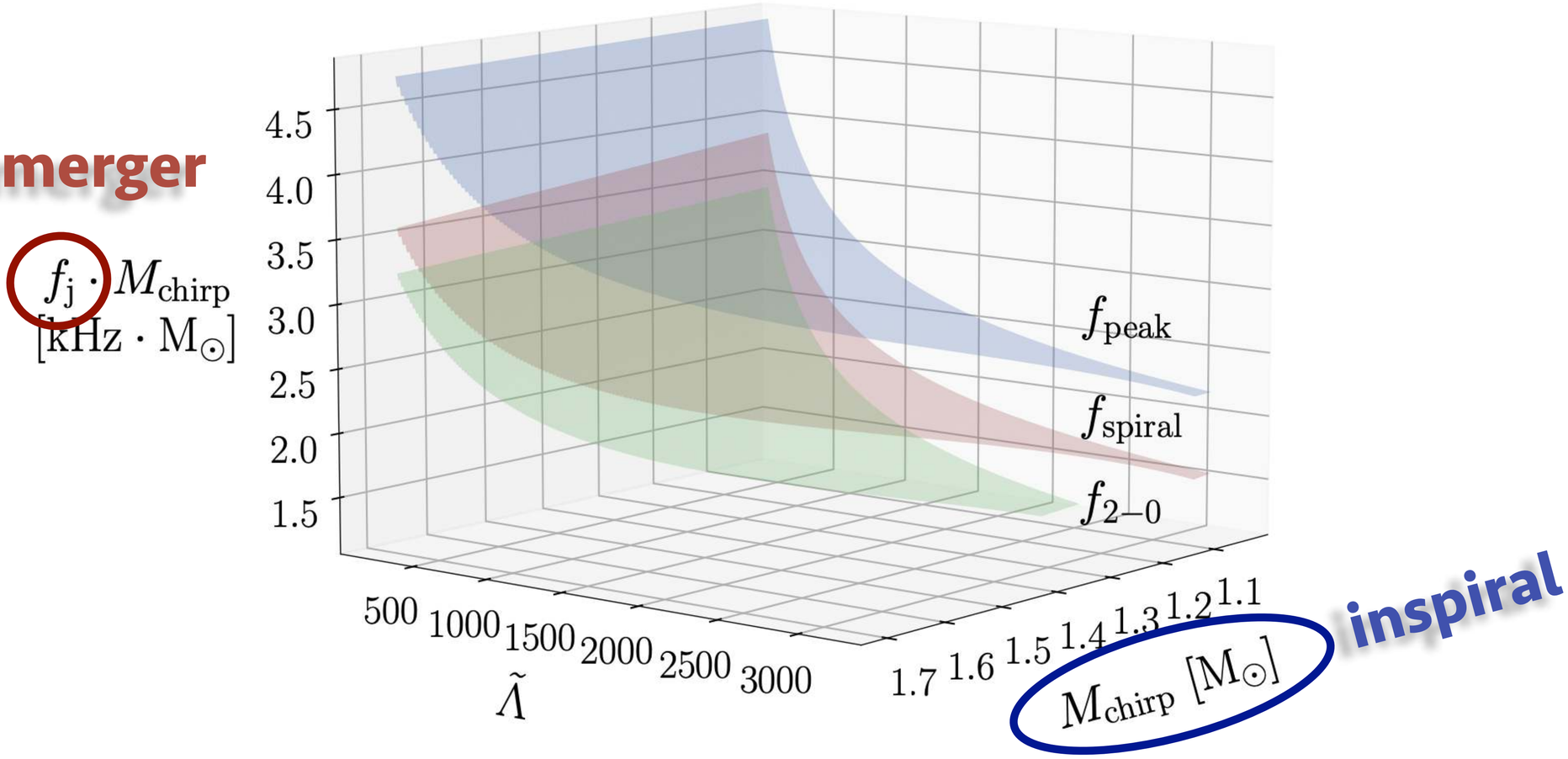
EMPIRICAL RELATIONS OF POST-MERGER FREQUENCIES

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

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

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

post-merger



inspiral

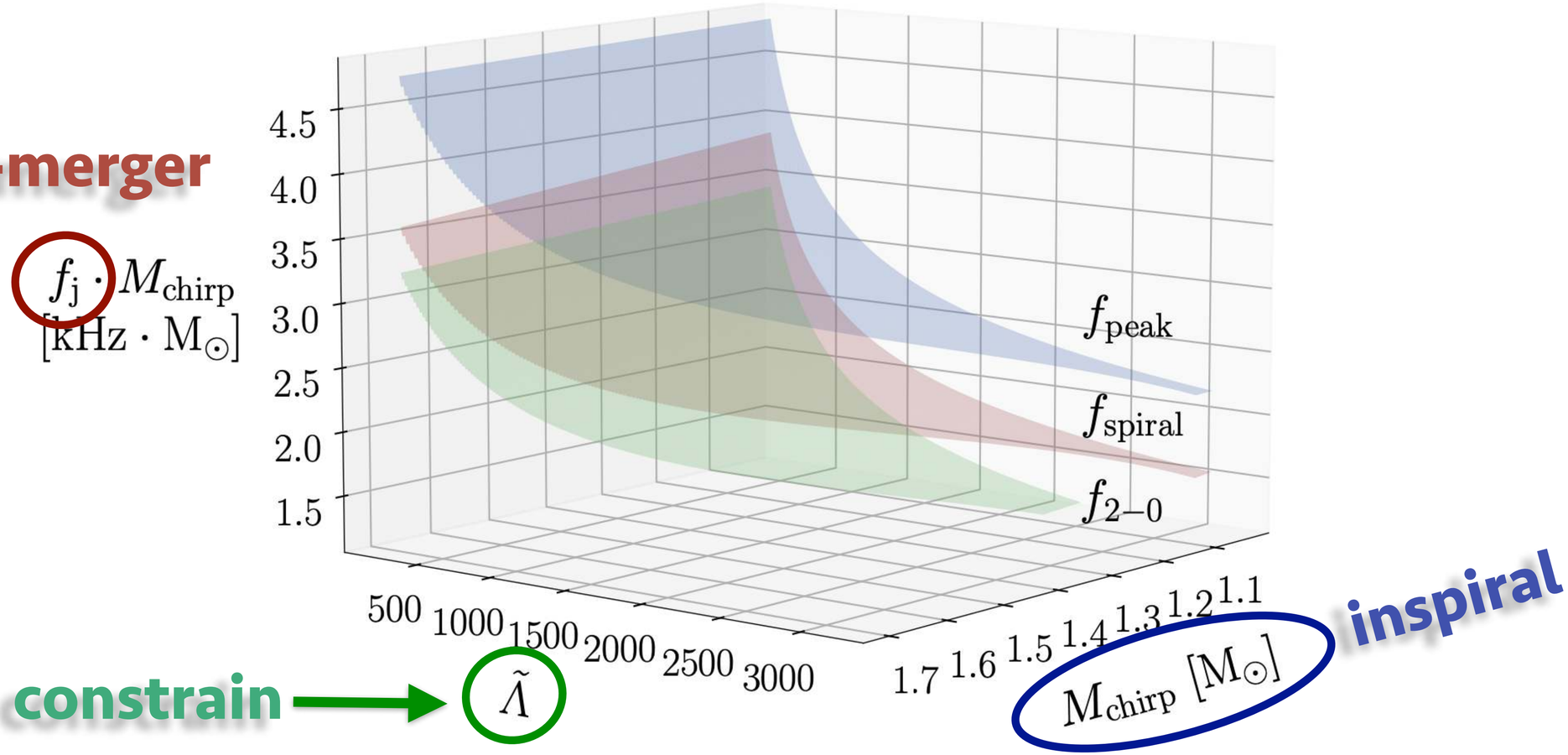
EMPIRICAL RELATIONS OF POST-MERGER FREQUENCIES

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

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

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

post-merger



constrain →

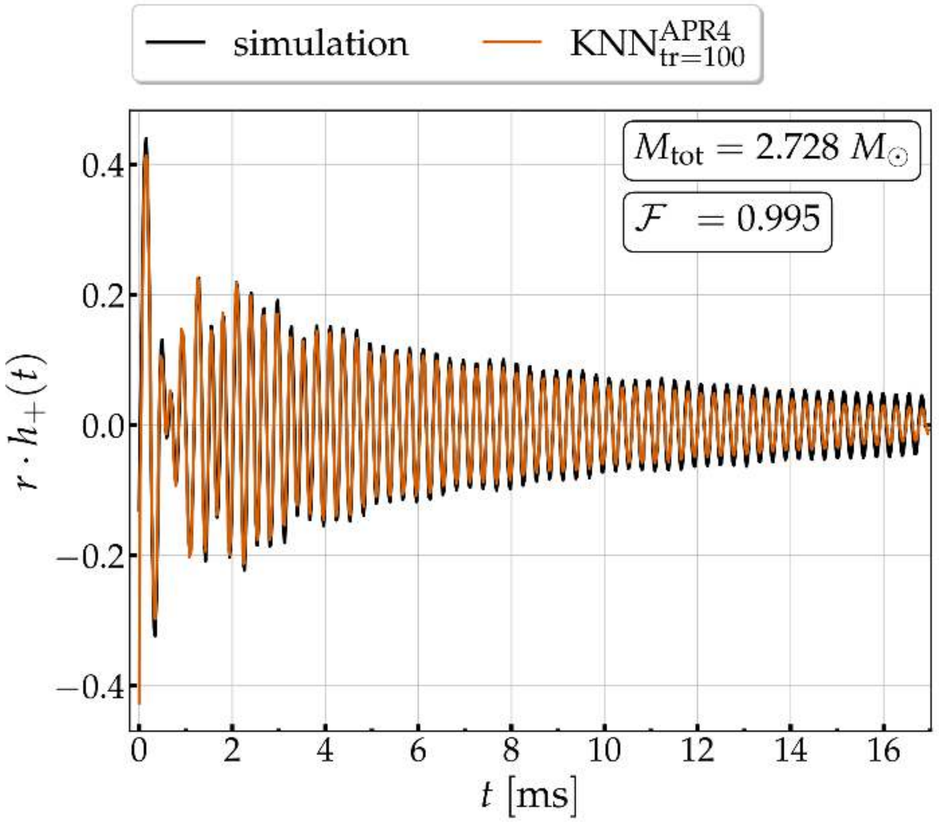
inspiral

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:



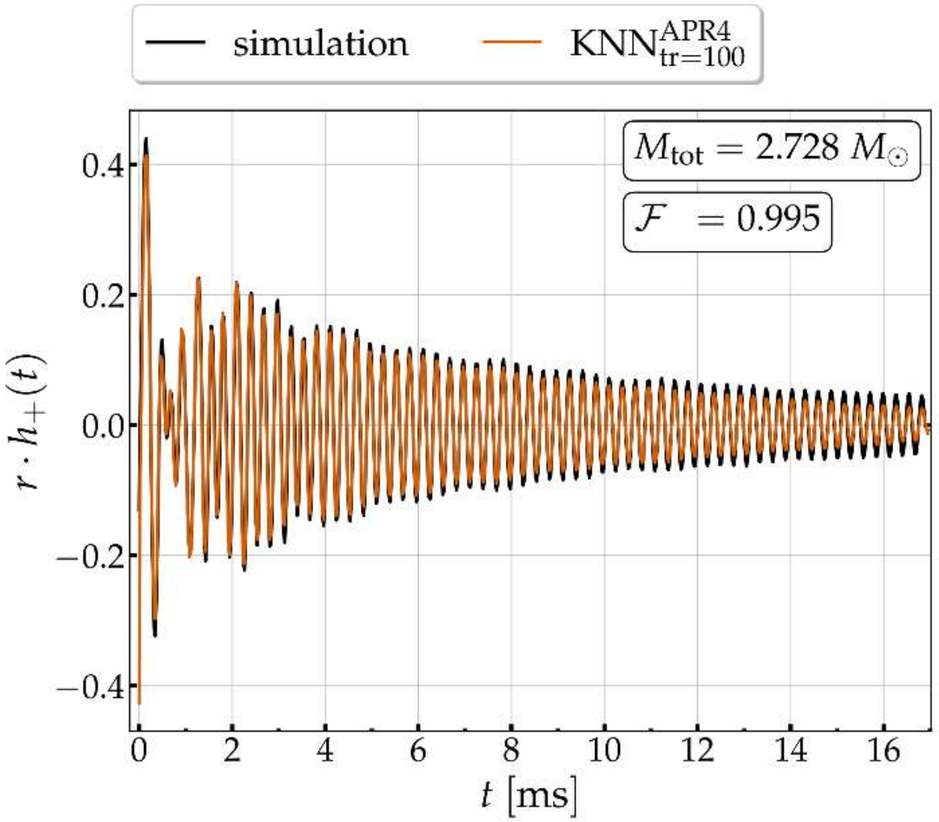
Soultanis et al. (2025)

APPLICATION OF MACHINE LEARNING TO THE POST-MERGER PHASE

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

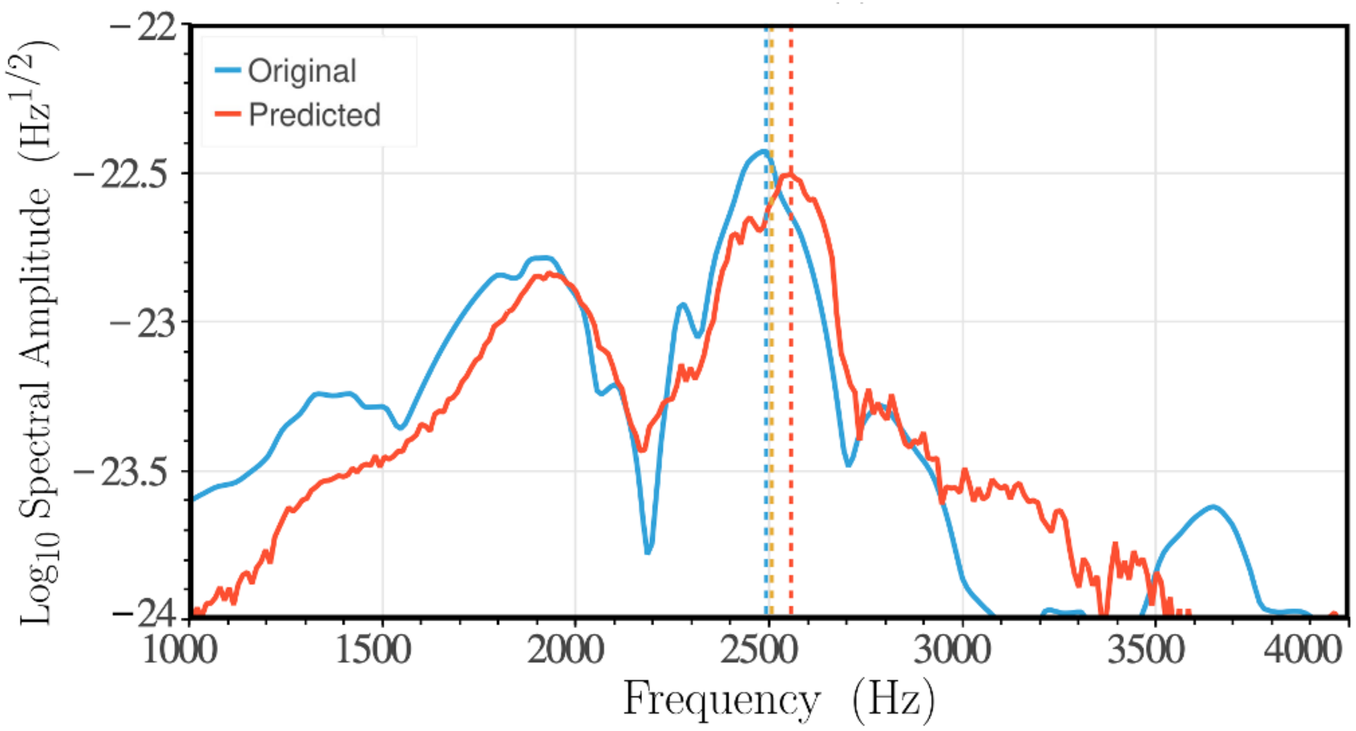
Solution: Construct surrogate model of post-merger GWs as function of e.g. M, q, EOS(Λ)

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)

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}.

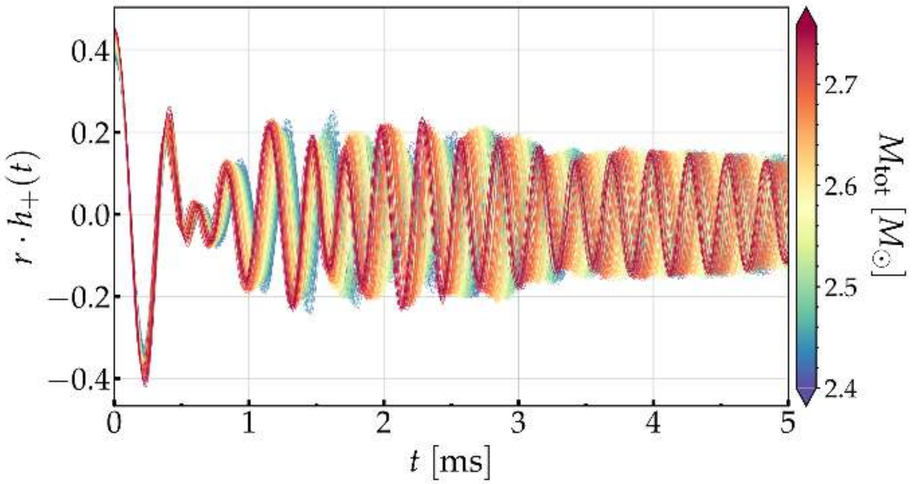
Choose specific EOS (APR4; SFHX).

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}.

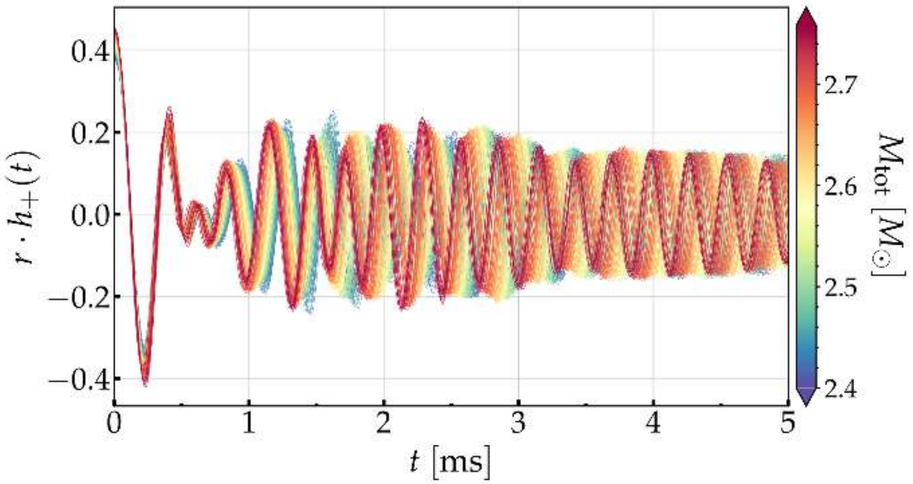
Choose specific EOS (APR4; SFHX).

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)

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:

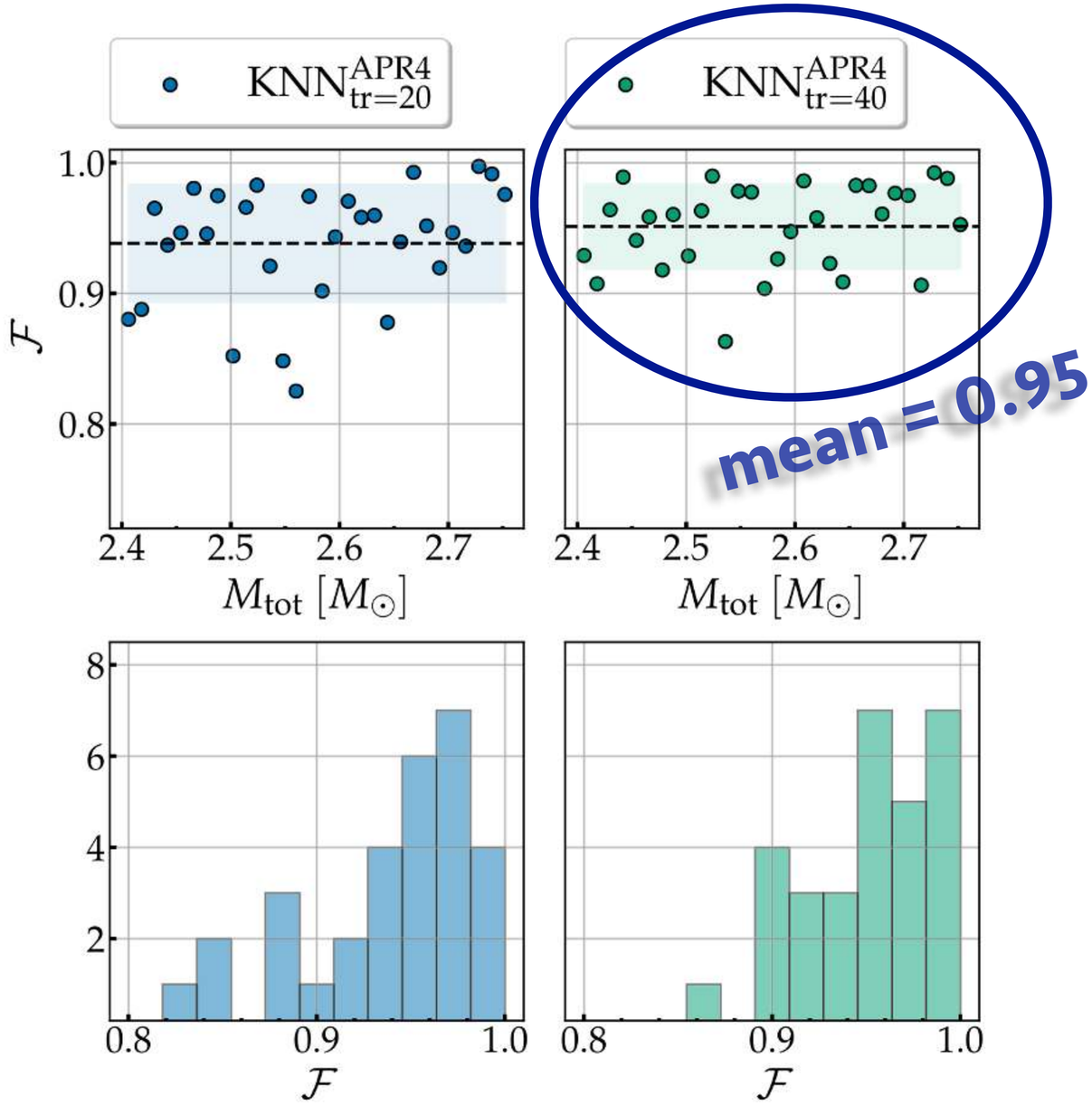
$$\langle h_1(t), h_2(t) \rangle \equiv 4\,\mathrm{Re} \int_0^\infty df \frac{\tilde{h}_1(f) \cdot \tilde{h}_2^*(f)}{S_h(f)}$$

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)

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

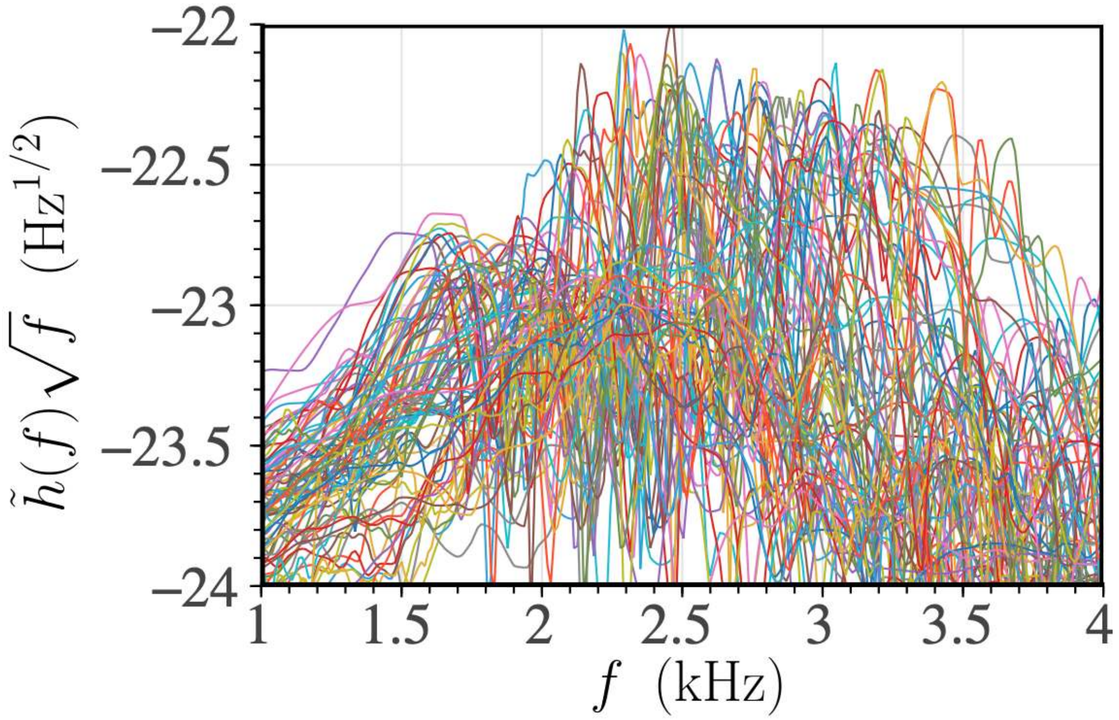


Soultanis et al. (2025)

ANN REGRESSION IN THE FREQUENCY DOMAIN

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

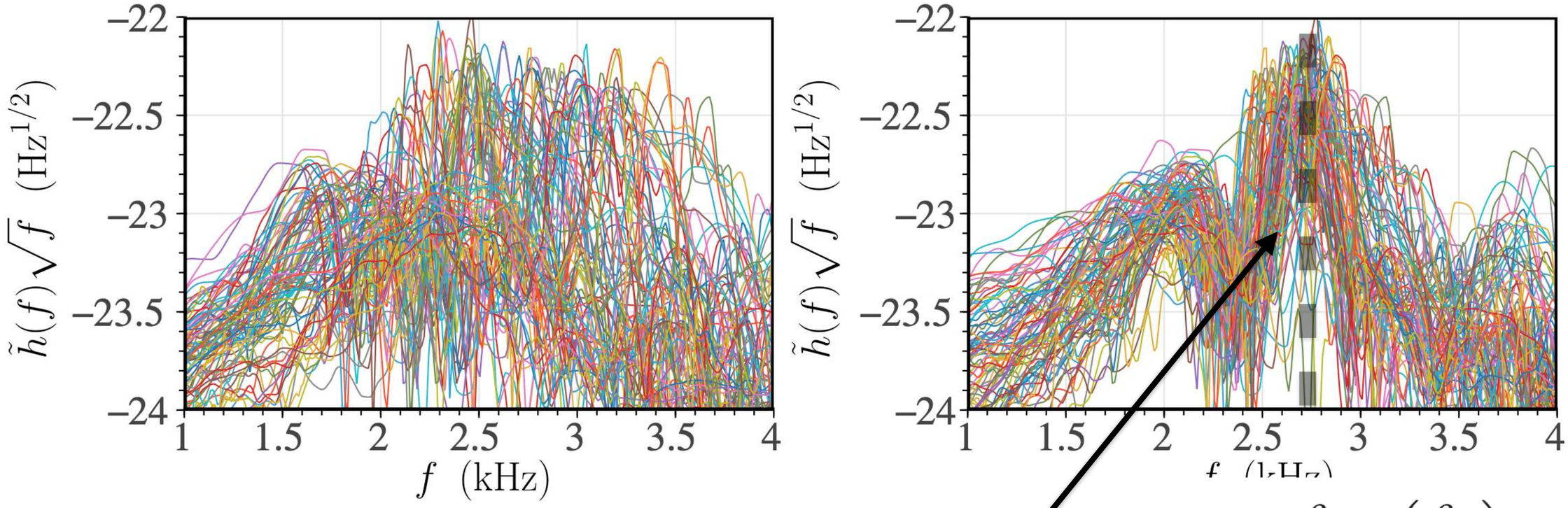
Surrogate model now depends on both mass and tidal deformability.



ANN REGRESSION IN THE FREQUENCY DOMAIN

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

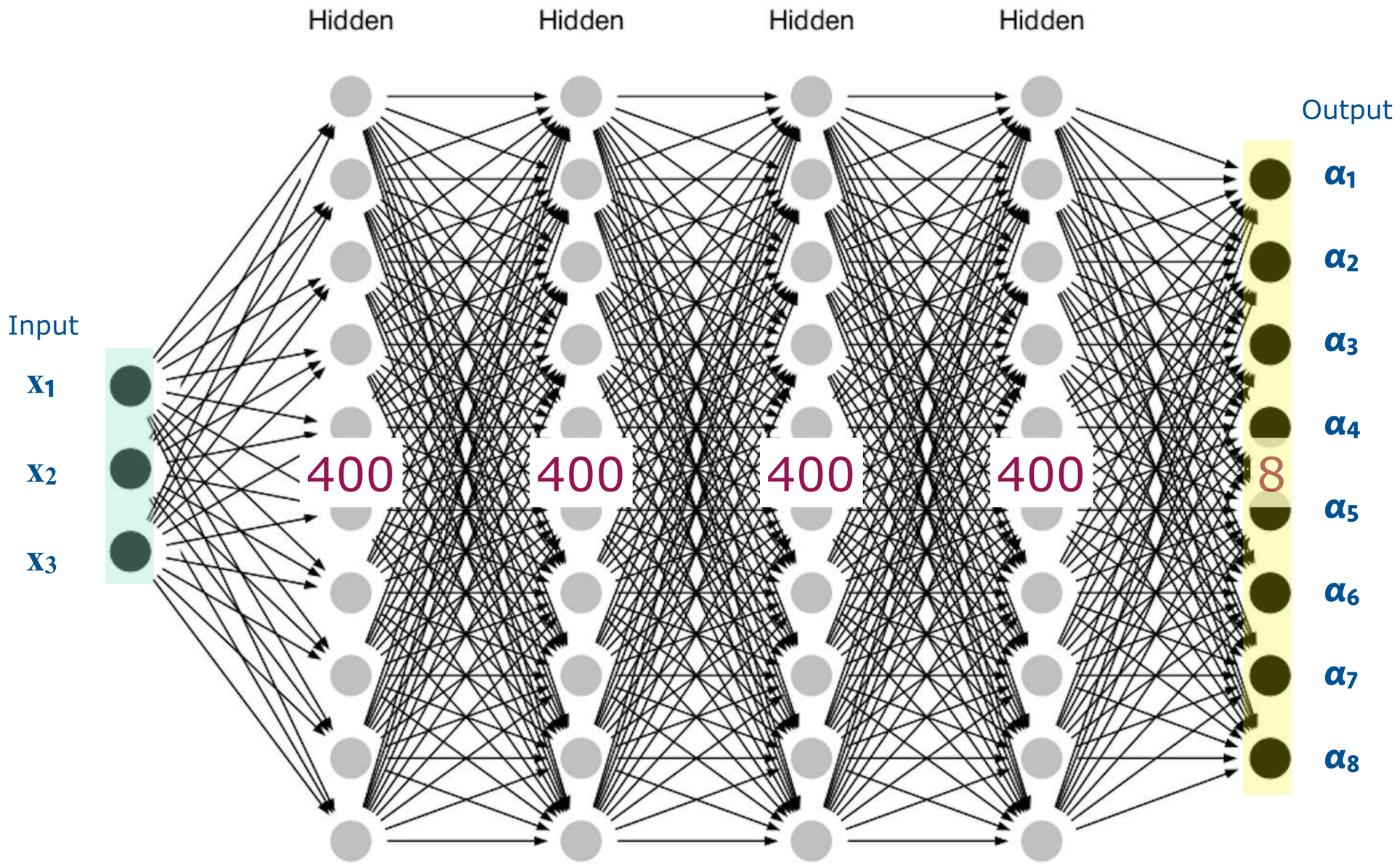
Surrogate model now depends on both mass and tidal deformability.



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

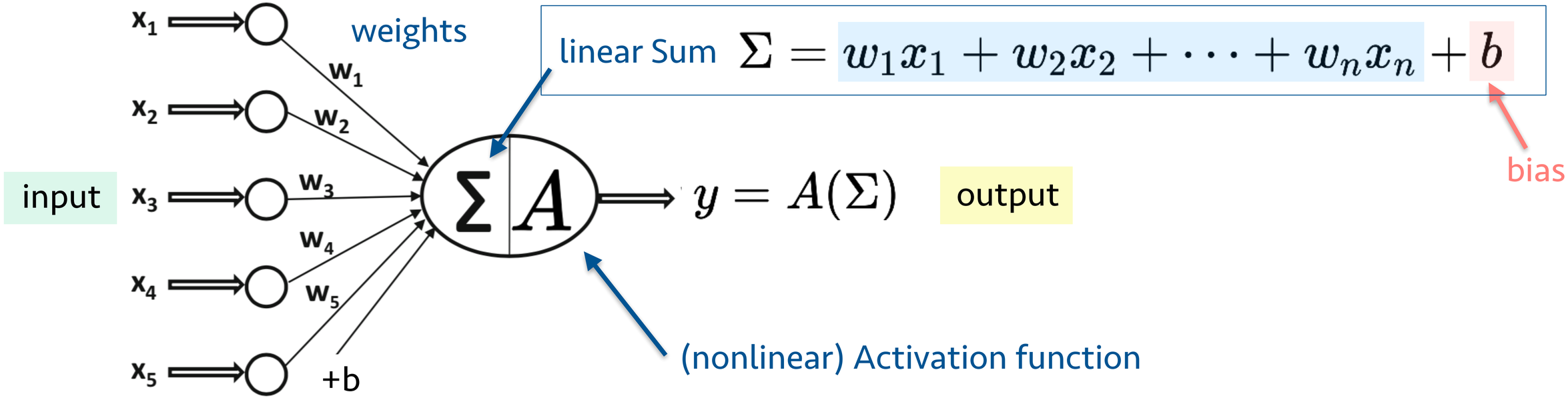
Pesios, Koutalios, Kugiumtzis, Stergioulas (2024)

ARTIFICIAL NEURAL NETWORKS



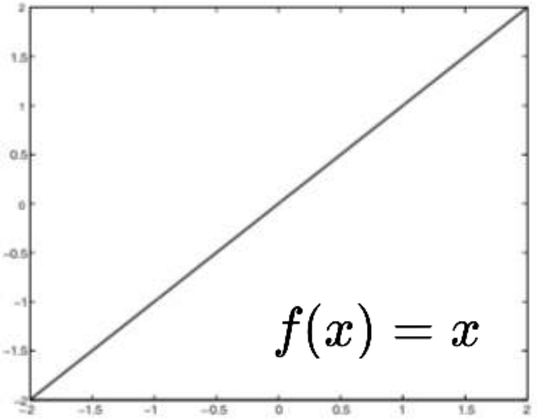
ARTIFICIAL NEURON

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

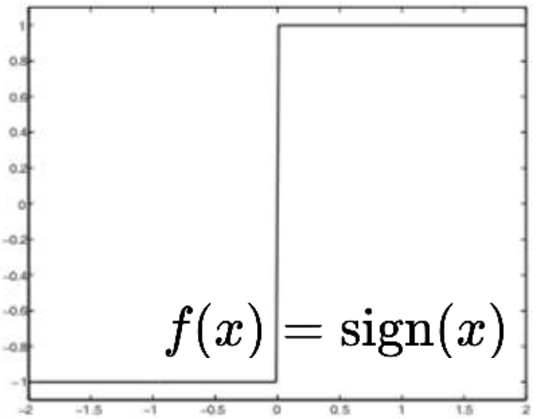


ACTIVATION FUNCTIONS

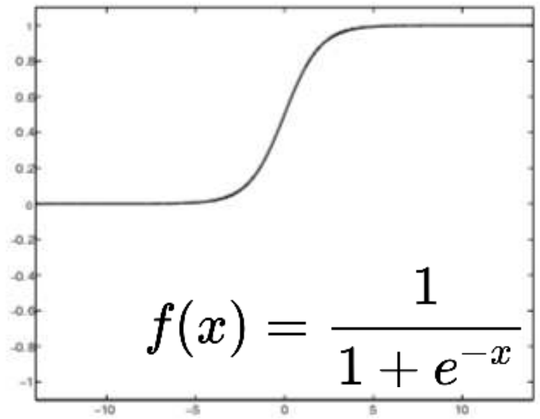
- Different common activation functions:



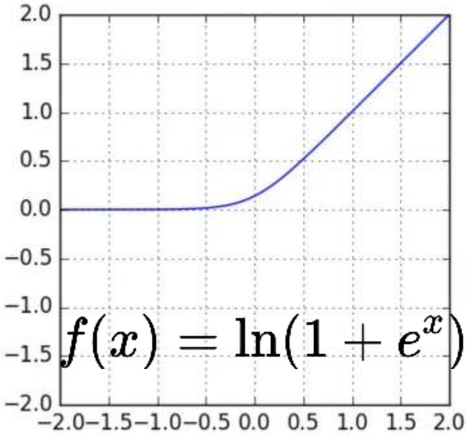
(a) Identity



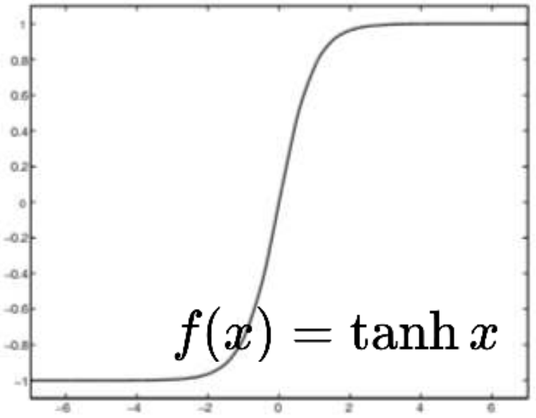
(b) Sign



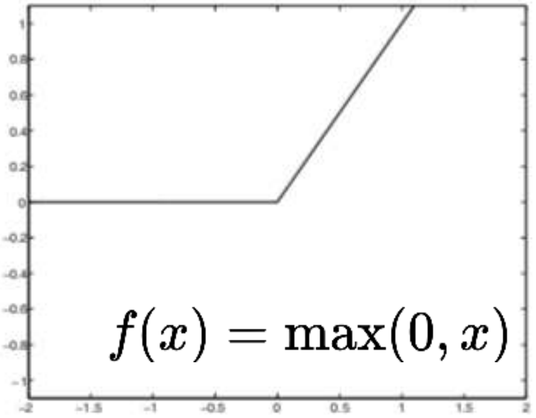
(c) Sigmoid (logistic function)



Softplus

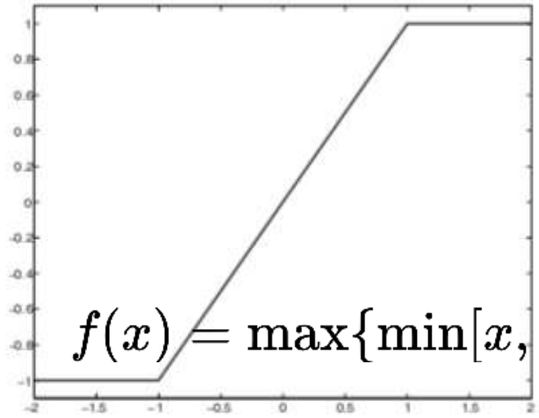


(d) Tanh



(e) ReLU

(Rectified Linear Unit)



(f) Hard Tanh

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)

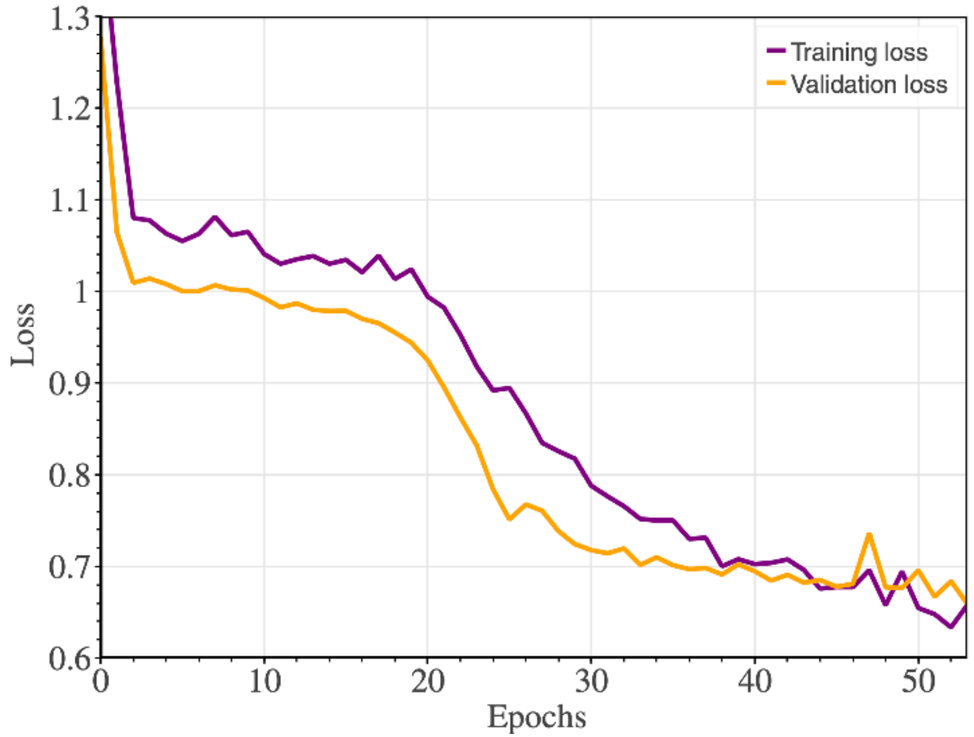
ANN SURROGATE IN THE FREQUENCY DOMAIN

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)

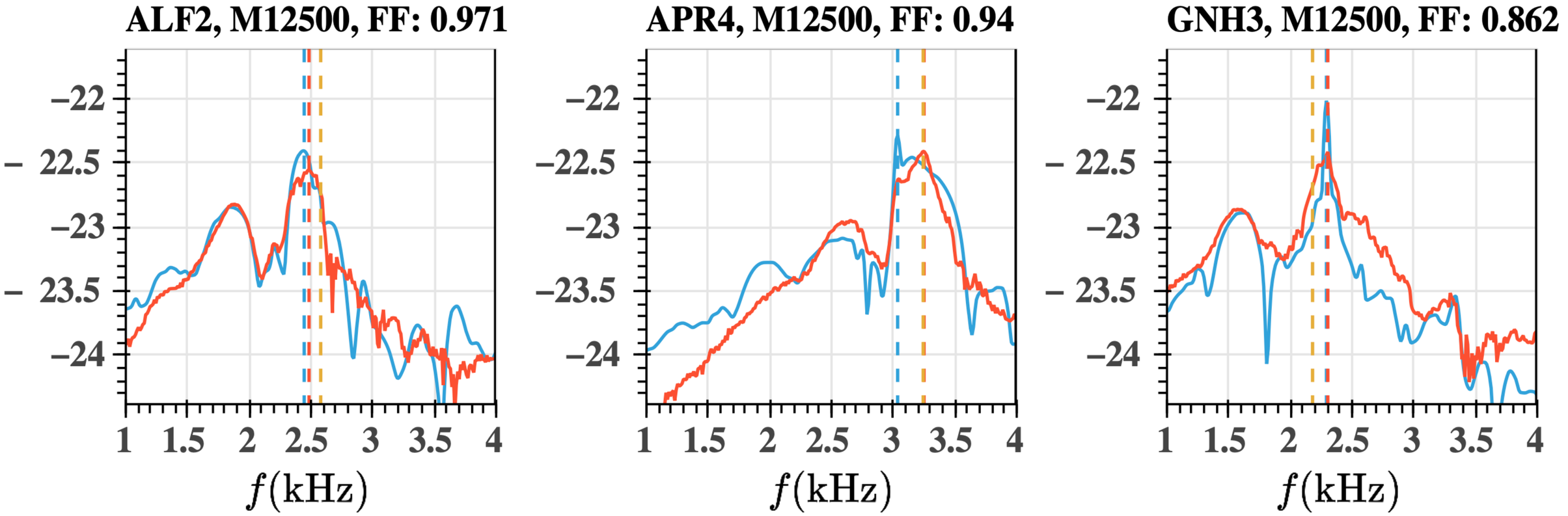
ANN SURROGATE IN THE FREQUENCY DOMAIN

Typical examples of predicted magnitudes of GW spectra

FF = O = overlap

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

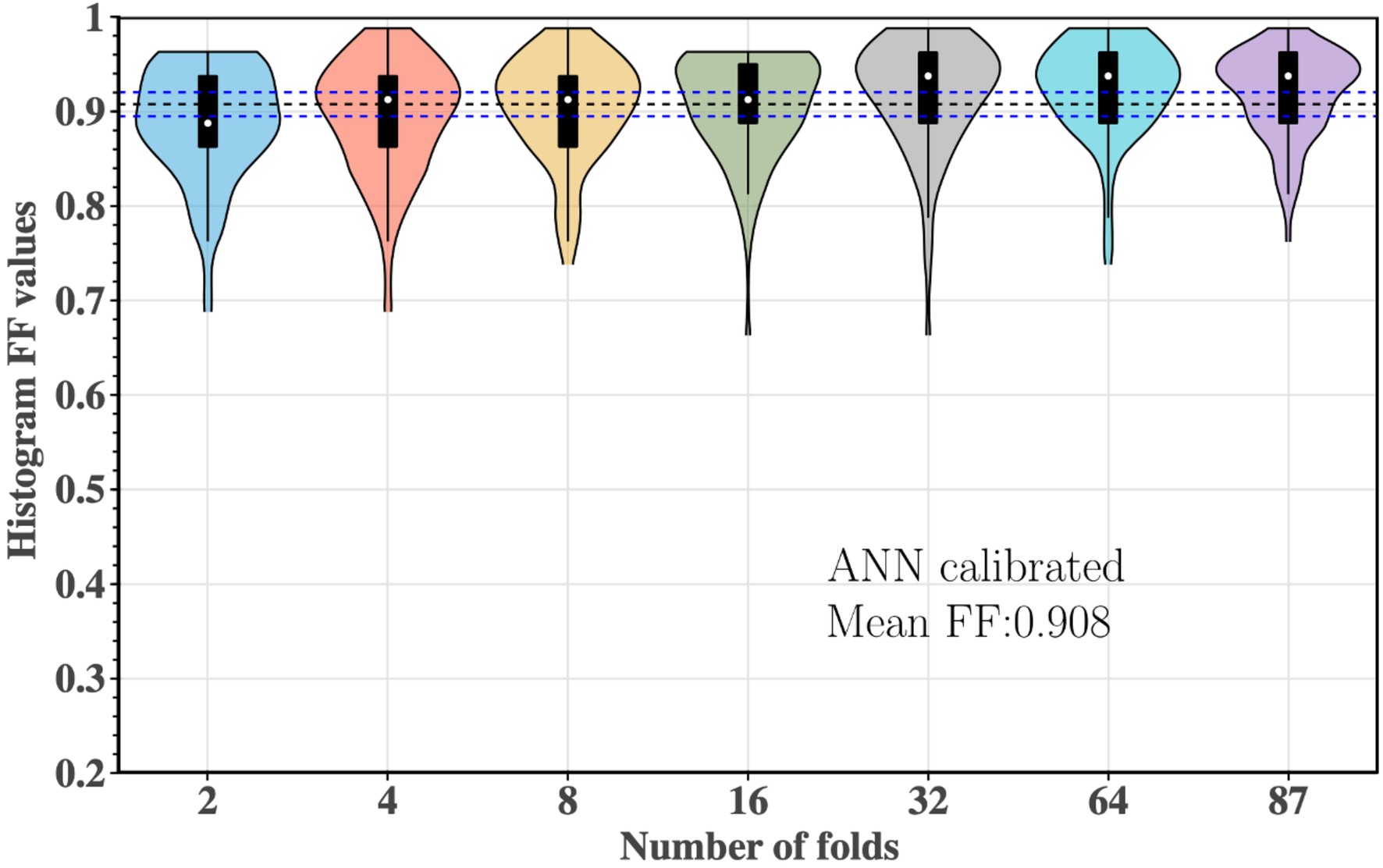
ANN outperforms multivariate linear regression.



Pesios et al. (2024)

ANN SURROGATE IN THE FREQUENCY DOMAIN

Cross-validation study of fitting factors distribution:

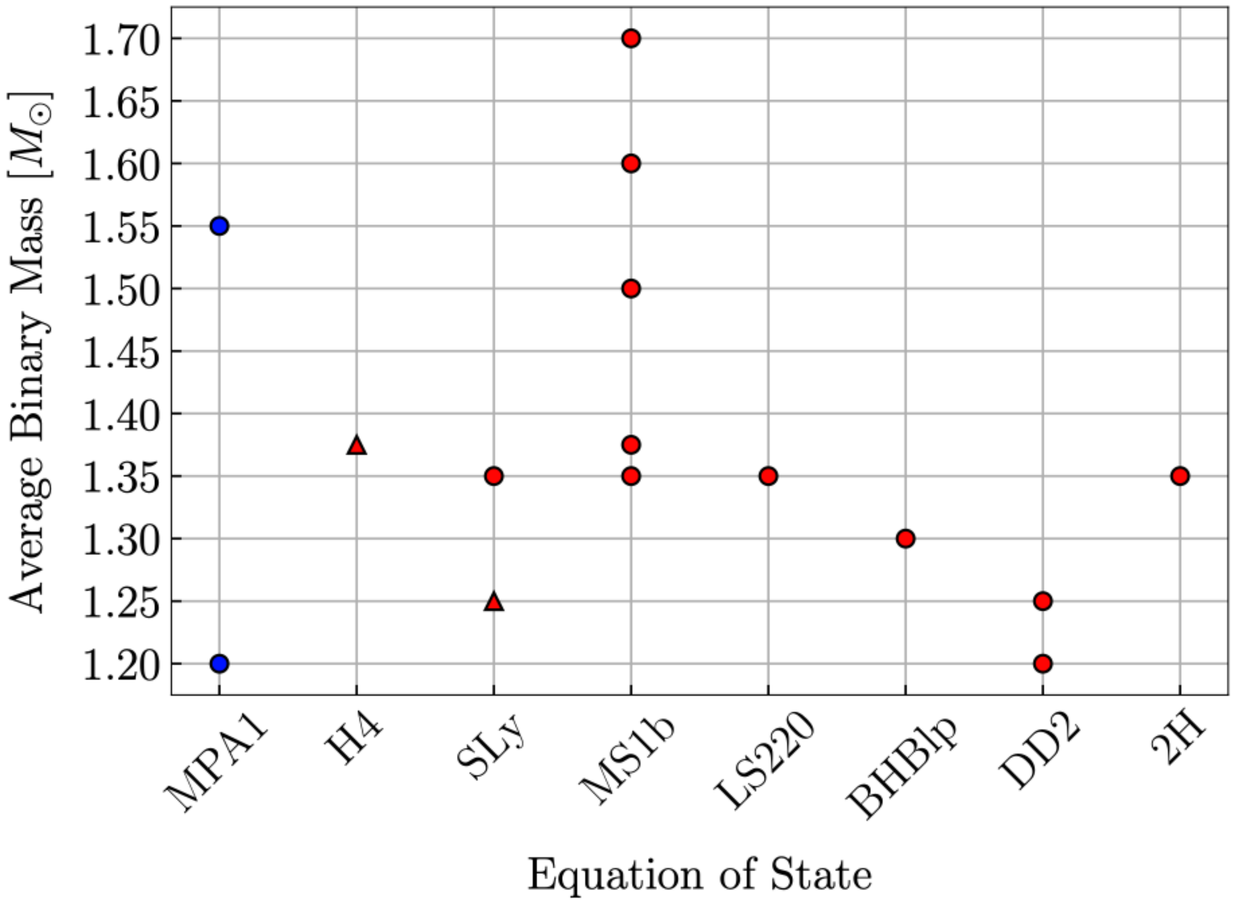


ACCELERATION OF POST-MERGER INFERENCE

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]



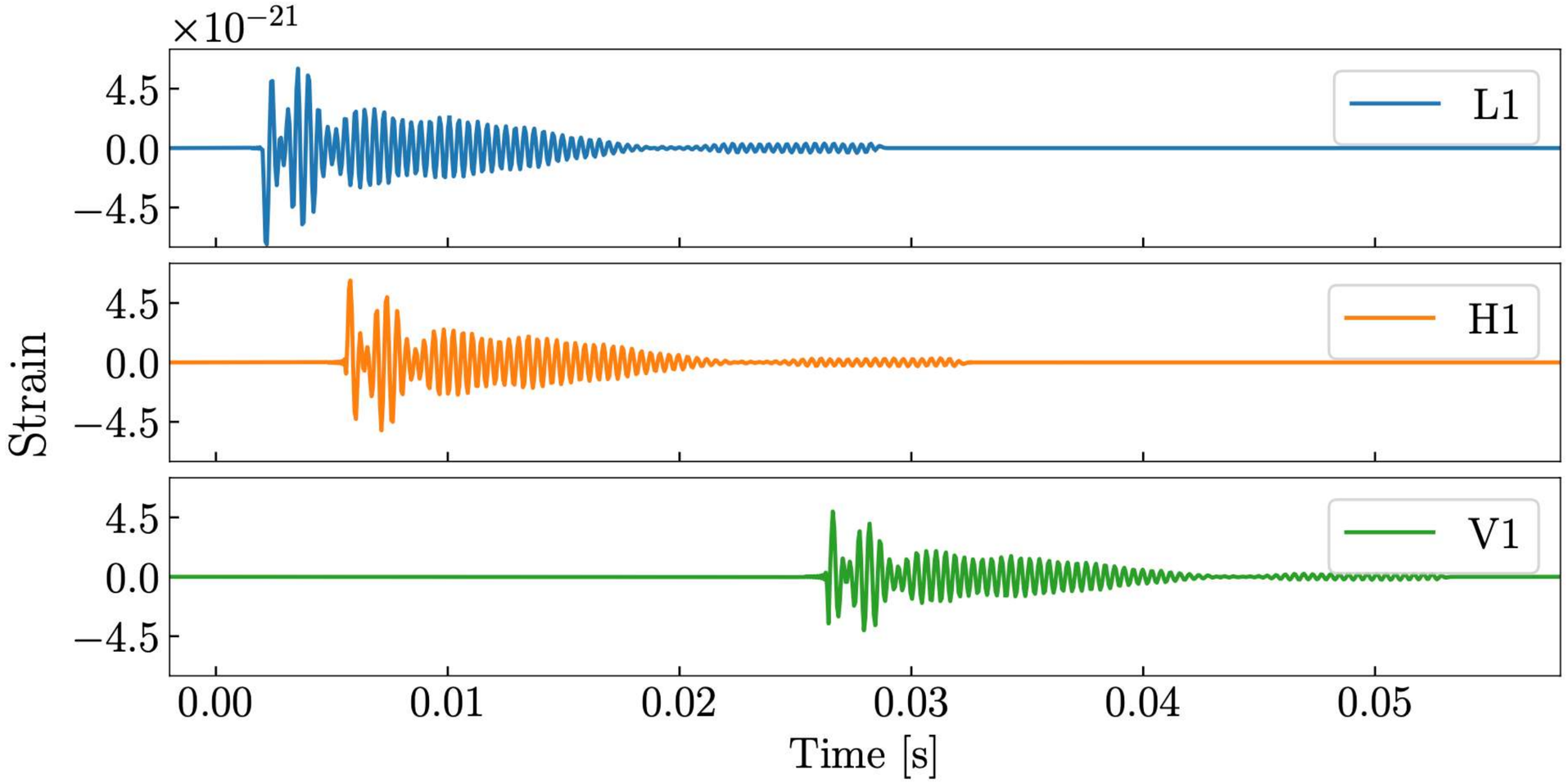
Extension of the set of 9 waveforms used in previous work by Easter et al. (2020)

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

$h_{j,X}(\boldsymbol{\theta},t)$ is obtained by applying a $\pi/2$ 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 $\mathcal{U}(1,5)[\mathrm{kHz}]$ for f_{spiral}.
  - Type II: Gaussian priors, $\mathcal{N}(f_{2-0}, \sigma^2)$ for f_{2-0} and $\mathcal{N}(f_{\mathrm{spiral}}, \sigma^2)$ for f_{spiral}.
  - Type III: Gaussian priors, $\mathcal{N}(f_{\mathrm{spiral}}, \sigma^2)$ for f_{spiral} and uniform priors $\mathcal{U}(1,5)[\mathrm{kHz}]$ for f_{2-0}.

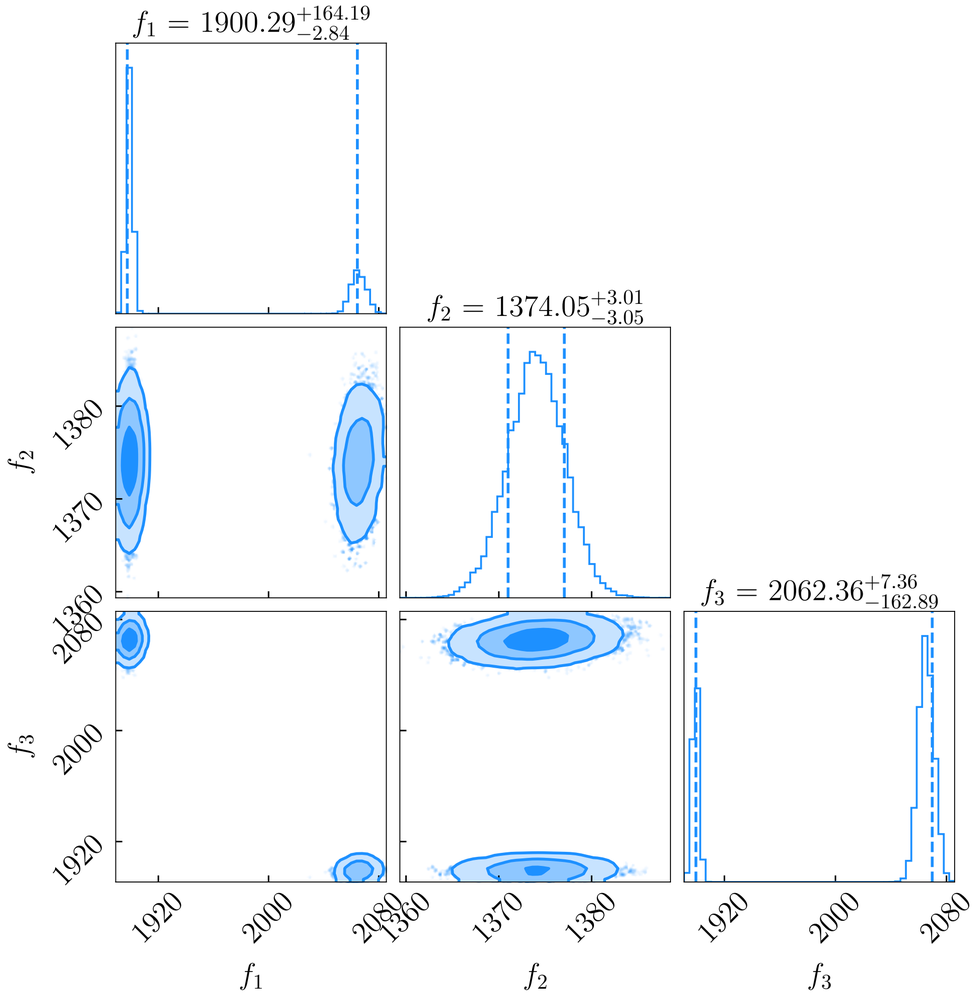
where σ = 3 x max error of empirical relations. For f_{peak}: Gaussian ; for f_4: flat in [f_{peak}+0.3kHz, 5 kHz].

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)

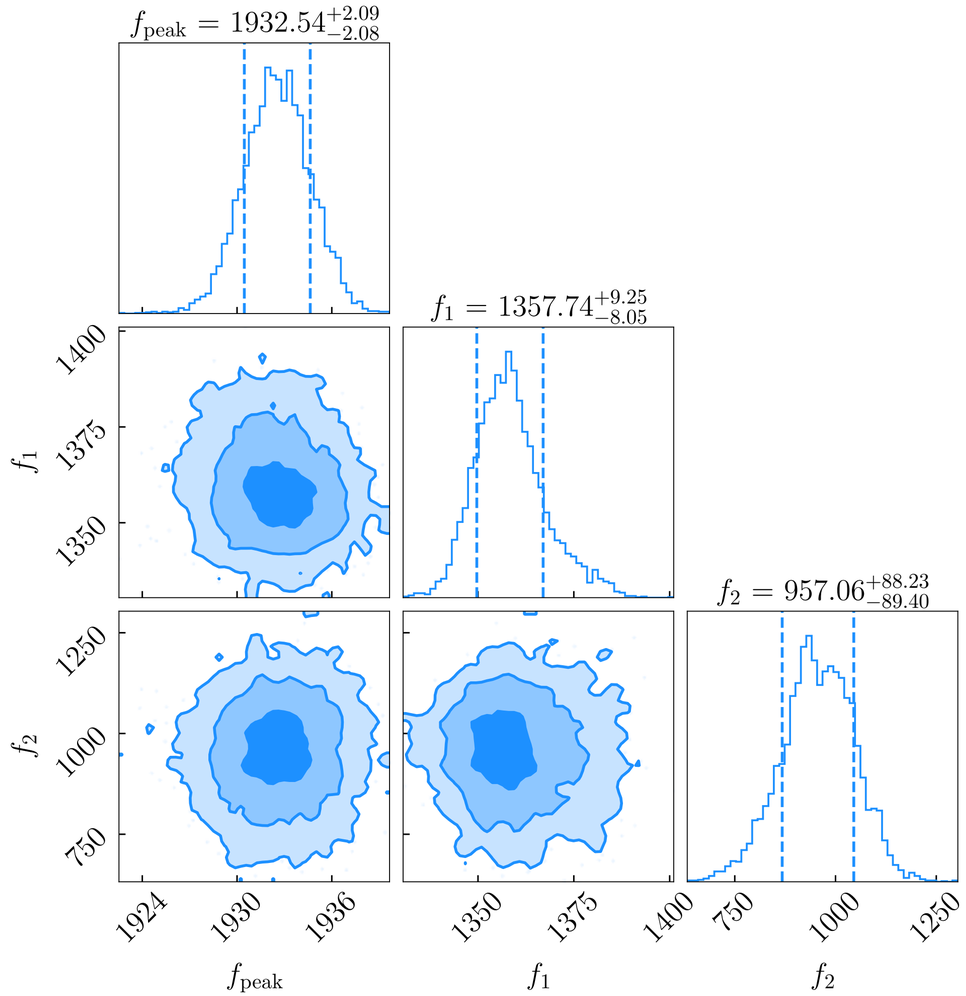
POSTERIOR DISTRIBUTIONS

EOS 2H 1.35+1.35, SNR = 50

using flat priors



using informed priors (through empirical relations)

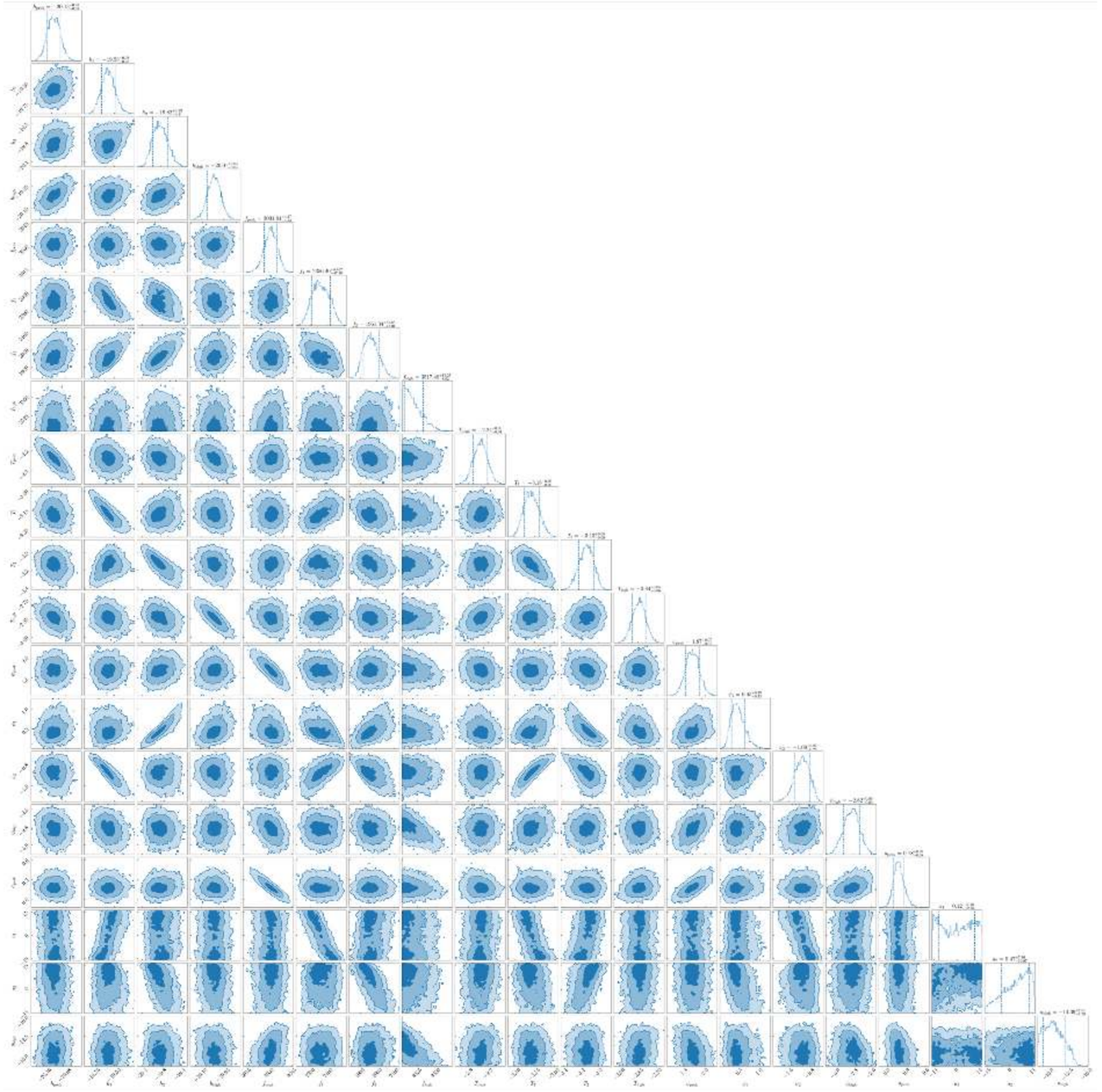


POSTERIOR DISTRIBUTIONS

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.

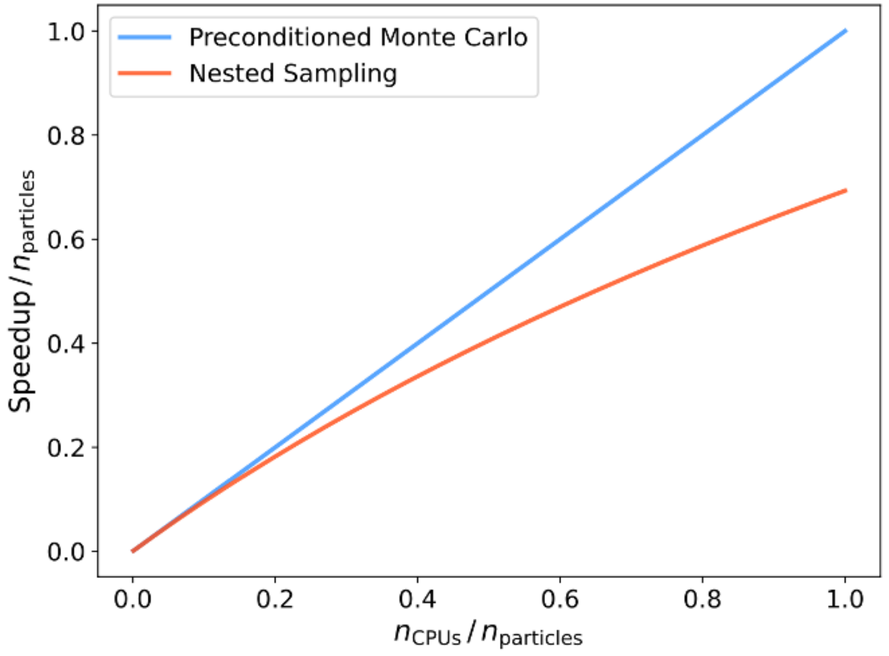
We use 2000 particles that transition from the prior ($\beta_0 = 0$) to the posterior distribution ($\beta_T = 1$), through a sequence of reweighing, resampling and mutation steps.

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.

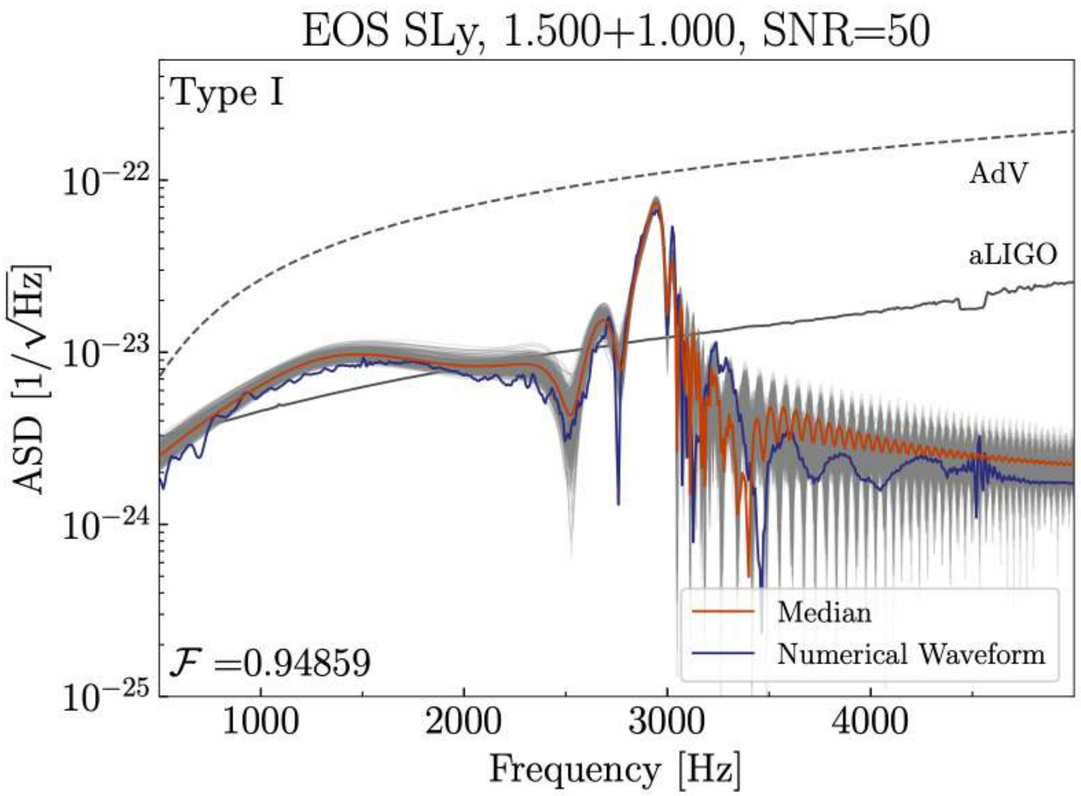
pocoMC

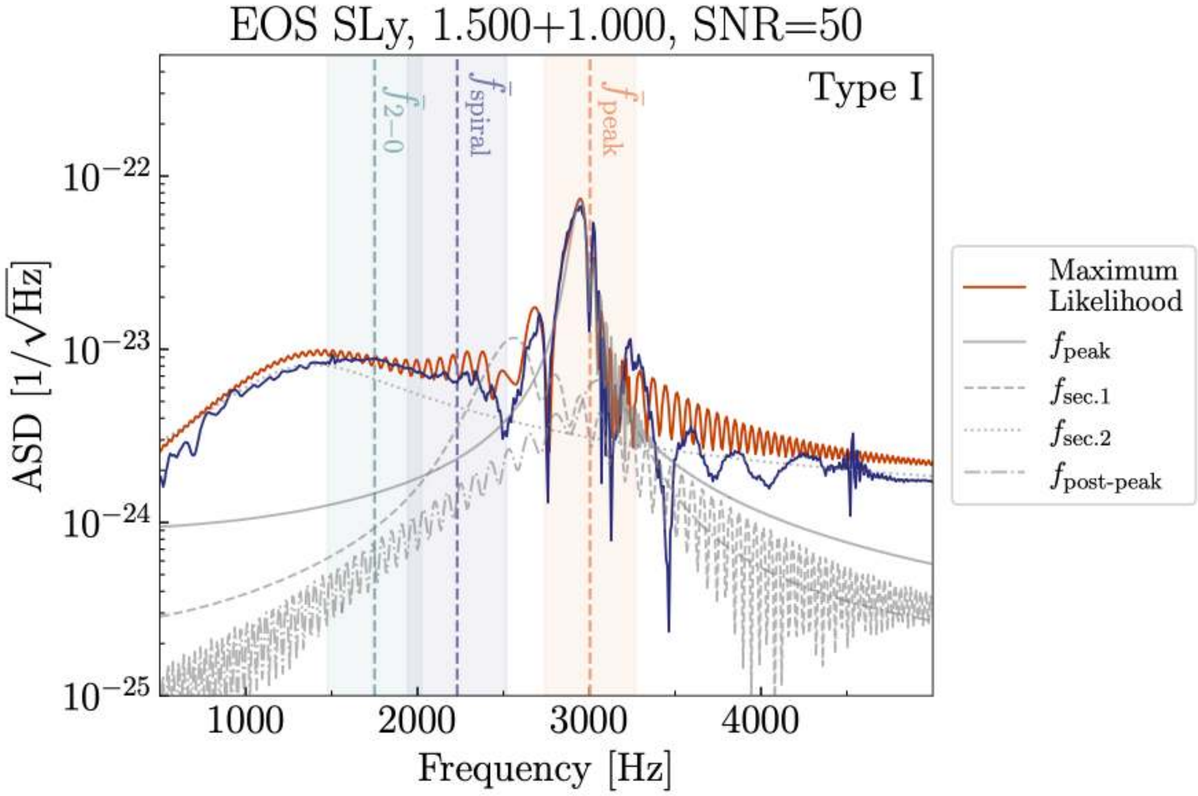
https://github.com/minaskar/pocomc



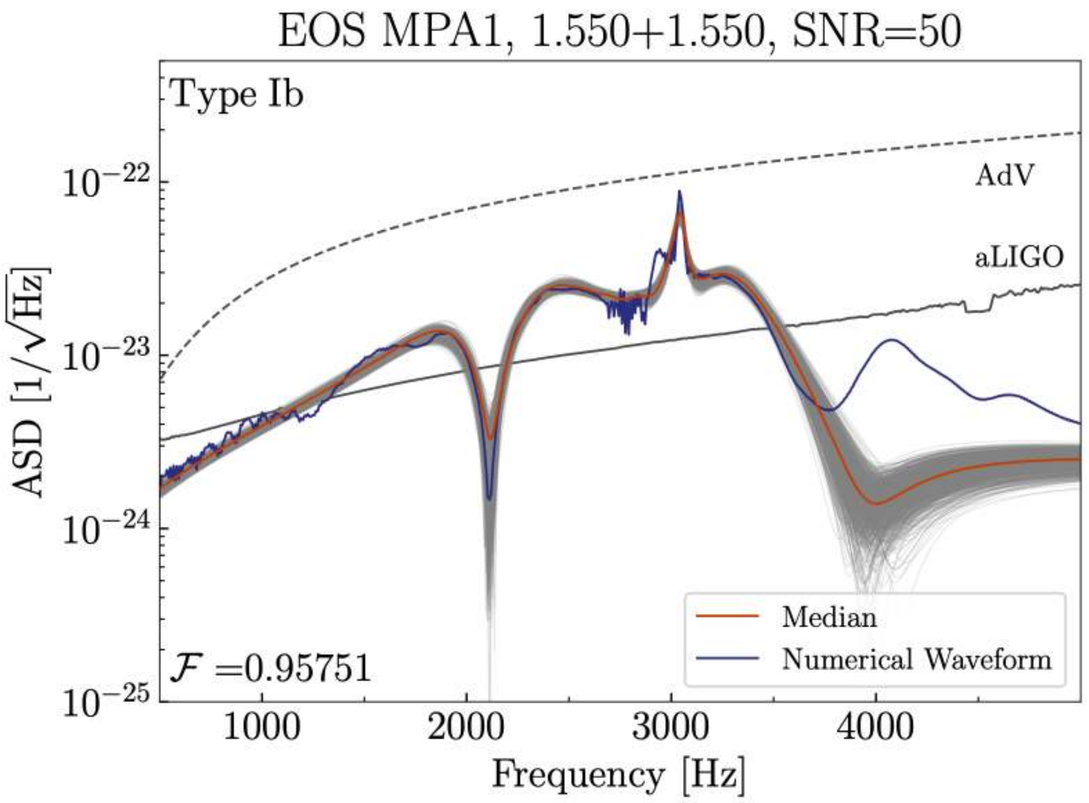
PocoMC is highly parallelizable

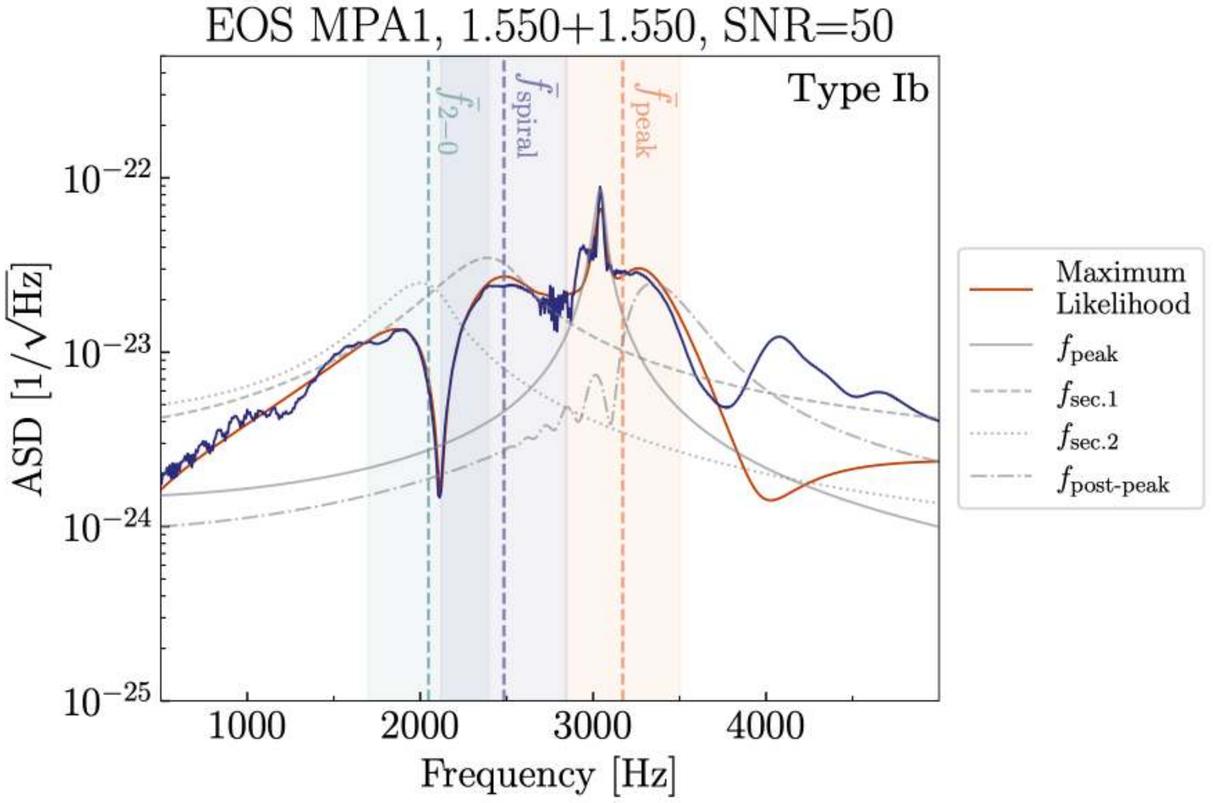
RECONSTRUCTION IN THE FREQUENCY DOMAIN



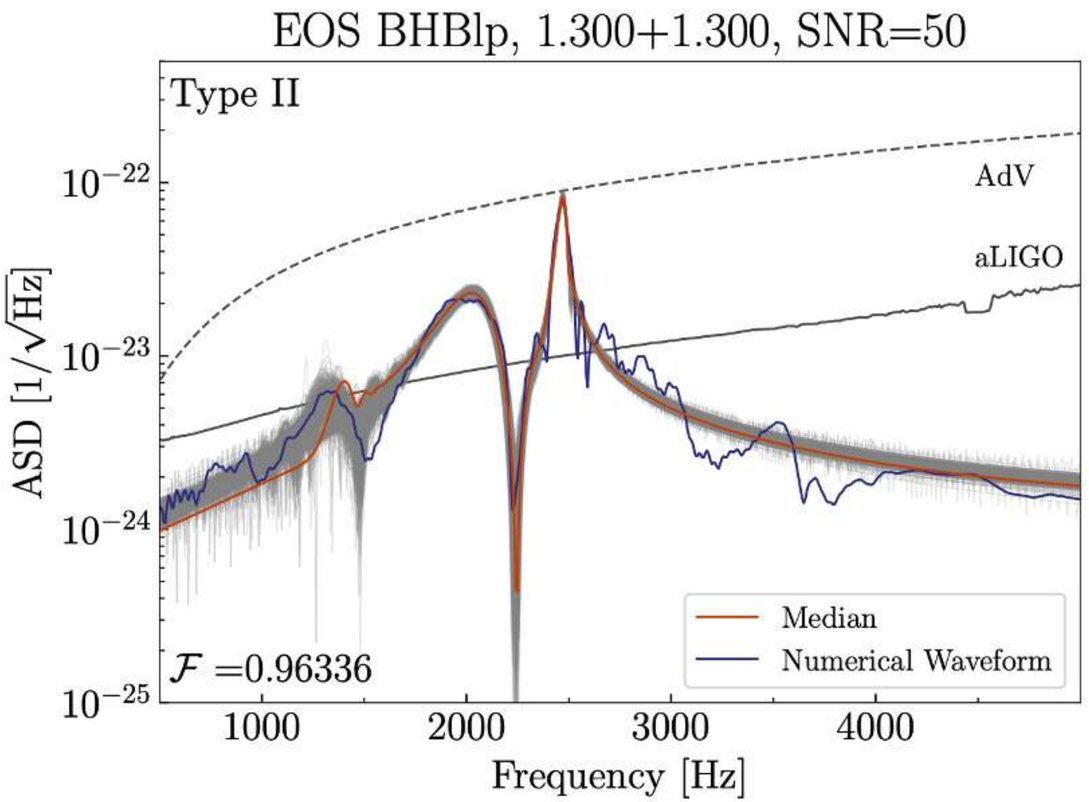


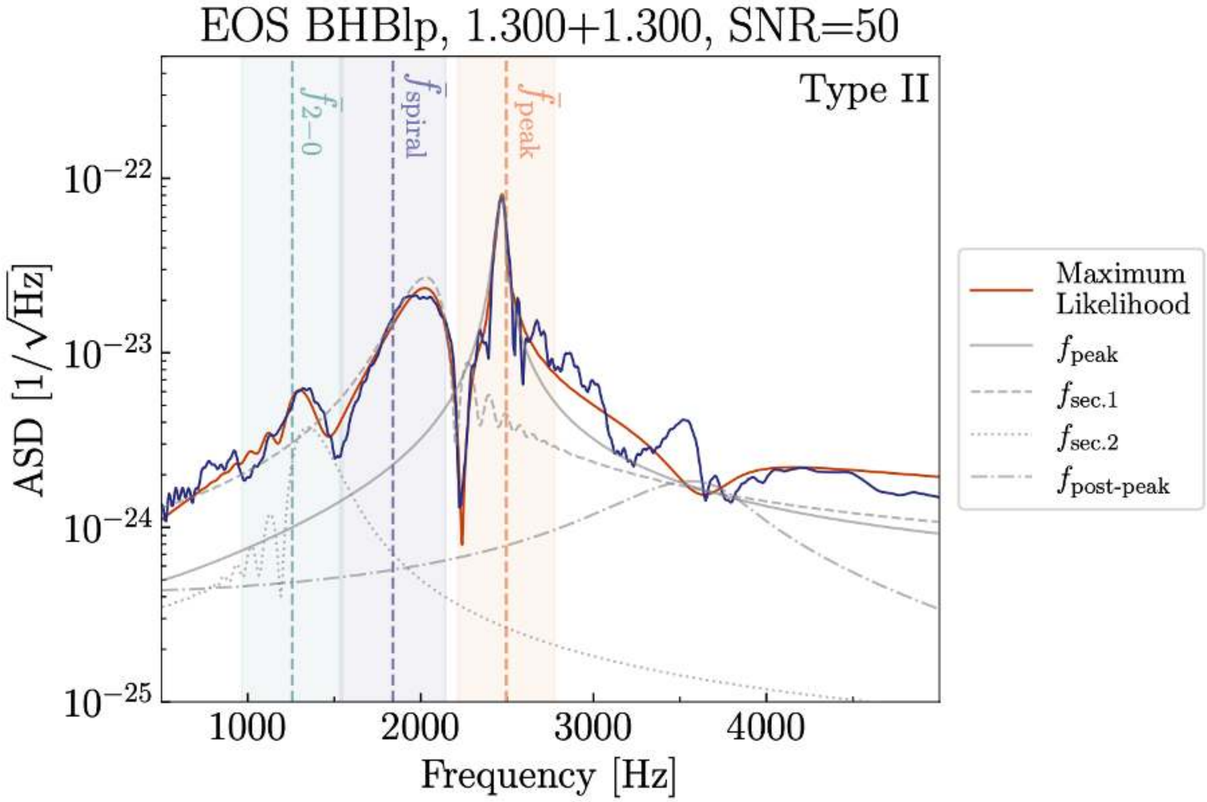
RECONSTRUCTION IN THE FREQUENCY DOMAIN



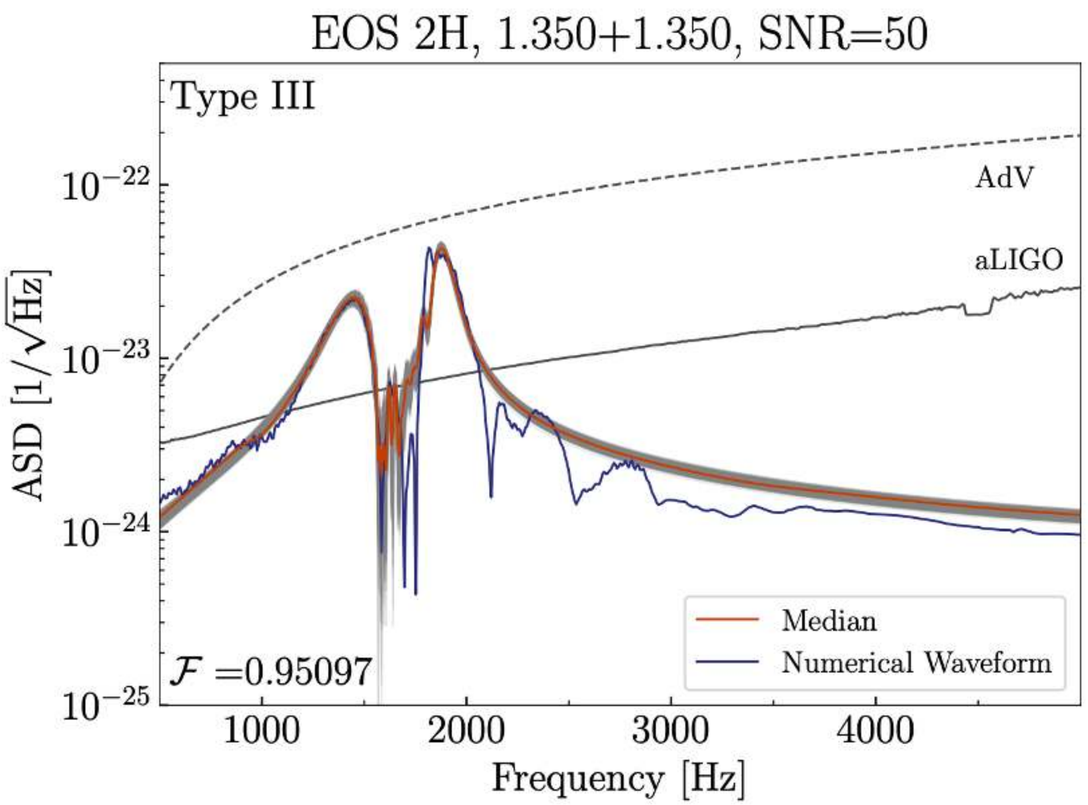


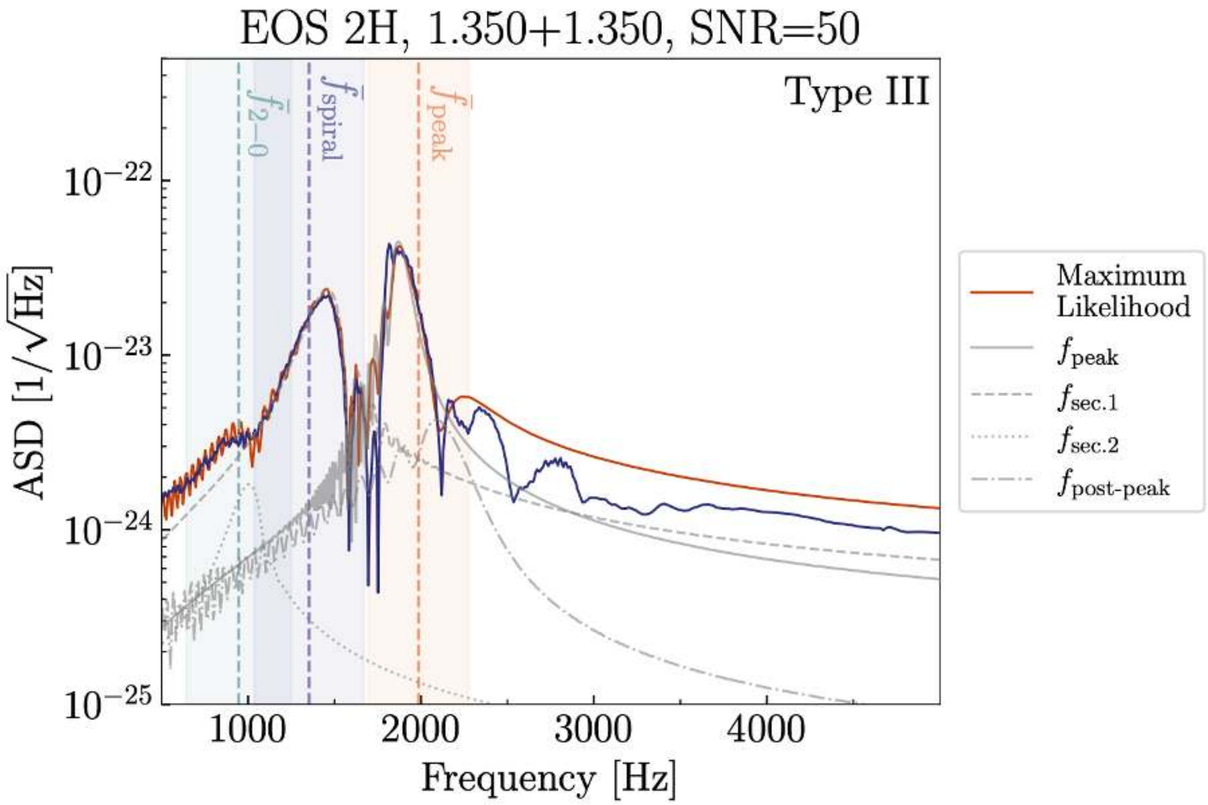
RECONSTRUCTION IN THE FREQUENCY DOMAIN



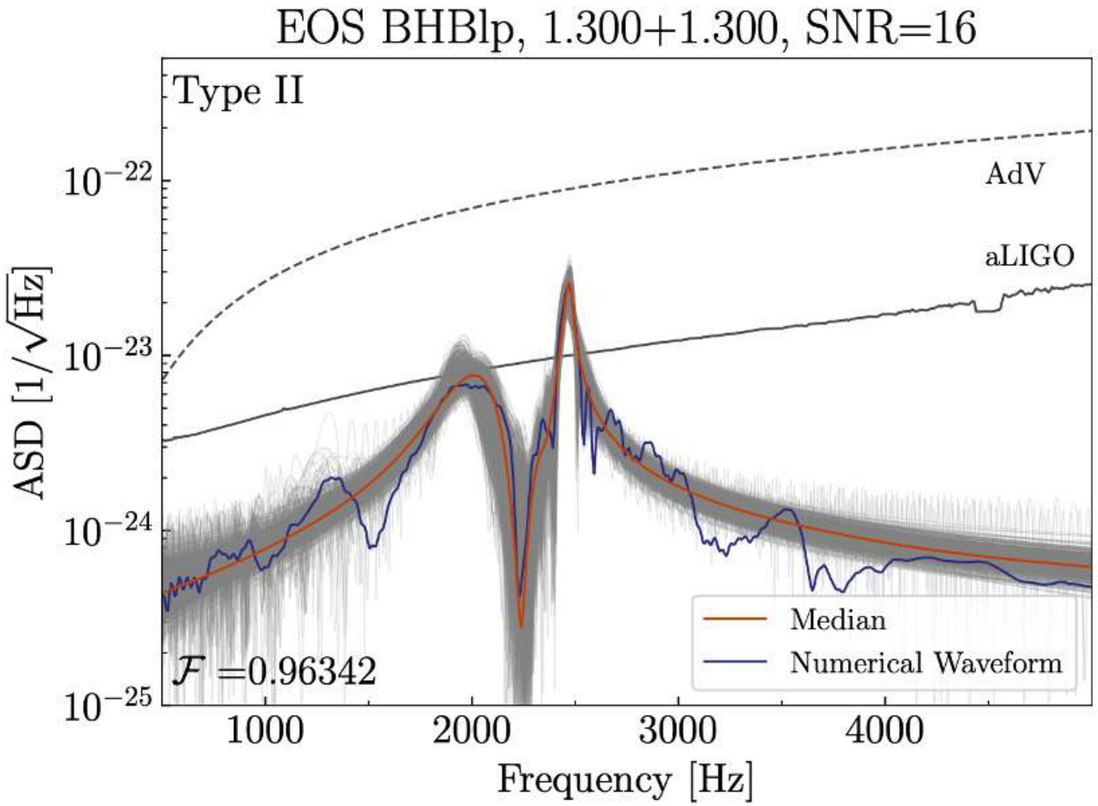


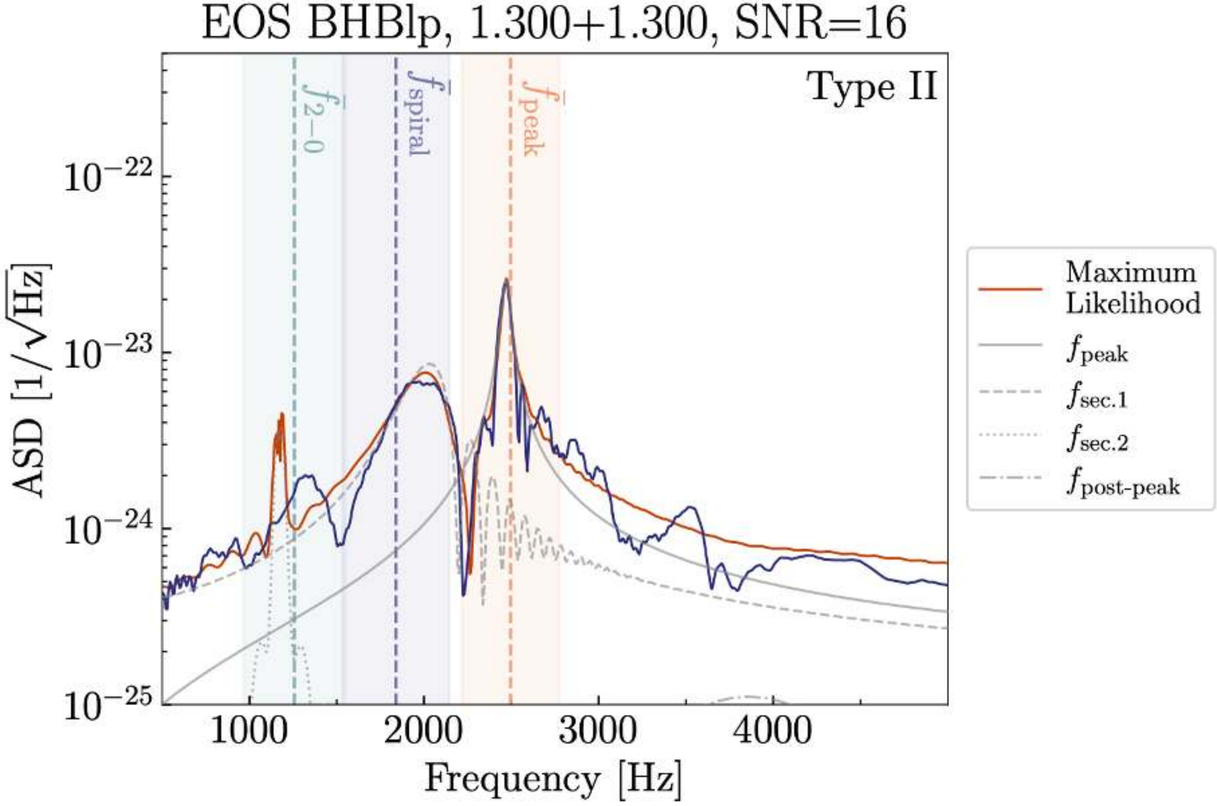
RECONSTRUCTION IN THE FREQUENCY DOMAIN



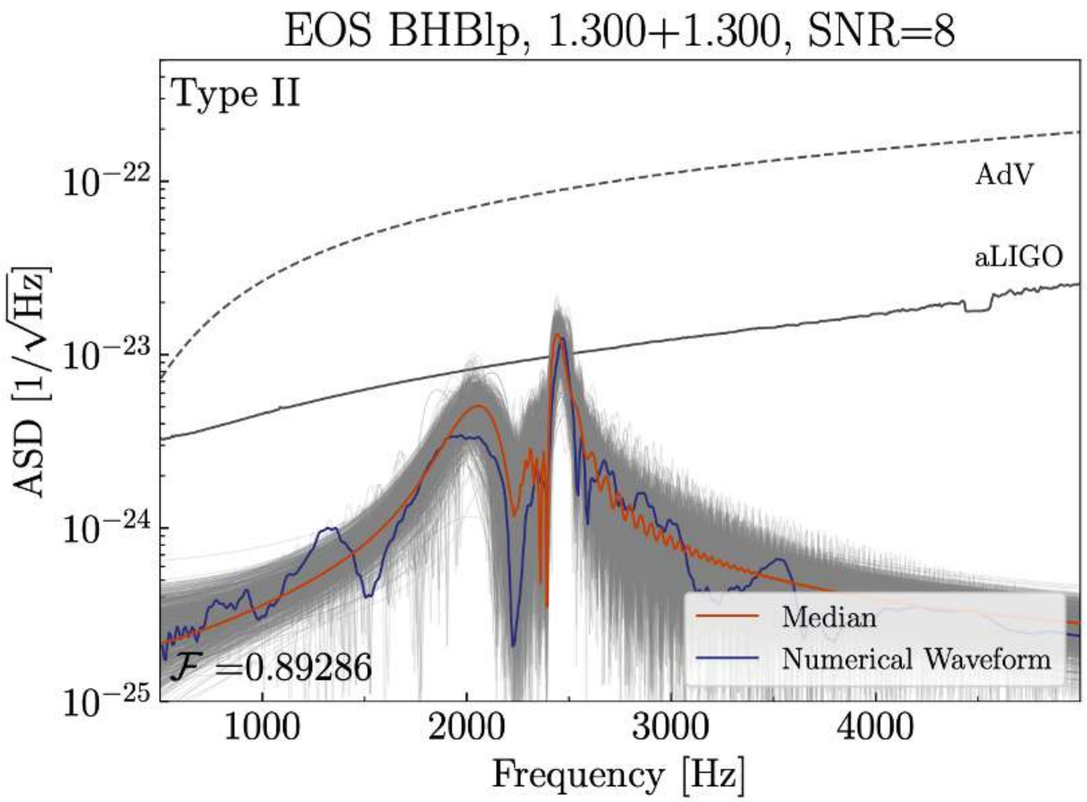


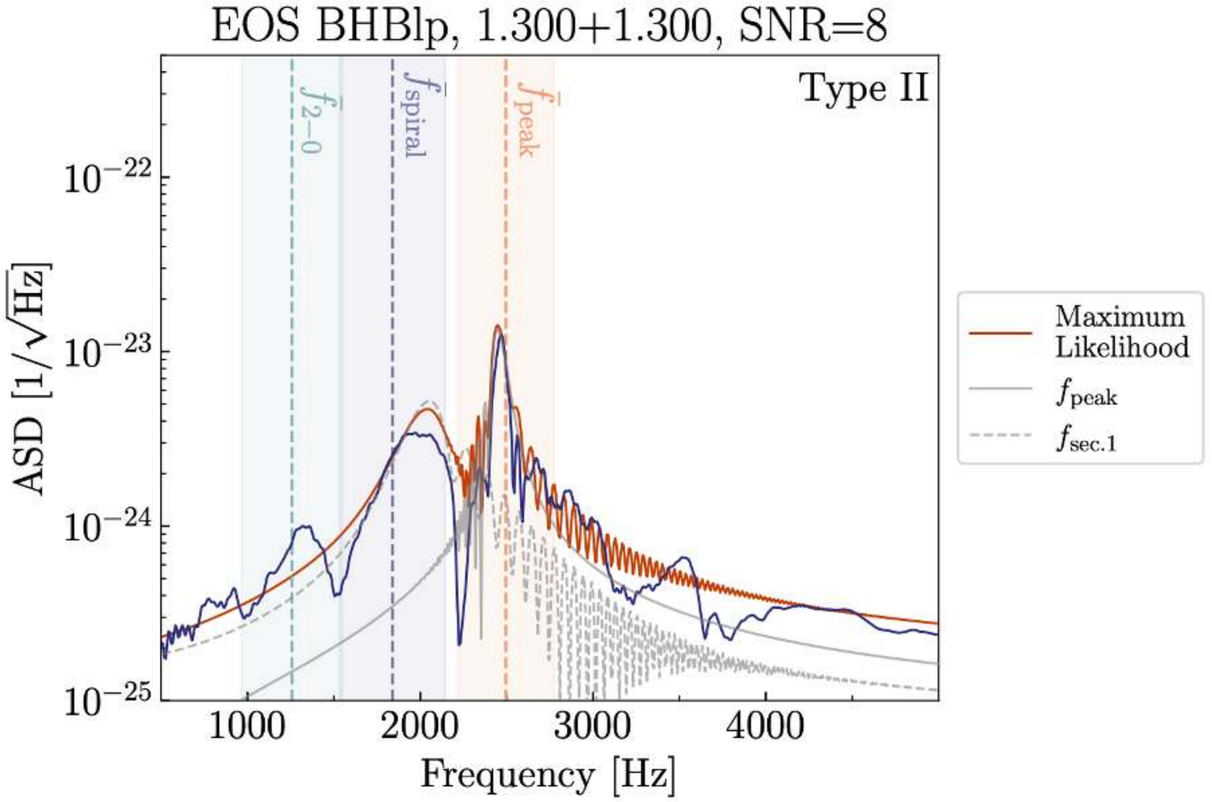
RECONSTRUCTION IN THE FREQUENCY DOMAIN





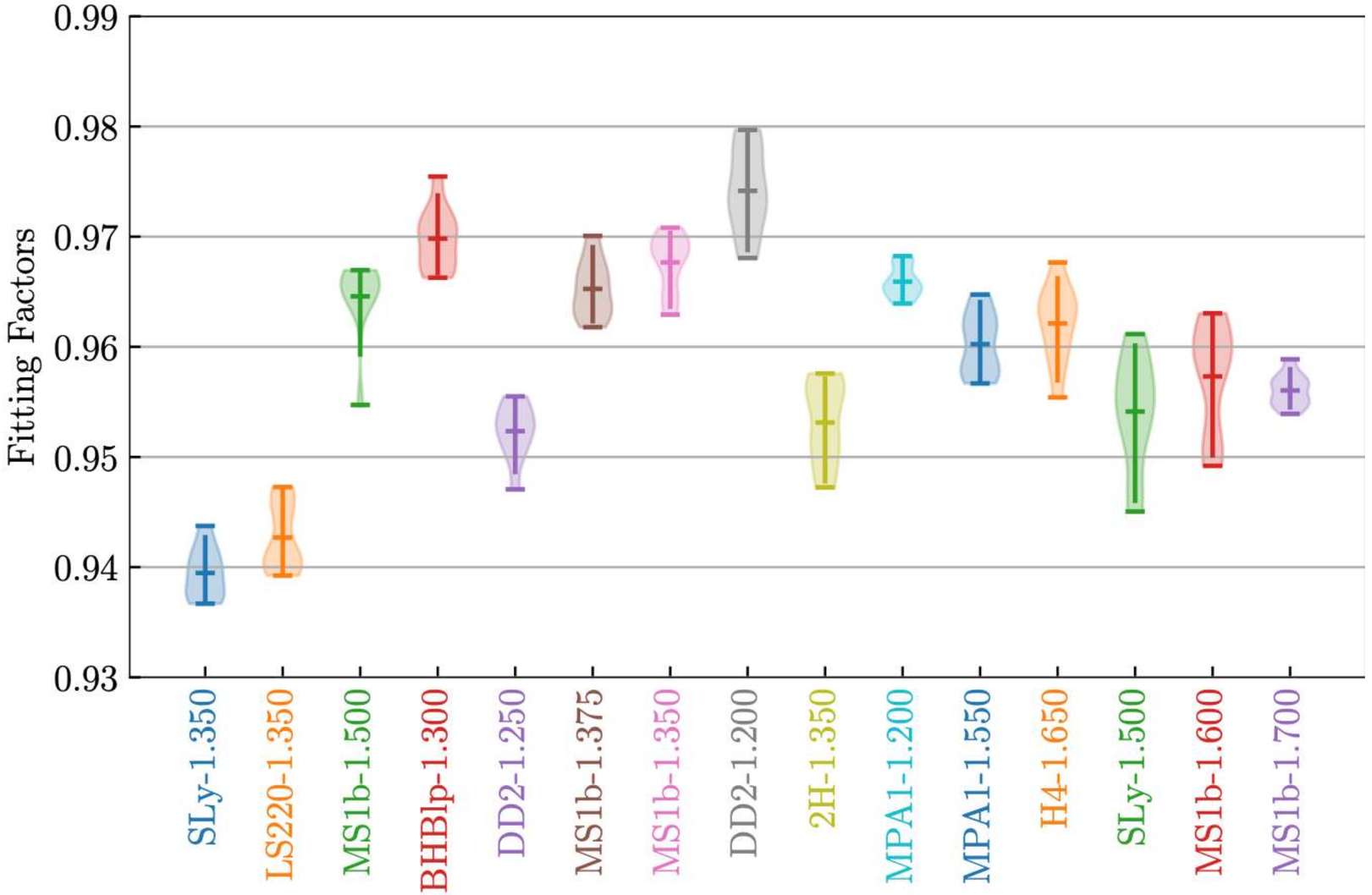
RECONSTRUCTION IN THE FREQUENCY DOMAIN





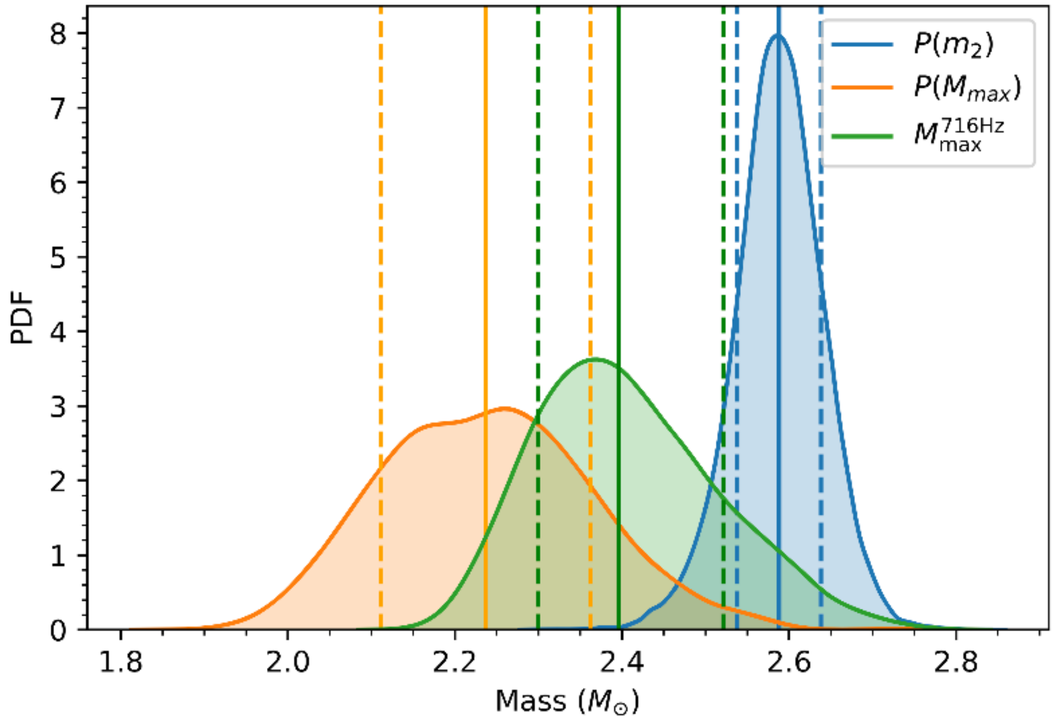
FITTING FACTORS FOR SNR=50

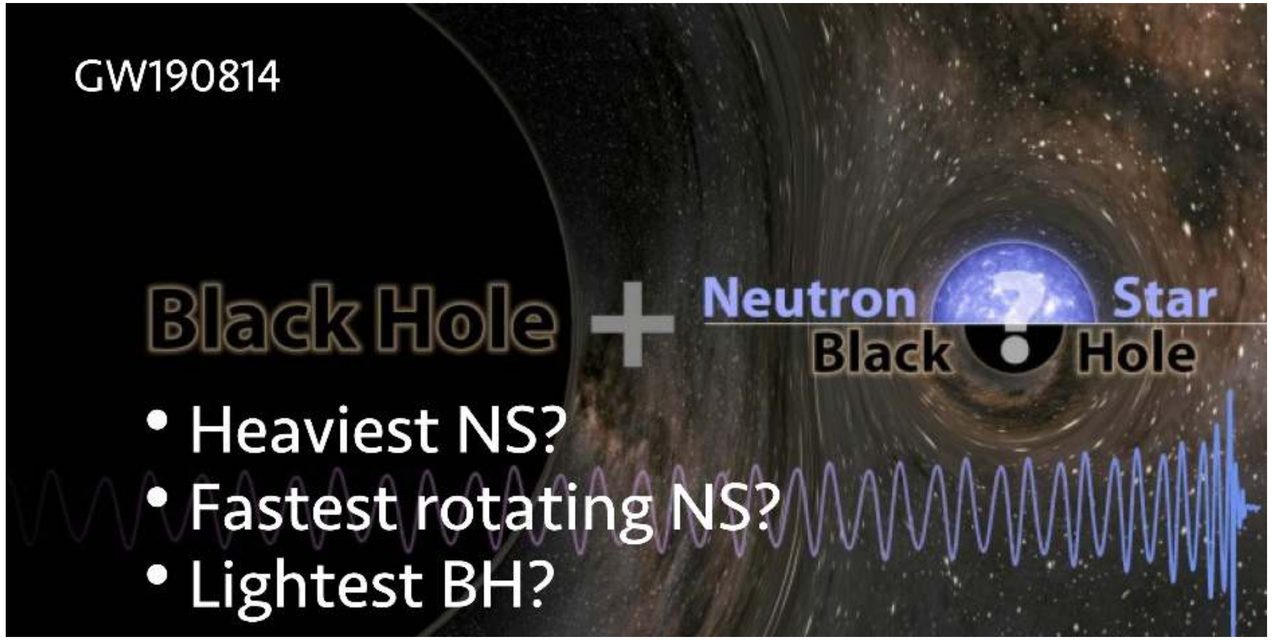
The fitting factors achieved are consistent with those in Easter et al. (2020) for a smaller set of EOS.



HIGH MASS NEUTRON STARS?

What was the nature of the lighter component in GW190814?





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

Another possibility: 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

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

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)

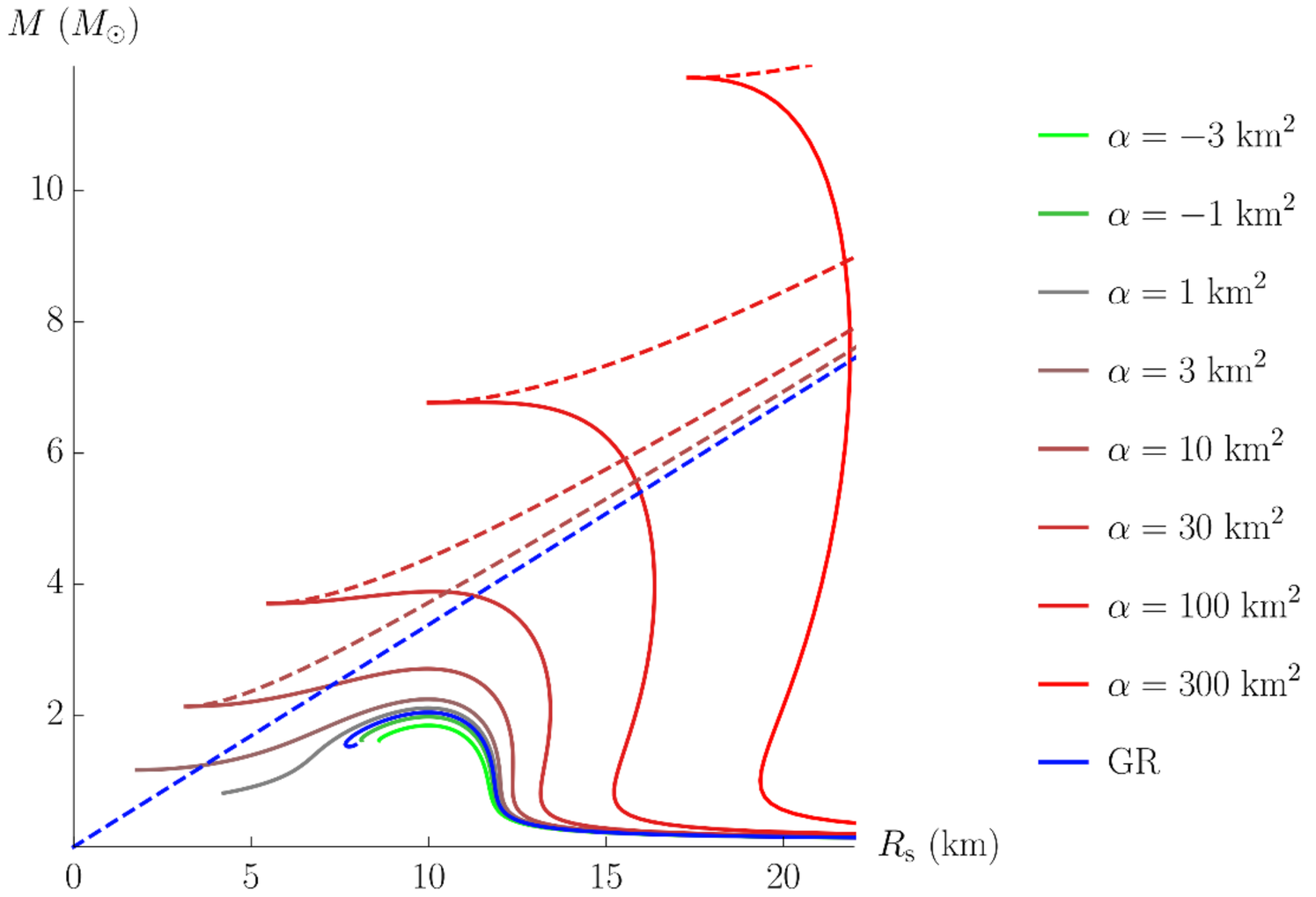
and the shift-symmetric scalar field is

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

Charmousis, Lehebel, Smyrniotis, Stergioulas (2022)

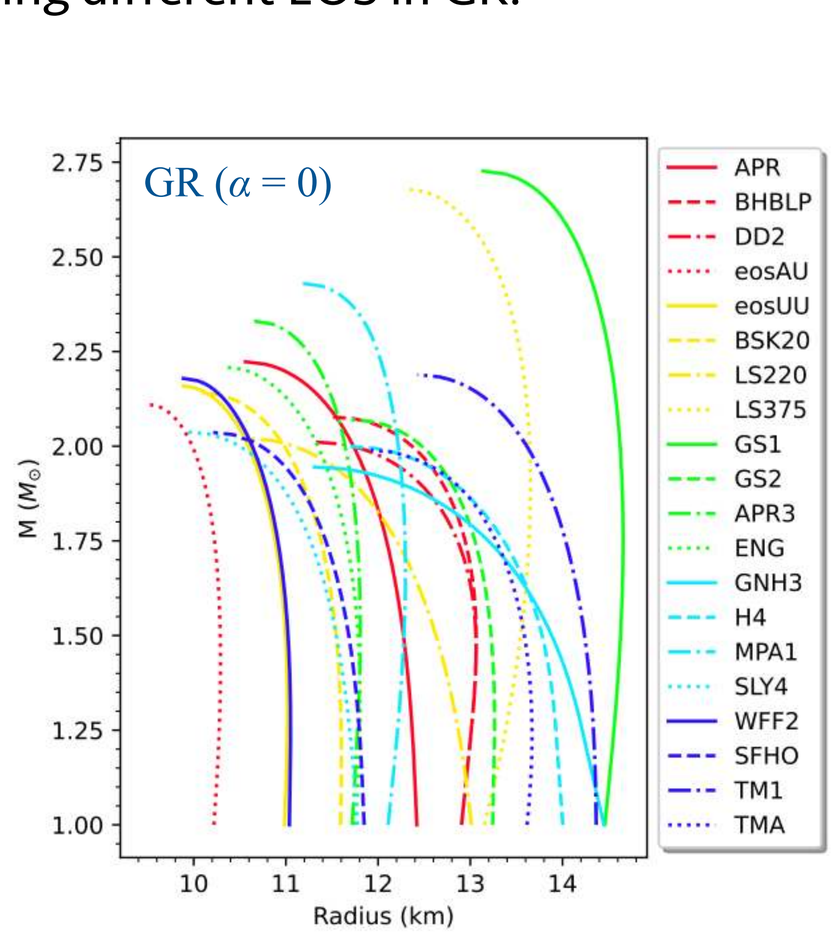
SEQUENCES OF EQUILIBRIUM MODELS

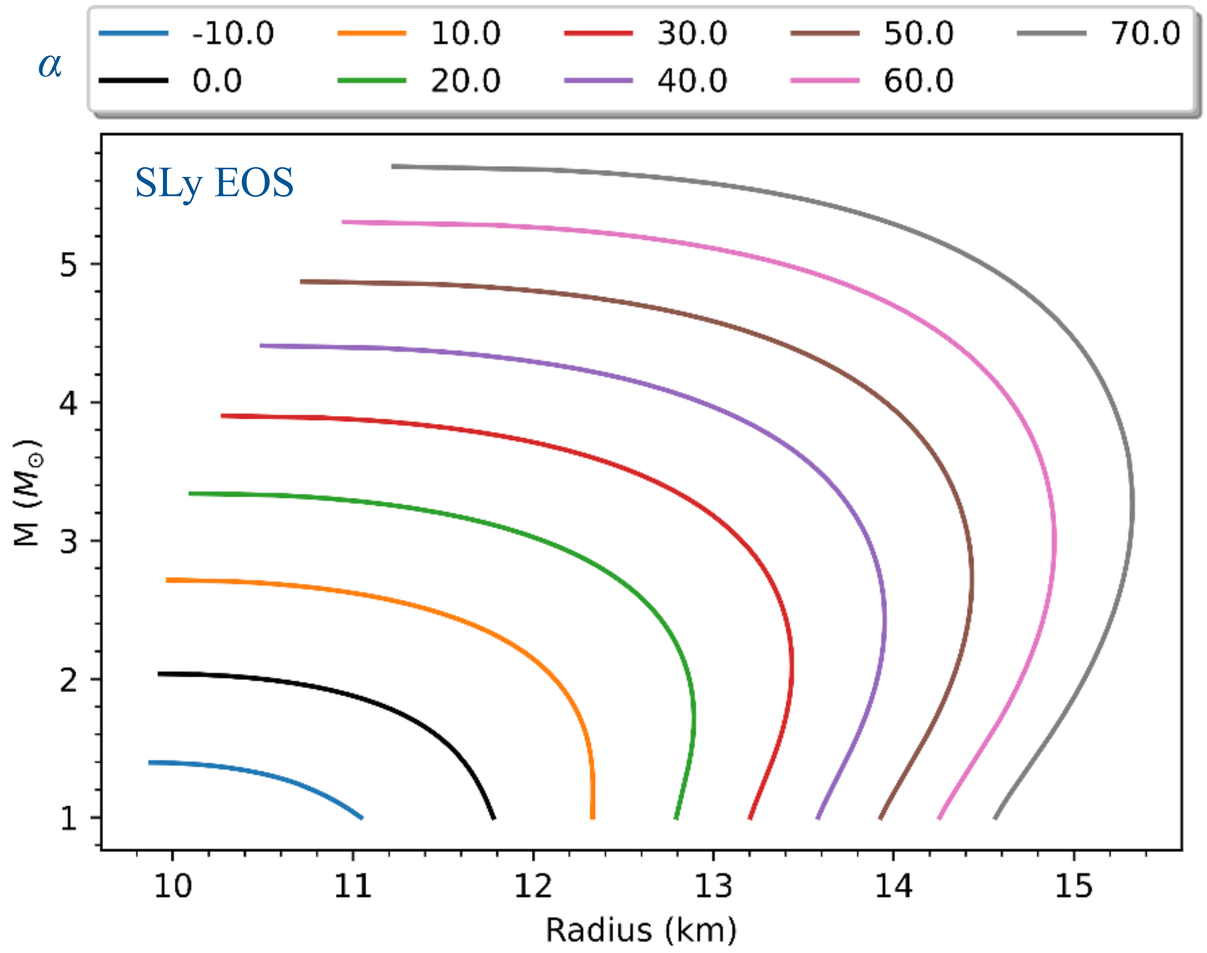
For the **SLy EOS**, we show representative cases of sequences of NS equilibrium models (solid lines) and BHs (dashed lines). For α>0, the NS solutions merge with the minimum mass BH solution.



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

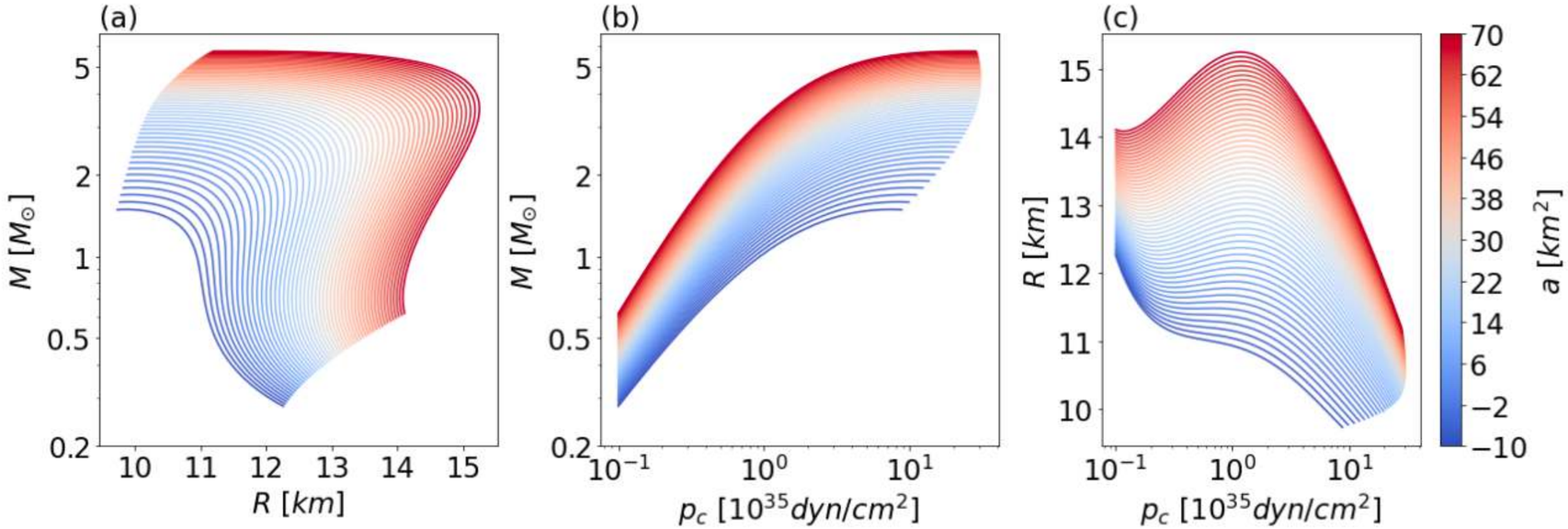
$$f_1(\mathrm{EoS};\alpha,\mathrm{p_c}) \rightarrow (\mathrm{M},\mathrm{R})$$

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

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

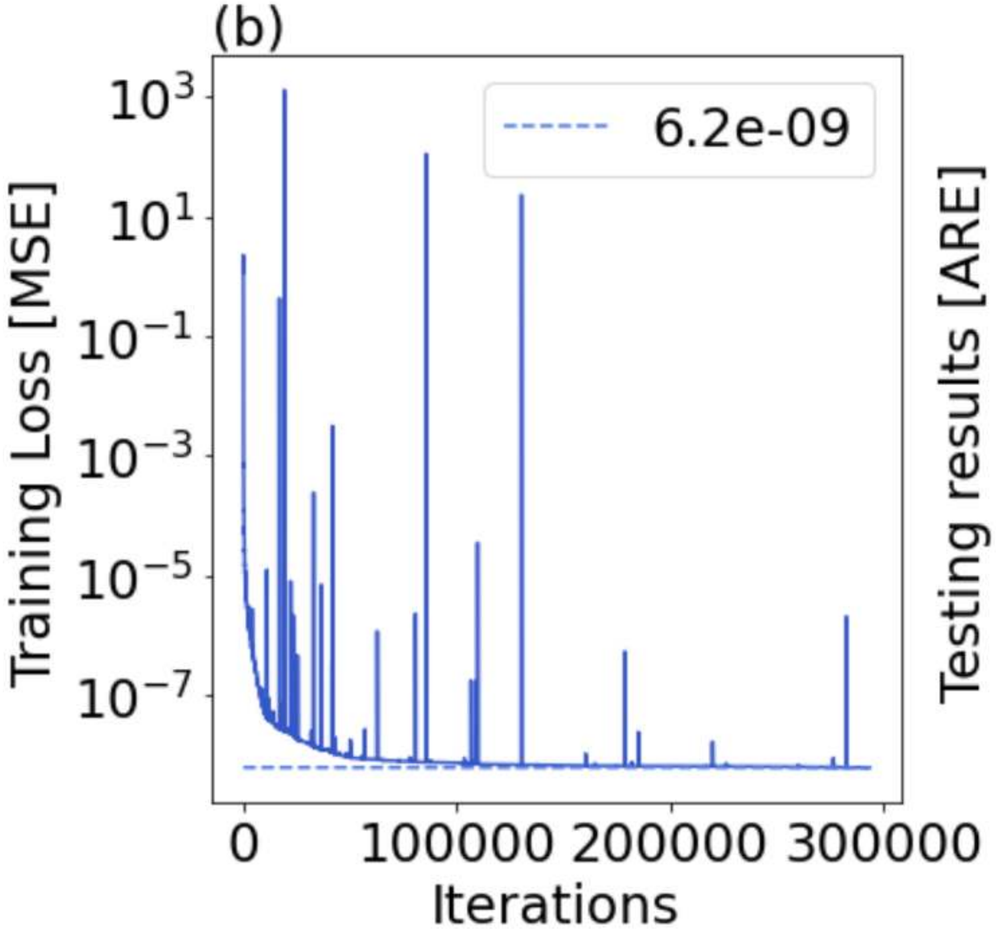


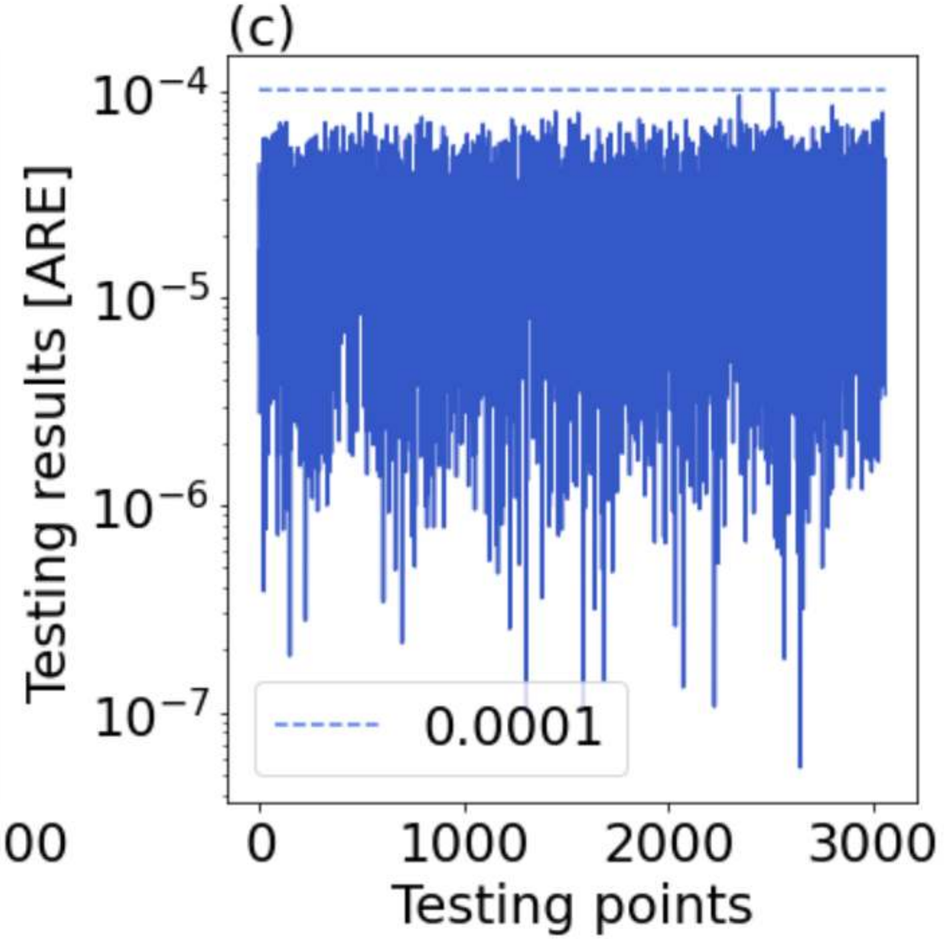
Liodis, Smyrniotis, Stergioulas (2024)

ACCURACY OF f_1 ANN SURROGATE MODEL FOR EOS BSk20

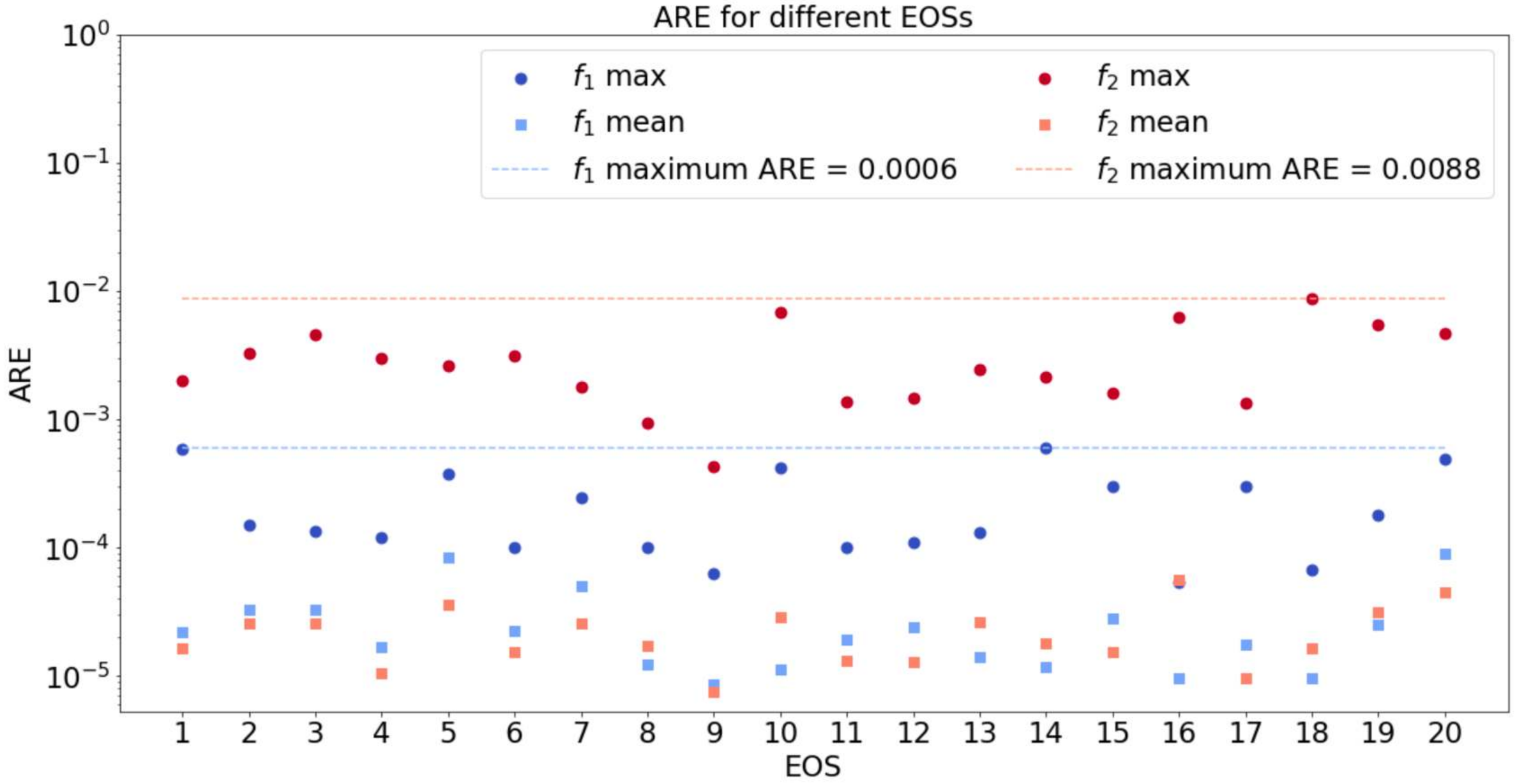
Loss function:
Mean Square Error (MSE)

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



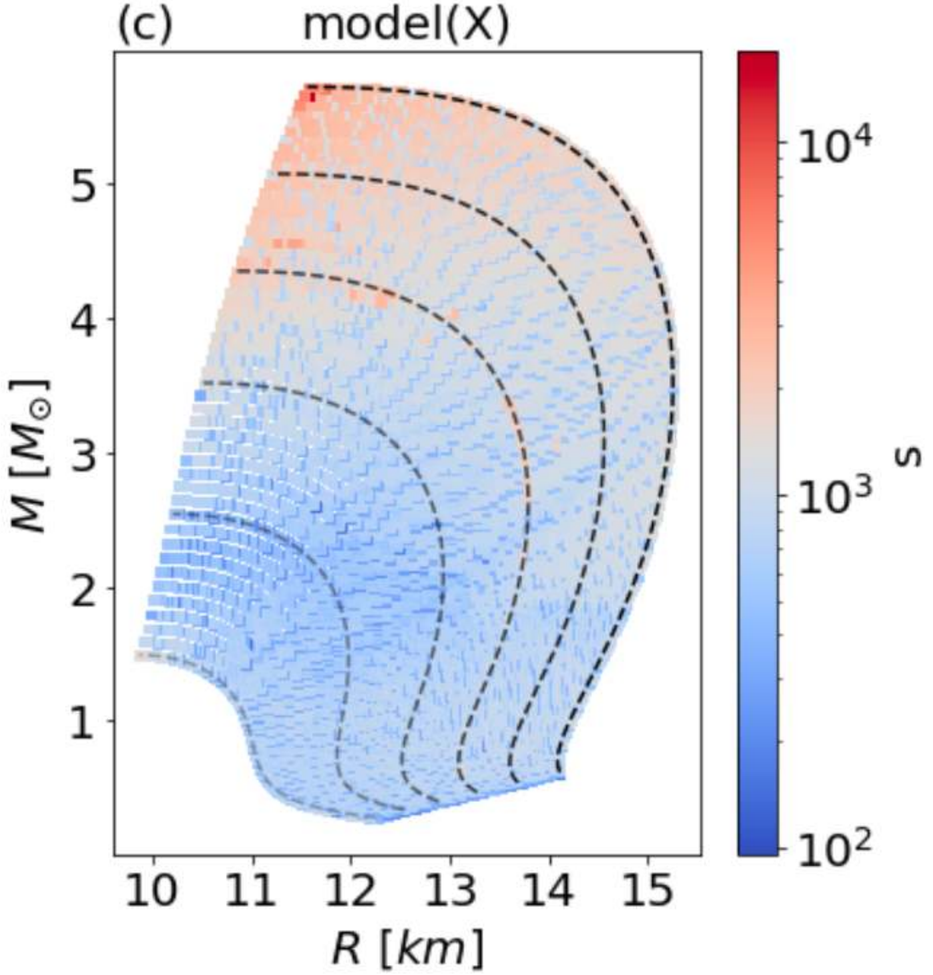
SPEED-UP ACHIEVED WITH THE ANN SURROGATE MODELS

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

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

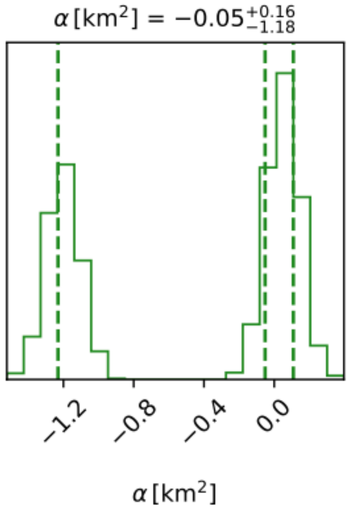
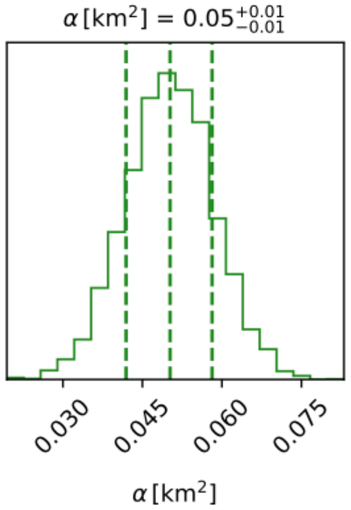
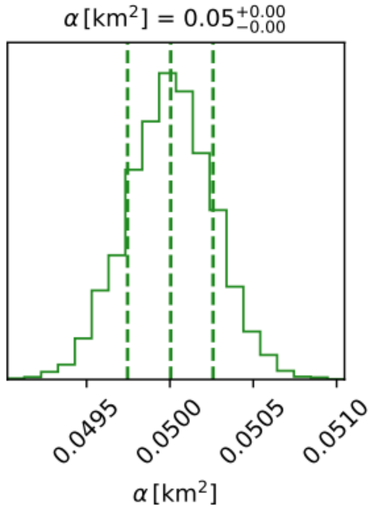
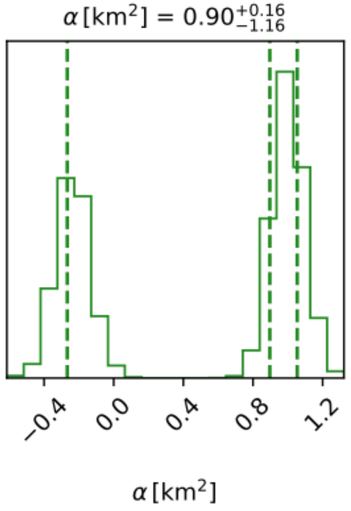
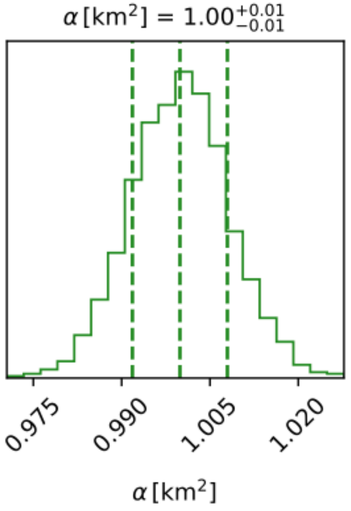
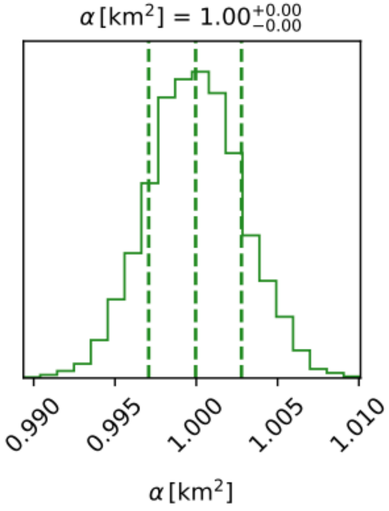
Numerical code run time:	(Mean) 1003.5 ms	(Min) 147.96 ms	(Max) 18122.4 ms
Output method	Speed Up (Mean)	Speed Up (Minimum)	Speed Up (Maximum)
model.predict(X)	25.12	0.97	464.56
model(X)	921.9	95.6	18102.1
model.predict($\mathbf{X}$)	31295.5	4614.5	565157.2

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



RECOVERY OF INJECTED α VALUES

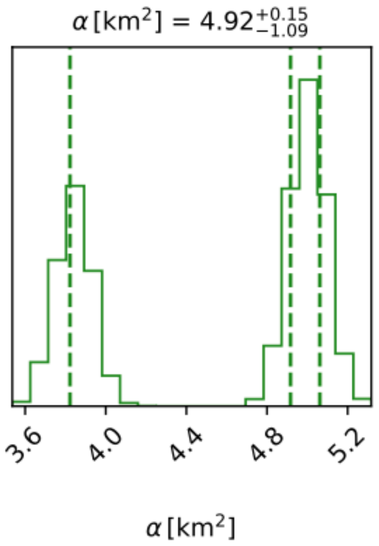
Hierarchical Bayesian Inference using astrophysical constraints for a set of 20 EOS:

	A+	Voyager	3G
$\alpha = 0.05\,\mathrm{km}^2$			
$\alpha = 1.0\,\mathrm{km}^2$			

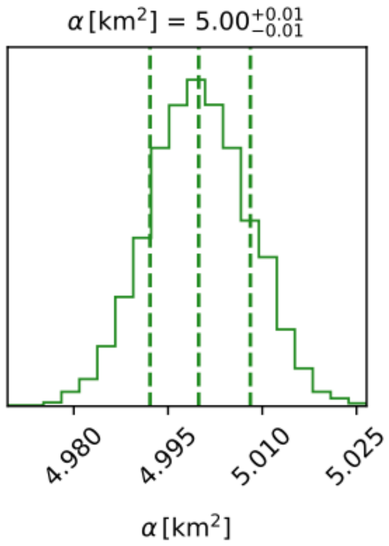
RECOVERY OF INJECTED α VALUES

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

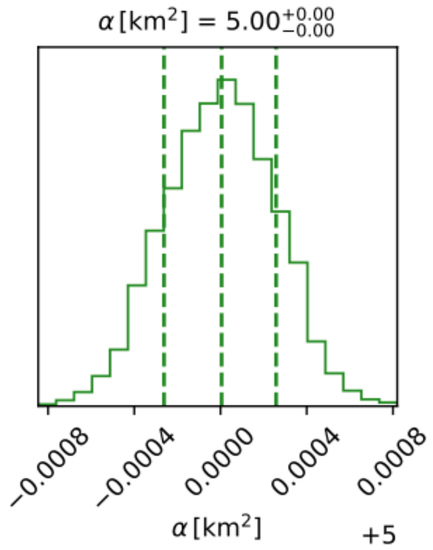
A+



Voyager



3G



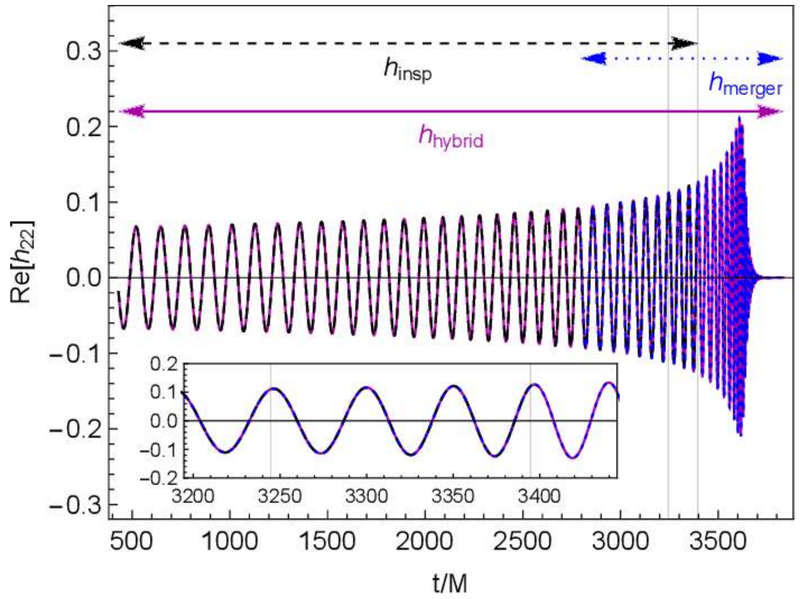
A+ : for an injection of α = 5 km^2, the GR value of α = 0 is excluded at the 3σ level.

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

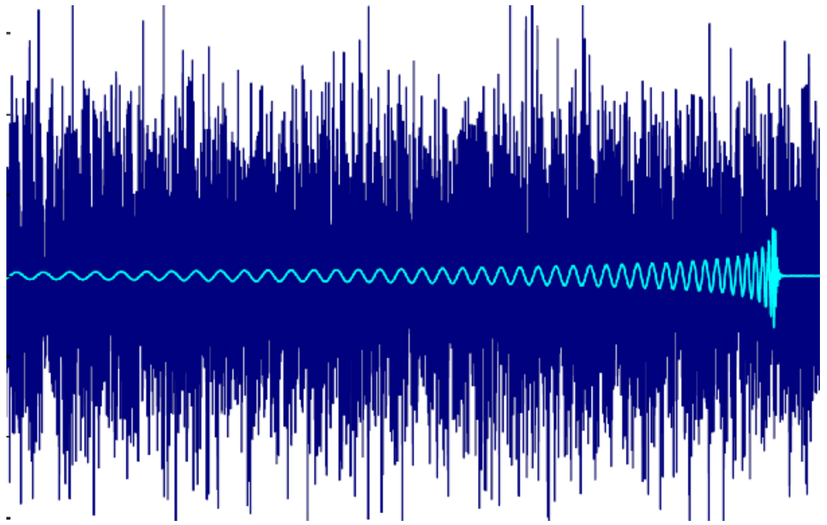
Biswas, Smyrniotis, Liodis, Stergioulas (2024)

GRAVITATIONAL WAVE ASTRONOMY

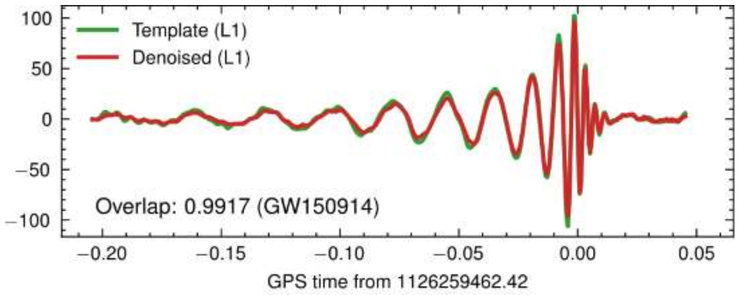
Waveform Modeling



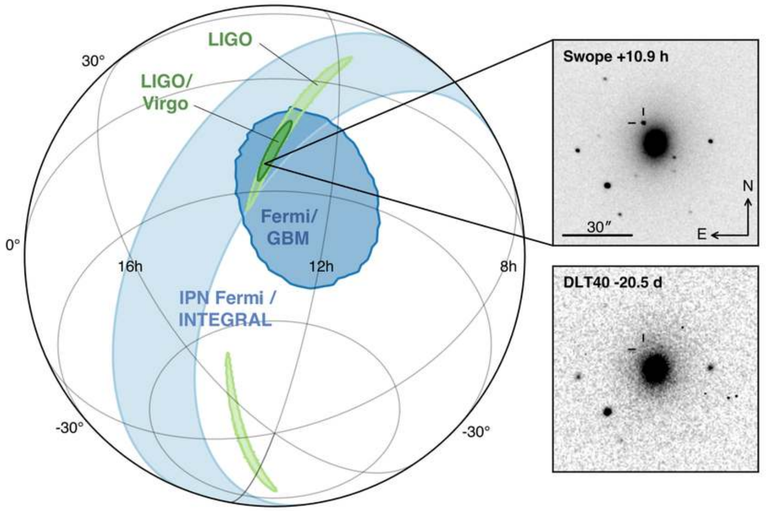
Detection



Denoising

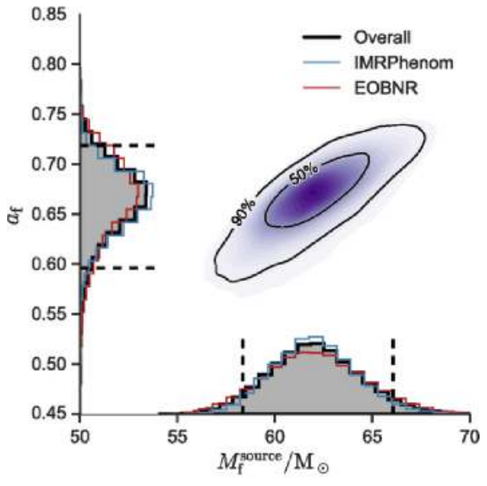


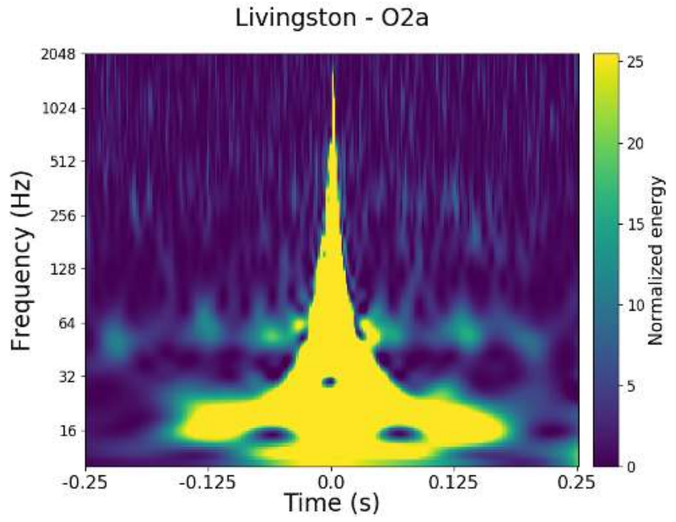
Glitch Classification



Parameter Estimation

Sky Localization





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

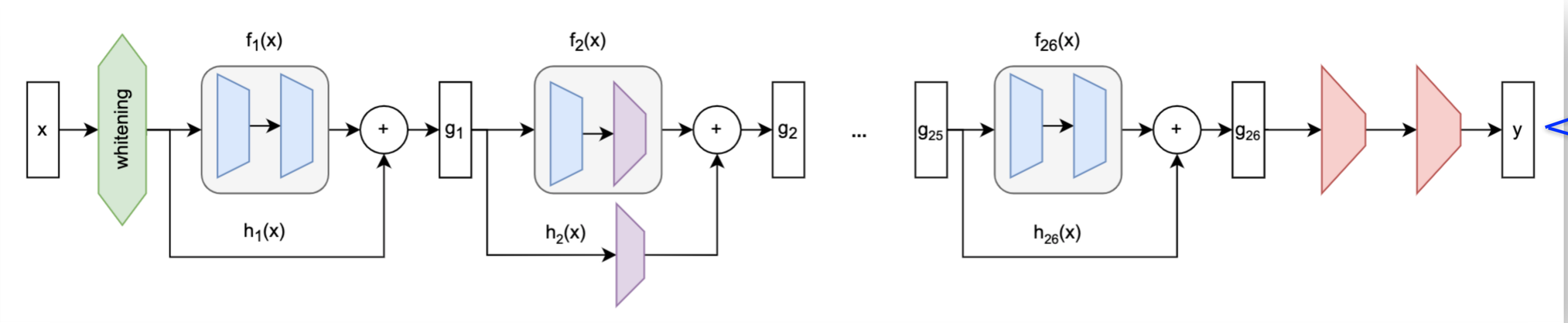
Mass range: $7M_\odot \leq M \leq 50M_\odot$

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)!

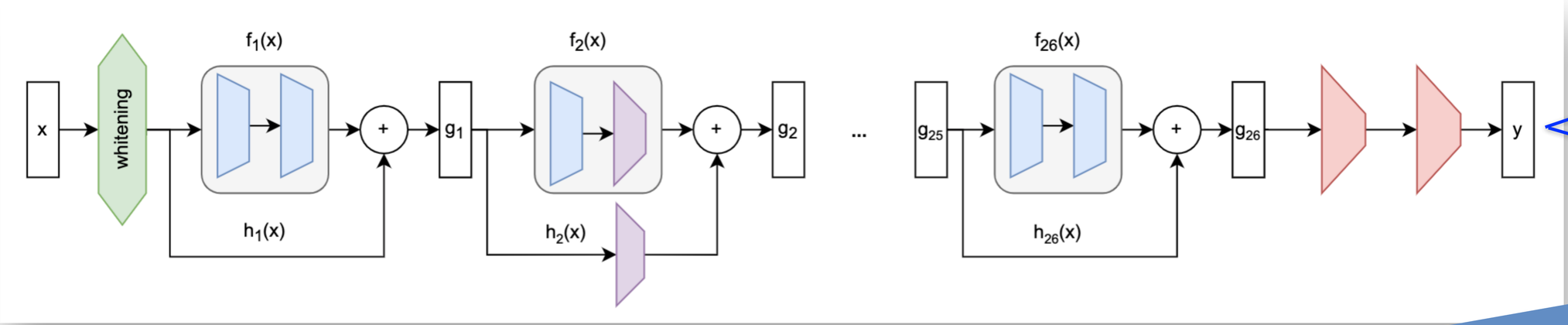


- Residual blocks with skip connections: g(**x**) = f(**x**) +h(**x**)
- h(**x**) = **x** *or* a strided convolutional layer
- After each convolutional layer: batch normalization + ReLU activation
- Mini-batch size of 400 segments
- Adam optimizer for back propagation
- Objective function = regularized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

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)!



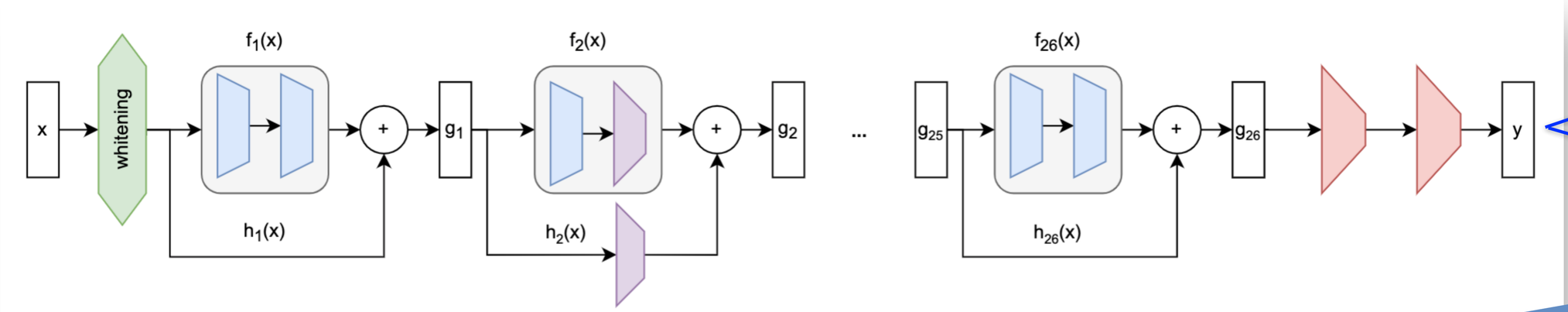
- Residual blocks with skip connections: g(**x**) = f(**x**)
- h(**x**) = **x** *or* a strided convolutional layer
- After each convolutional layer: batch no
  ReLU activation
- Mini-batch size of 400 segments
- Adam optimizer for back propagation
- Objective function = regularized binary cross entropy
- Gradient problem solved: much deeper networks!

Gradient problem solved!
-> much deeper networks

			Input D
			2×2048
		✓	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(**x**)
- h(**x**) = **x** *or* a strided convolutional layer
- After each convolutional layer: batch no
  ReLU activation
- Mini-batch size of 400 segments
- Adam optimizer for back propagatio
- Objective function = regularized bina
- Gradient problem solved: much deep works!

Gradient problem solved!
-> much deeper networks

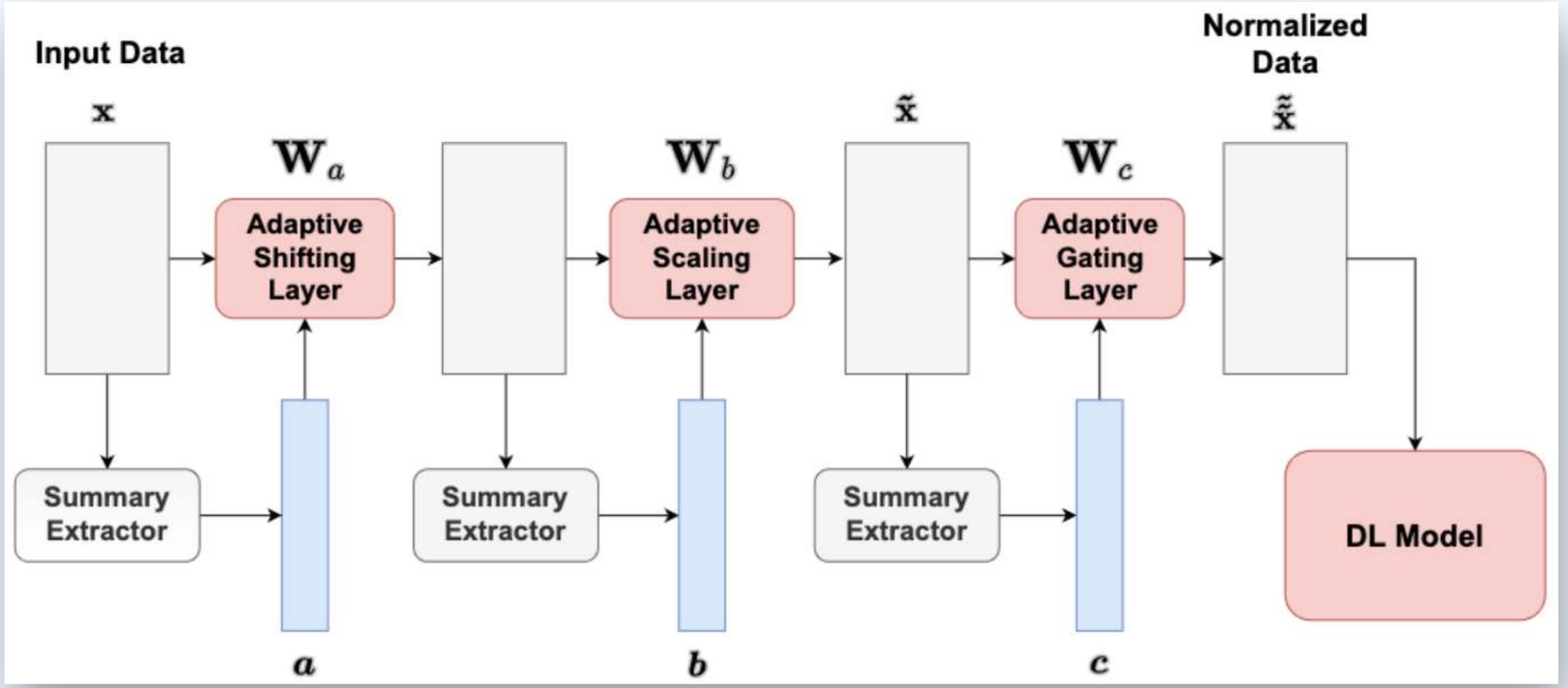
1-D ResNet-54 now used by most
ML GW detection codes

			Input D
			2×2048
		✓	8×2048
	16		
1			
2			
			64×256
			64×128
	64	✓	64×128
	64		64×64
5	32		64×64
3	16		32×64

ADAPTIVE INPUT NORMALIZATION

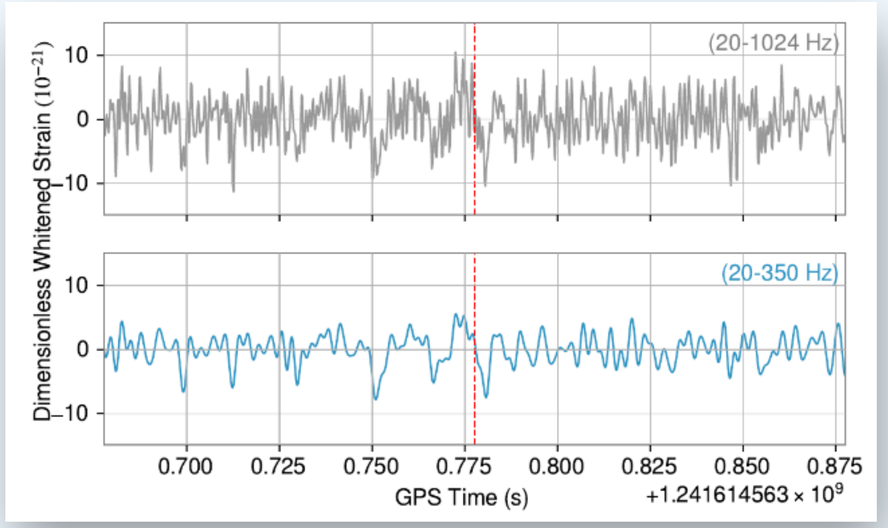
Deep Adaptive Input Normalization (DAIN) (Passalis et al. 2019)

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

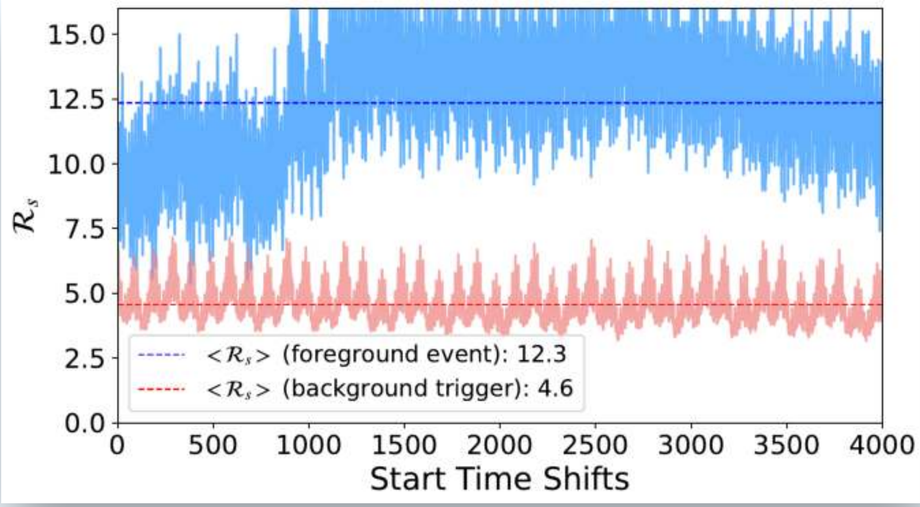


MAIN ENHANCEMENTS:

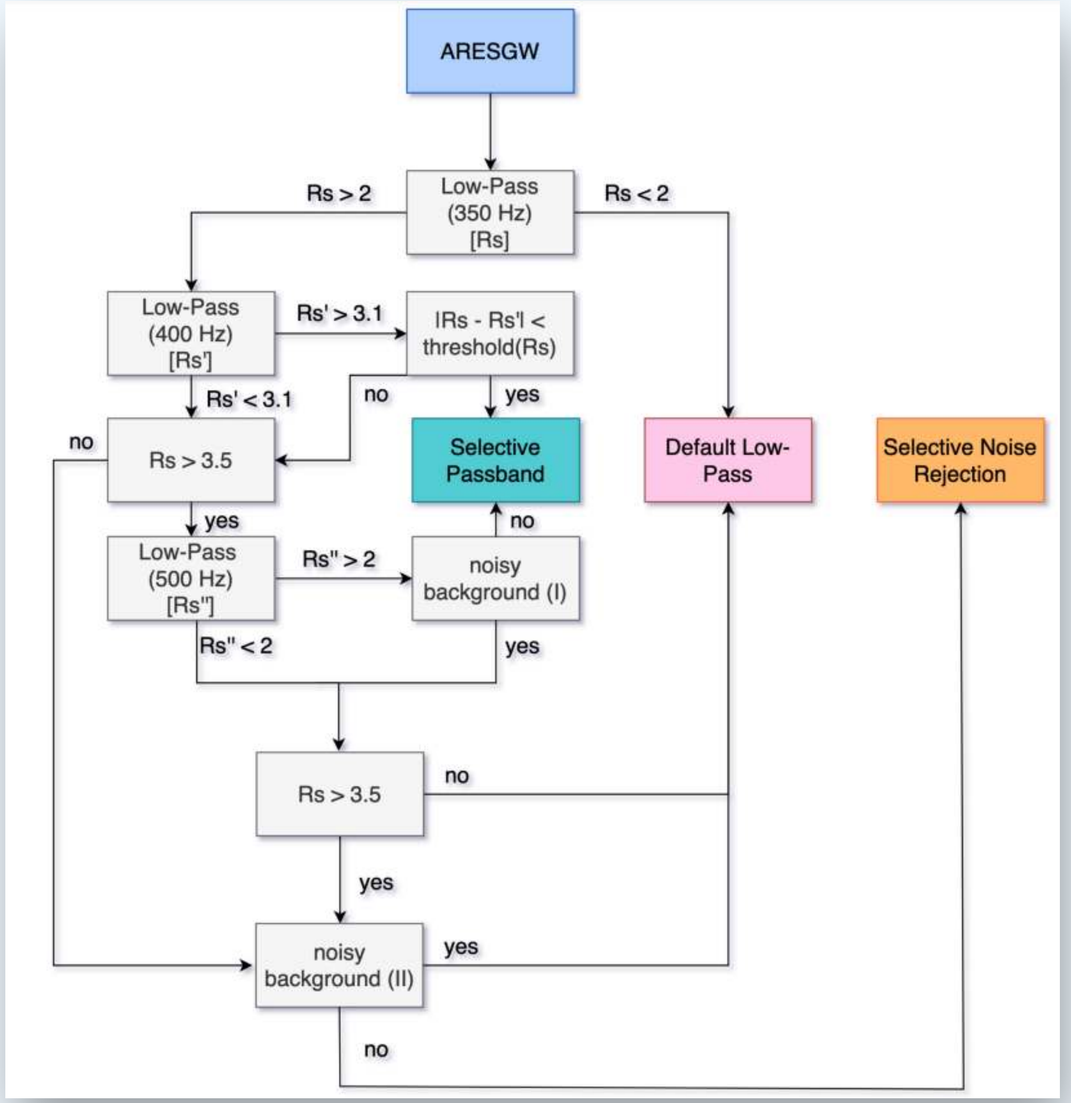
1) **Training** on **low-pass filtered** data



2) **Ensemble-averaged** ranking statistic



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



RECONSTRUCTED WAVEFORMS FOR NEW EVENTS

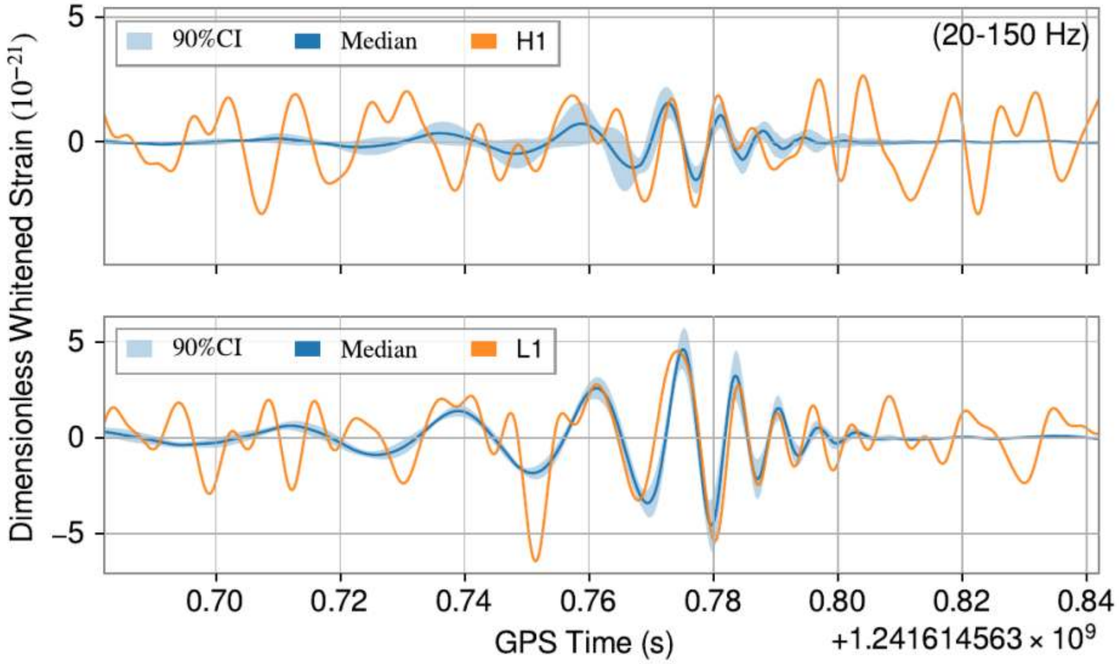


FIG. 35: Whitened, bandpassed strain data and reconstructed waveform for the new event GW190511_125545 identified by AresGW.

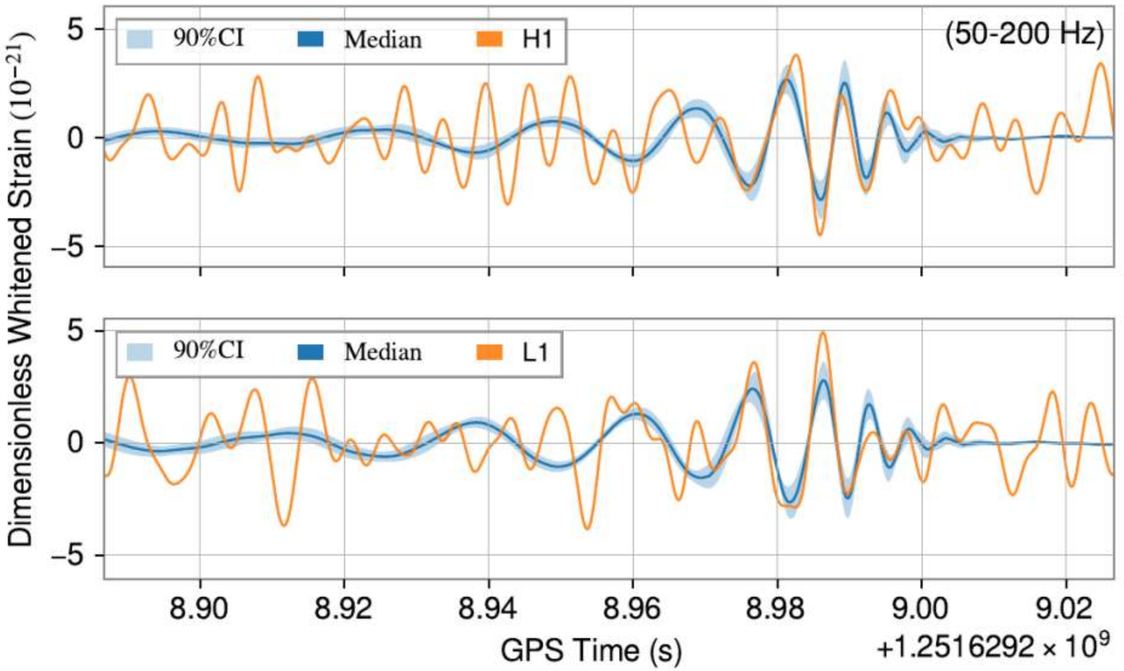
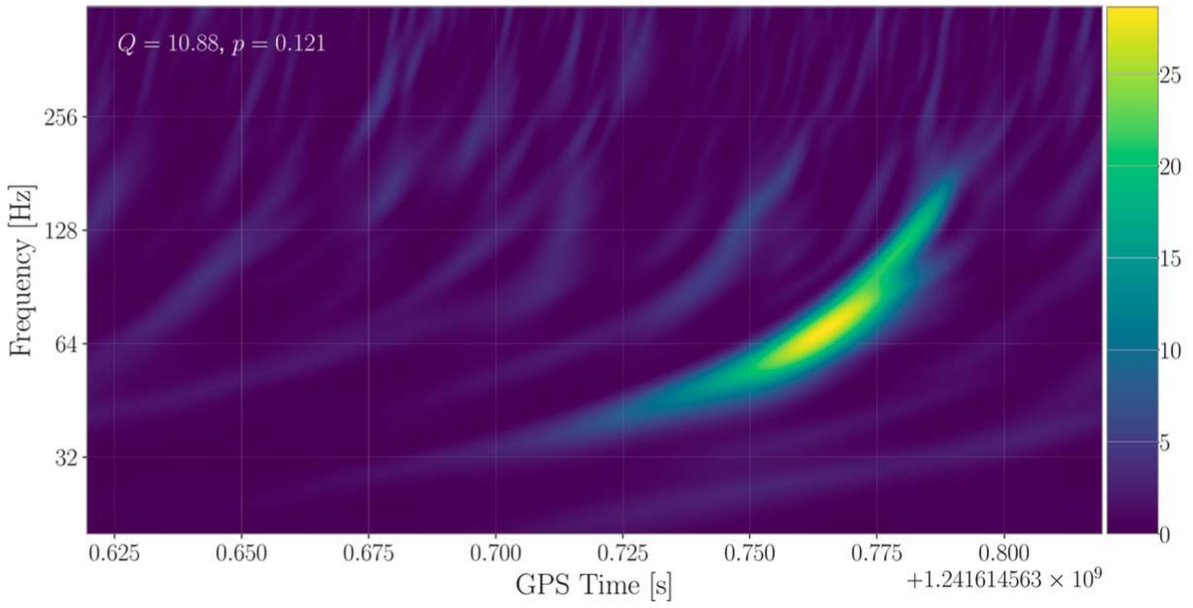
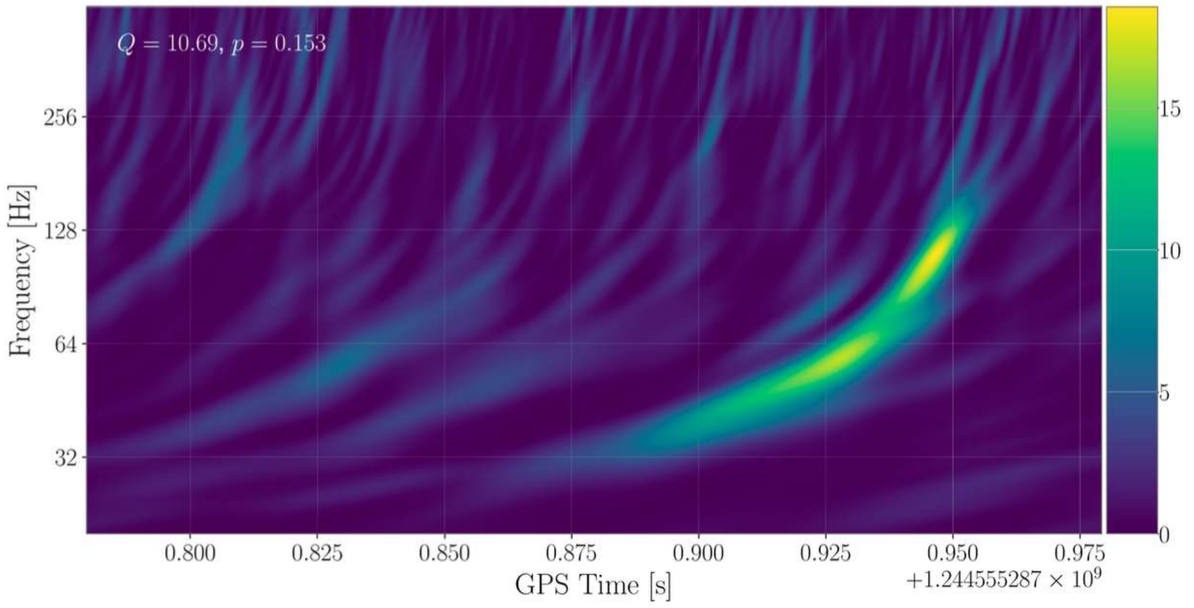


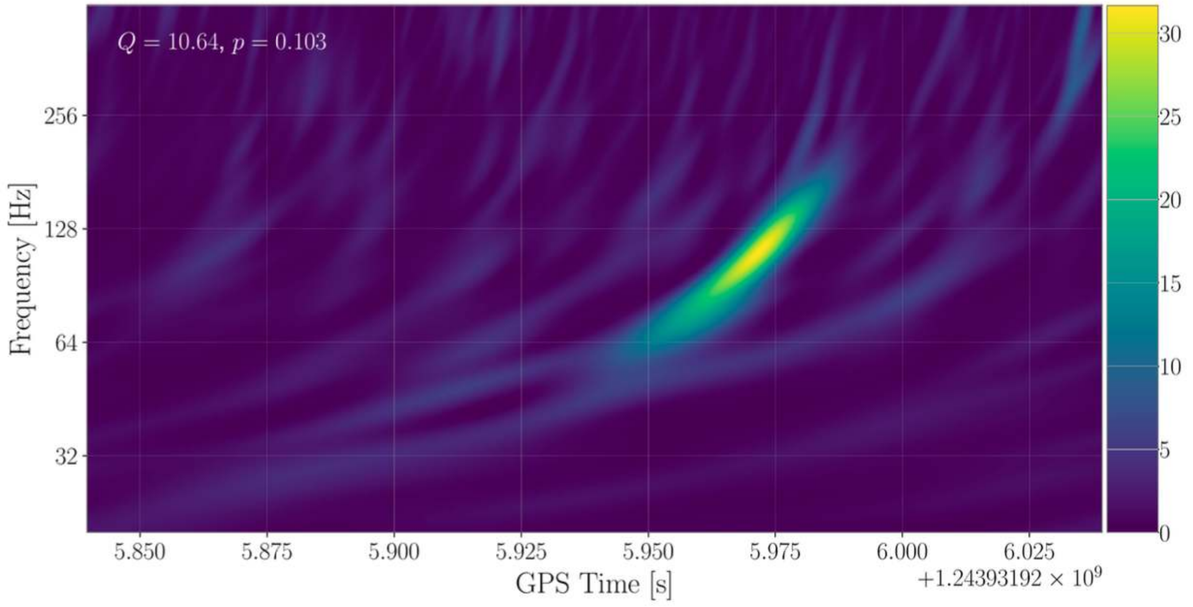
FIG. 38: Same as Fig. 35, but for the new event GW190904_104631.

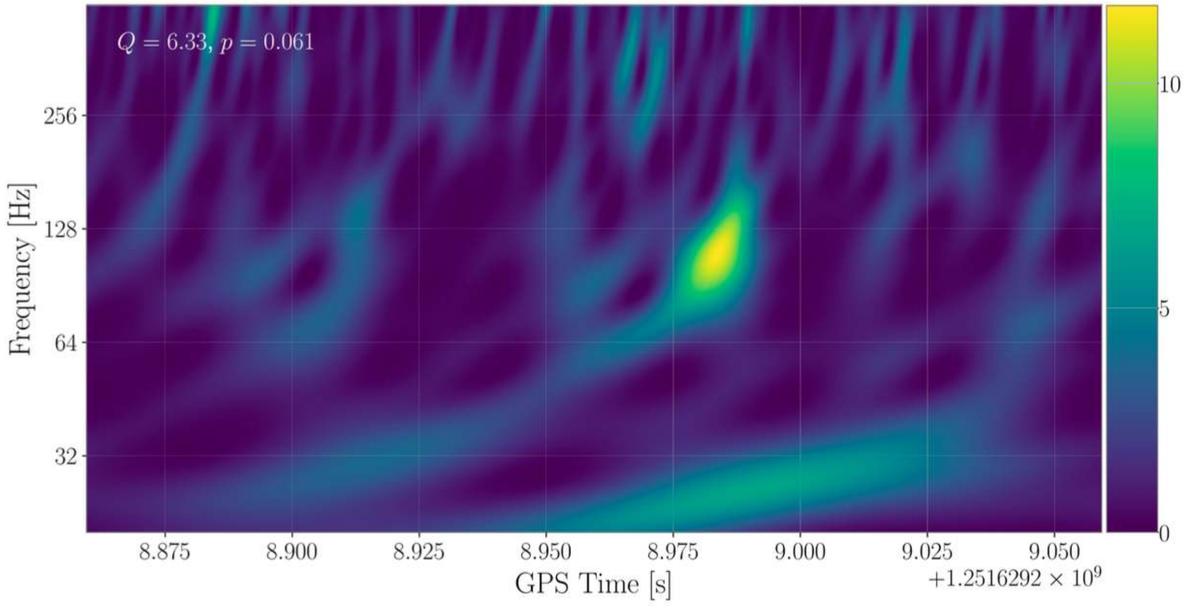
Koloniari, Koursoumpa, Nousi, Lampropoulos, Passalis, Tefas, Stergioulas (2025)

ARESGW NEW CANDIDATE EVENTS

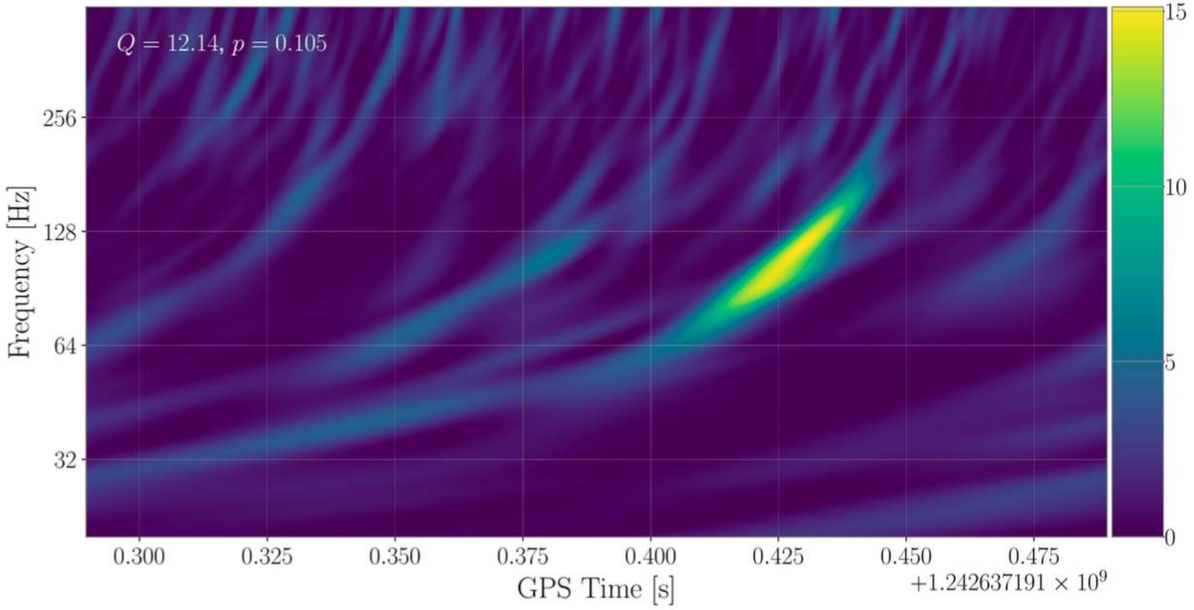


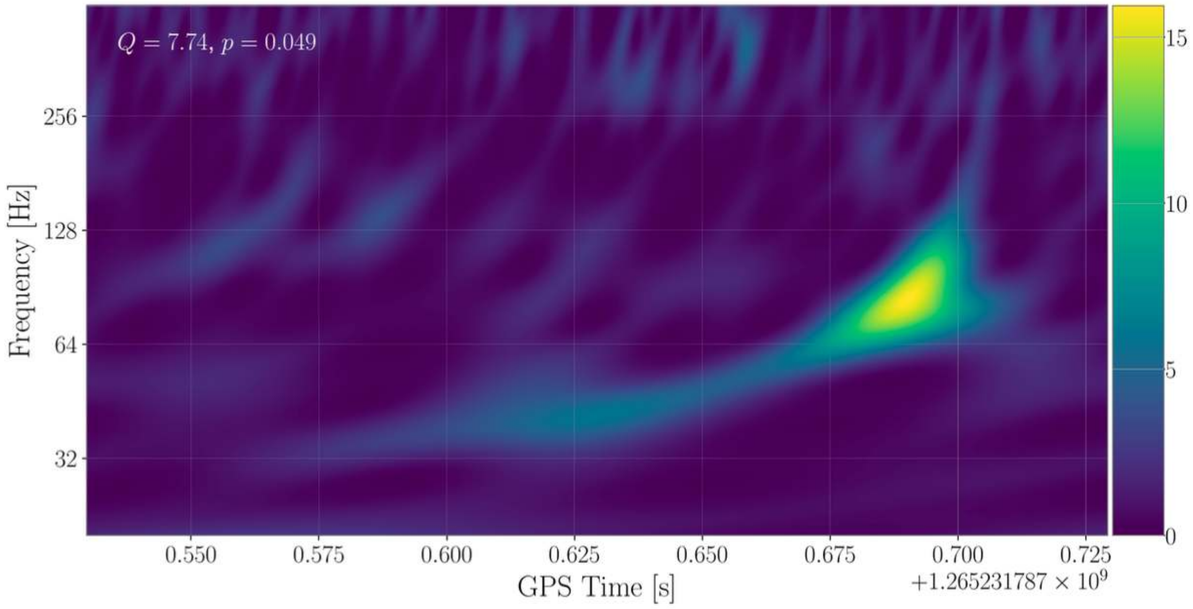


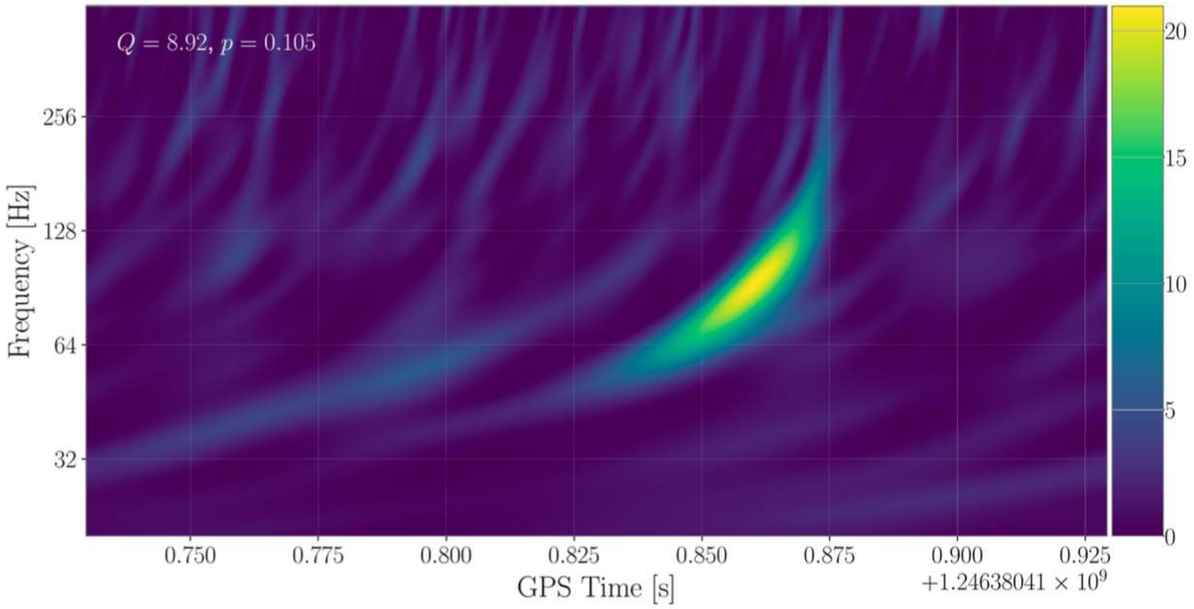


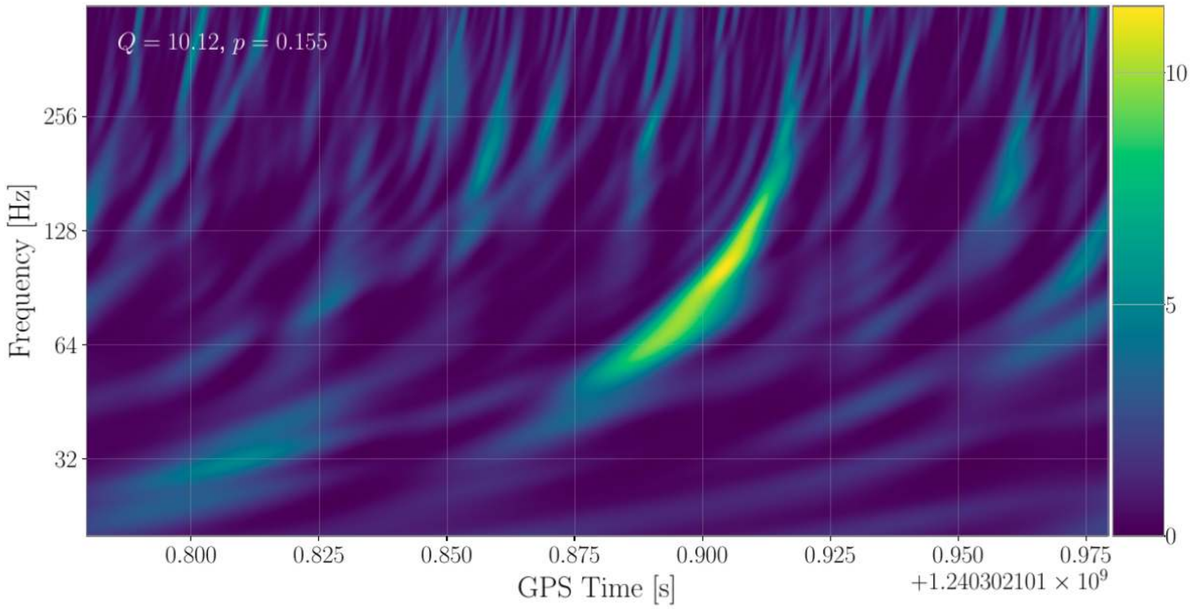


ARESGW NEW CANDIDATE EVENTS









THANK YOU FOR YOUR ATTENTION