SEARCHING GRAVITATIONAL WAVE SIGNALS WITH AUTOREGRESSIVE APPROACH AND DEEP LEARNING

DAVID C. Y. HUI

DEPARTMENT OF ASTRONOMY & SPACE SCIENCE
CHUNGNAM NATIONAL UNIVERSITY
SOUTH KOREA

65th Workshop on Gravitational Waves and Numerical Relativity   8 June 2022

COLLABORATORS

- **Sangin Kim (CNU, Korea)**
- **Jianqi Yan (MUST, Macau)**
- Kwangmin Oh (CNU, Korea)
- Albert Kong (NTHU, Taiwan)
- Lupin Lin (NCKU, Taiwan)
- Ray Kwan Lok Li (NCKU, Taiwan)
- Alex Leung (HKU, Hong Kong)
- Shengda Luo (MUST, Macau)
- Hisaaki Shinkai (OIT, Japan)

OUTLINE

- Development of GW astronomy
- Autoregressive Approach in GW detection
- Improving deep learning classification of GW spectrograms with GAN

EXISTENCE OF GW

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$$

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$$

$$(h_{\mu\nu} \ll 1)$$

$$\longrightarrow \quad \frac{\partial^2 h_{\mu\nu}}{\partial x_\alpha \partial x^\alpha} = 0$$

- In a limit of small disturbance from a flat spacetime, Einstein field equation reduces to a linear wave equation.
- Implication on the disturbance can be seen from the constant appears in Einstein field equation.

"STIFFNESS" OF SPACETIME

$$\frac{c^4}{8\pi G} \longrightarrow$$

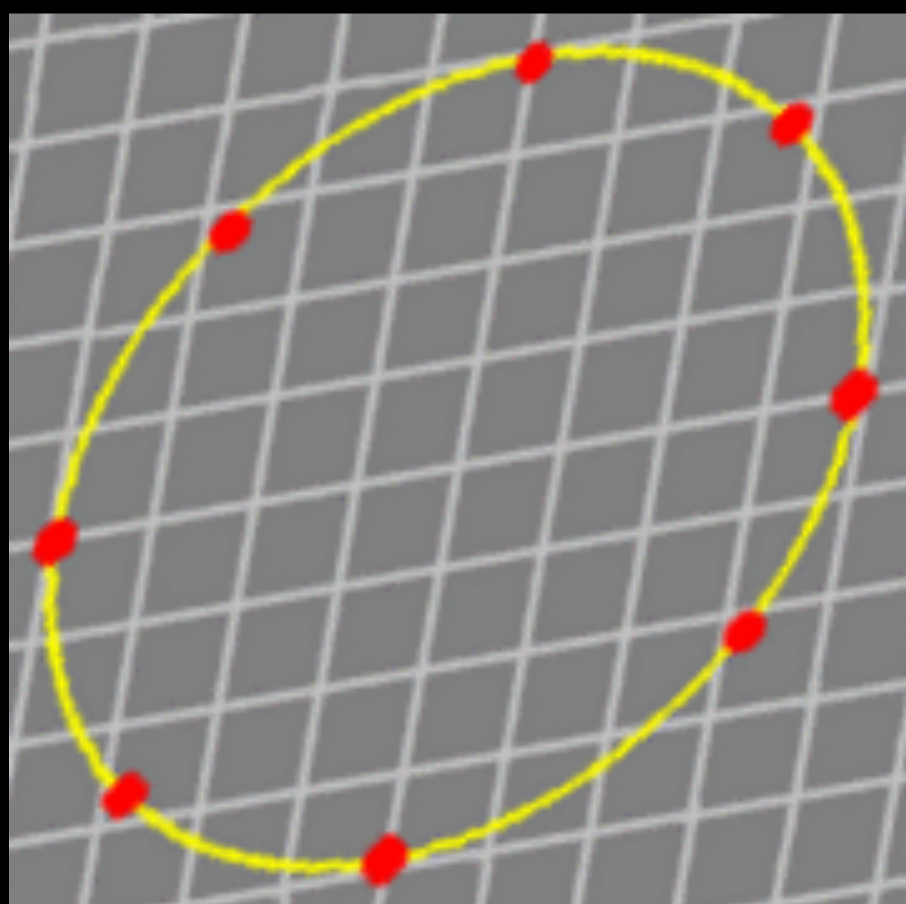
Require a stress of $\sim 10^{43}\ \mathrm{Nm^{-2}}$ to give a curvature of $\sim 1\mathrm{m}^{-2}$

- The amplitude of GW is hence very small.
- GW sources must involve huge masses and/or catastrophic events.
- Despite the small amplitude, a large amount of energy can be carried by GW.

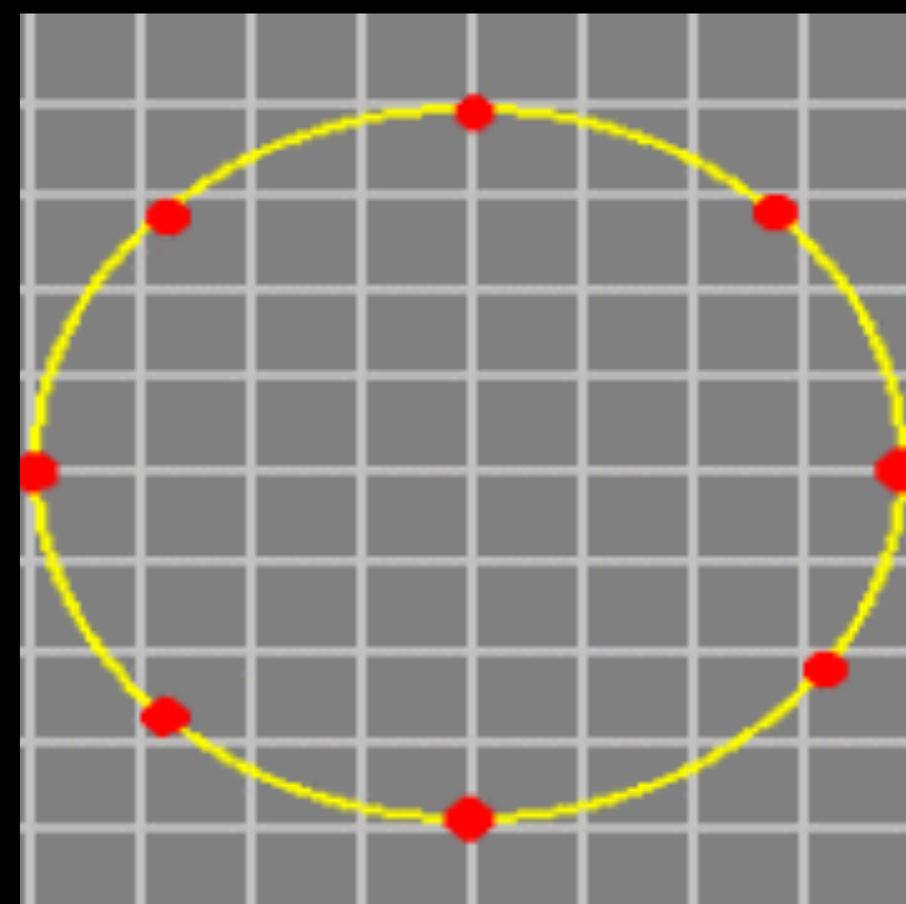
TWO POLARIZATION STATES

- Passage of GW can be revealed by investigating the motion of test masses lying in its path of propagation

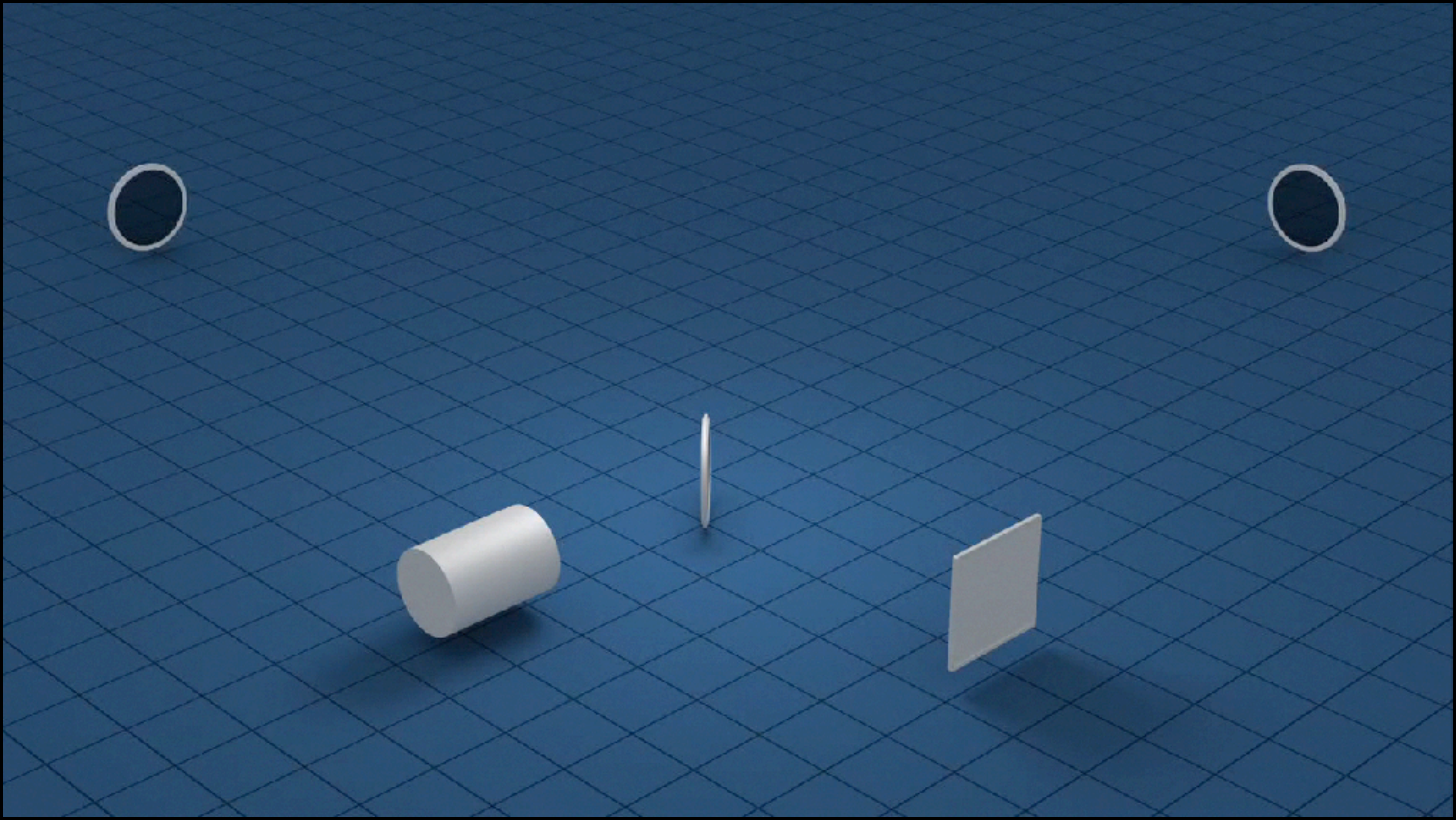
- Interferometry is most suitable for detecting GW



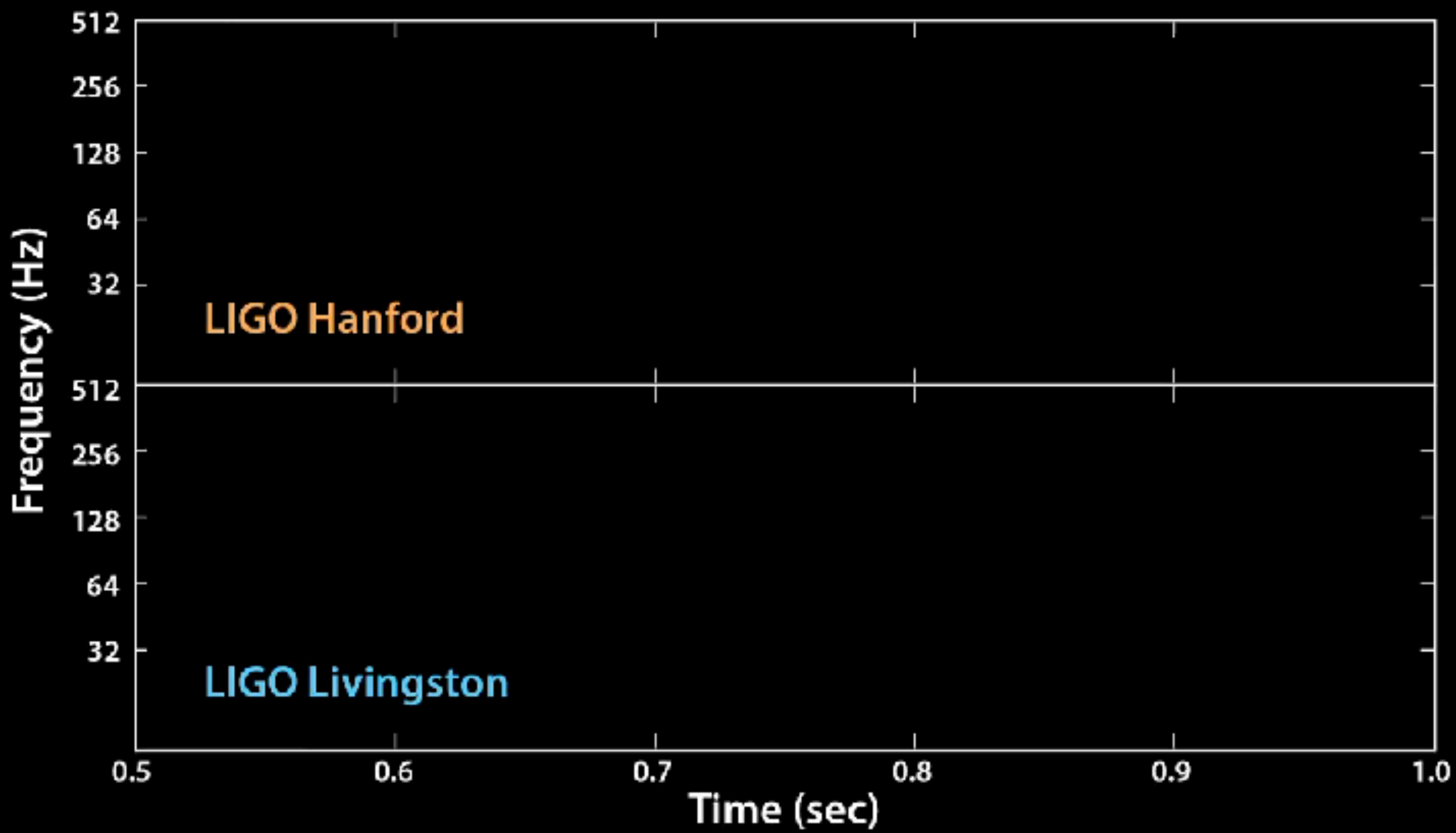
$h_{\times}$



h_{+}



FIRST BLACK HOLE MERGER - GW150914



On September 14, 2015, LIGO observed gravitational waves from the merger of two black holes, each about 30 times the mass of our sun. The powerful event releases $\sim 3M_{\odot}c^2$ within a fraction of a second.

90 GW events have been detected so far in GWTC-3

PROGRESS OF GW ASTRONOMY

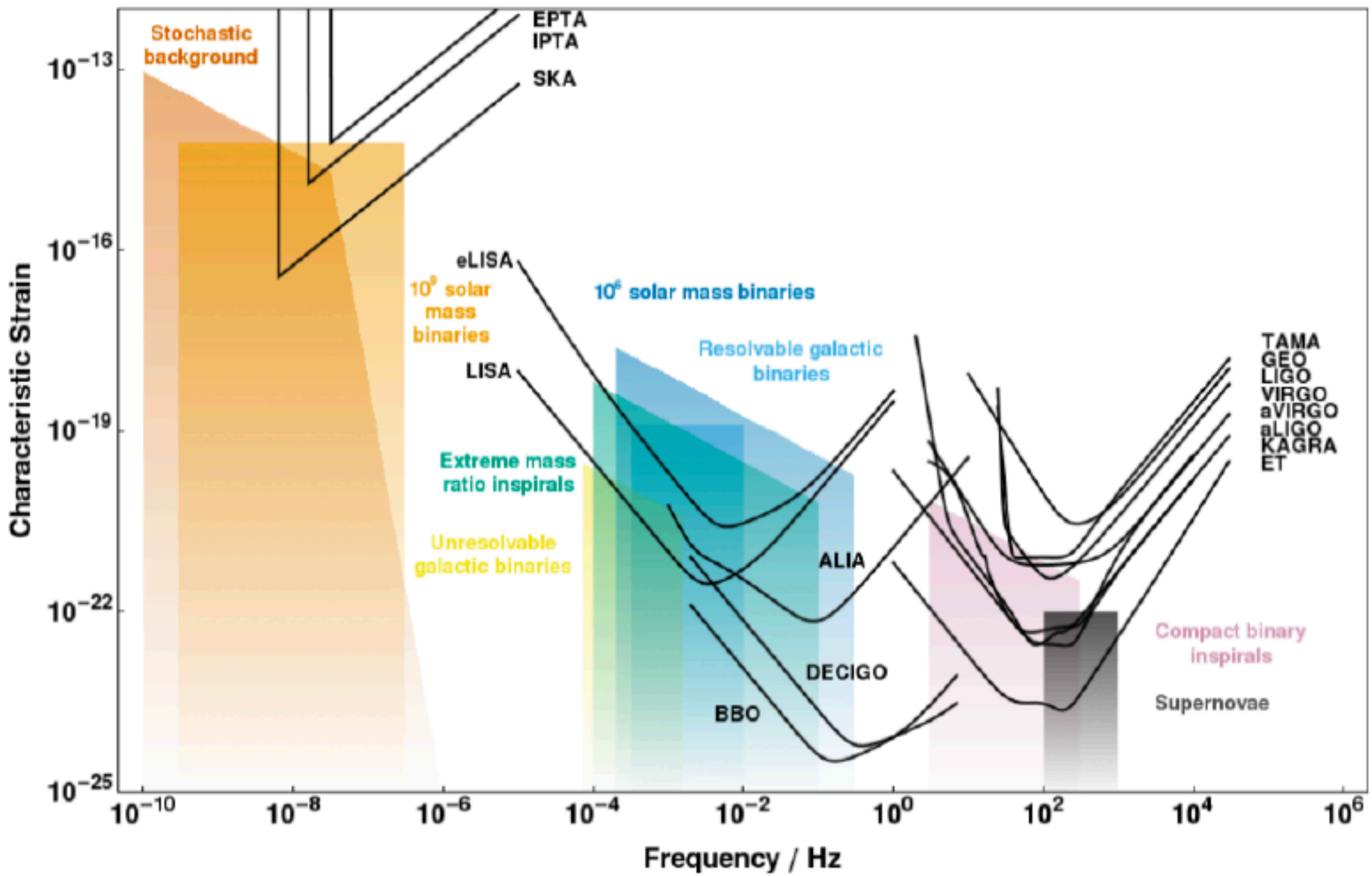
- Increasing the number of observatories
- Installation of space-based/Galaxy-based observatories
- Improving the sensitivity
- **Improving the performance in signal processing**

INTERNATIONAL GW NETWORK

- More precise source localization
- More certain parameter estimation
- Lower false alarm rate
- More information of polarization

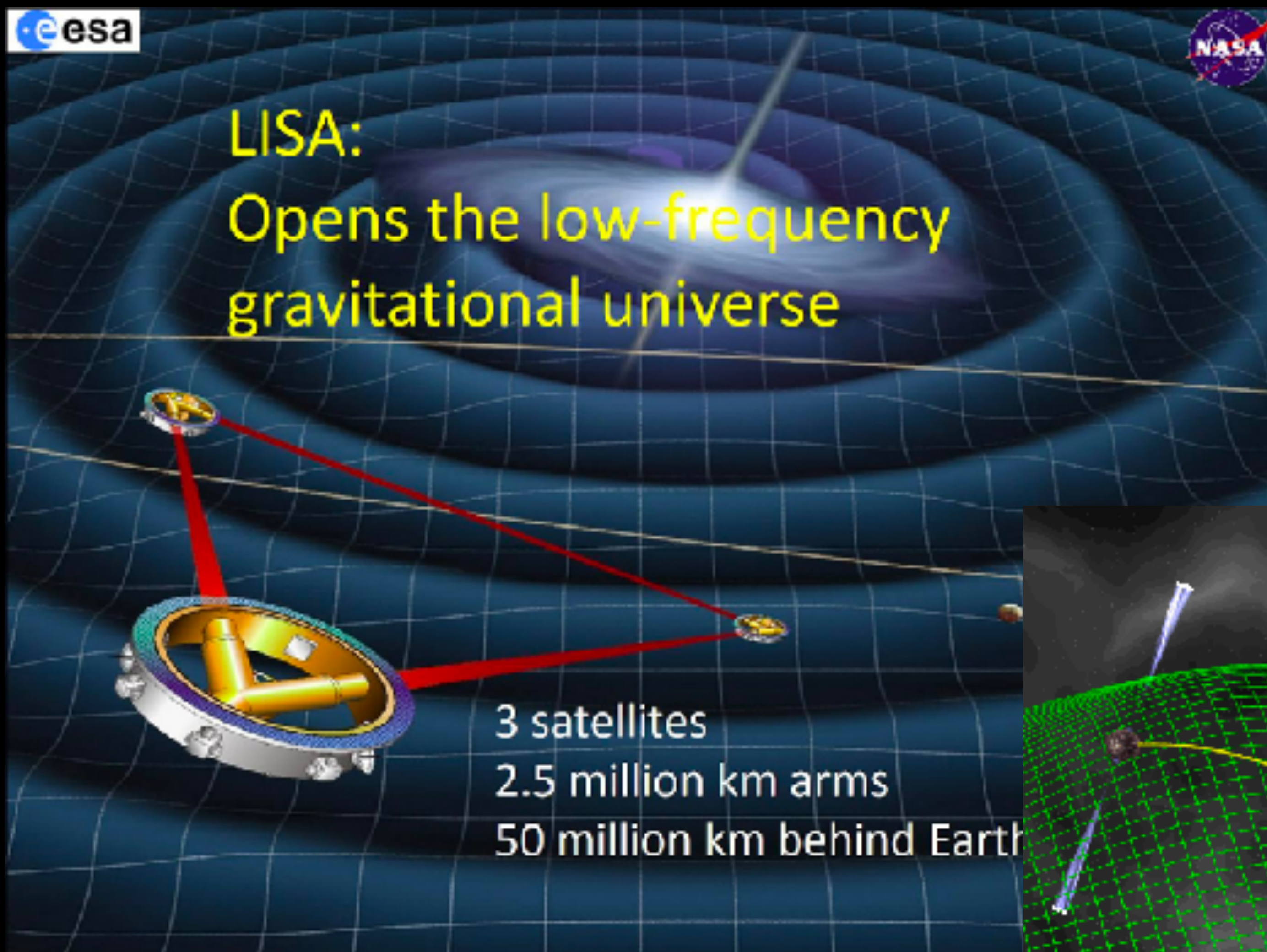
INTERNATIONAL GW NETWORK

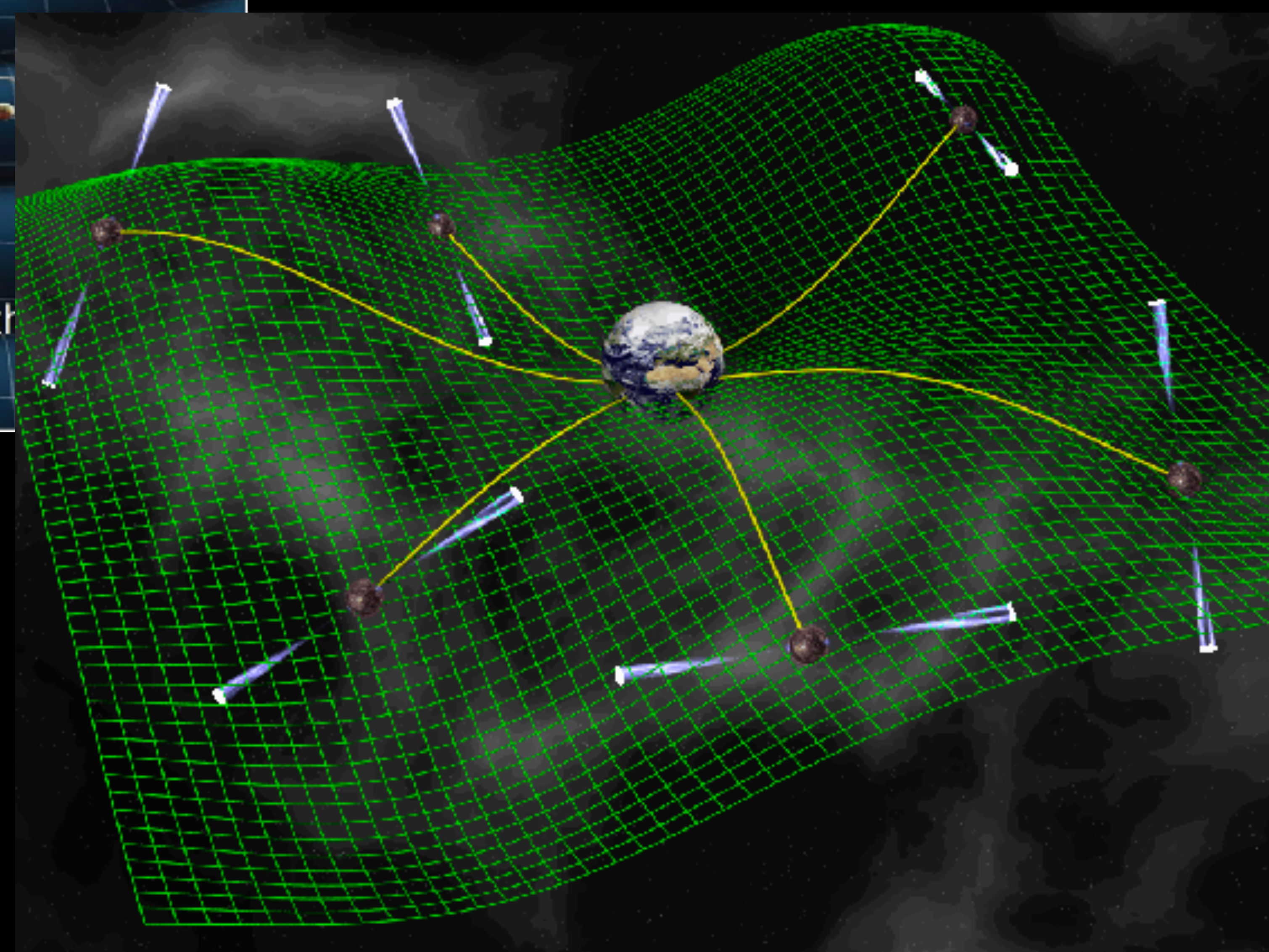
Hisaaki Shinkai (Osaka Institute of Technology) "GW physics, Status of KAGRA" Cosmology From Home 2021



Stochastic background
EPTA
IPTA
SKA
eLISA
10^9 solar mass binaries
10^6 solar mass binaries
Resolvable galactic binaries
LISA
Extreme mass ratio inspirals
Unresolvable galactic binaries
ALIA
DECIGO
BBO
TAMA
GEO
LIGO
VIRGO
aVIRGO
aLIGO
KAGRA
ET
Compact binary inspirals
Supernovae
Characteristic Strain
Frequency / Hz

SPACE/GALAXY BASED OBSERVATORIES





SENSITIVITY CURVE

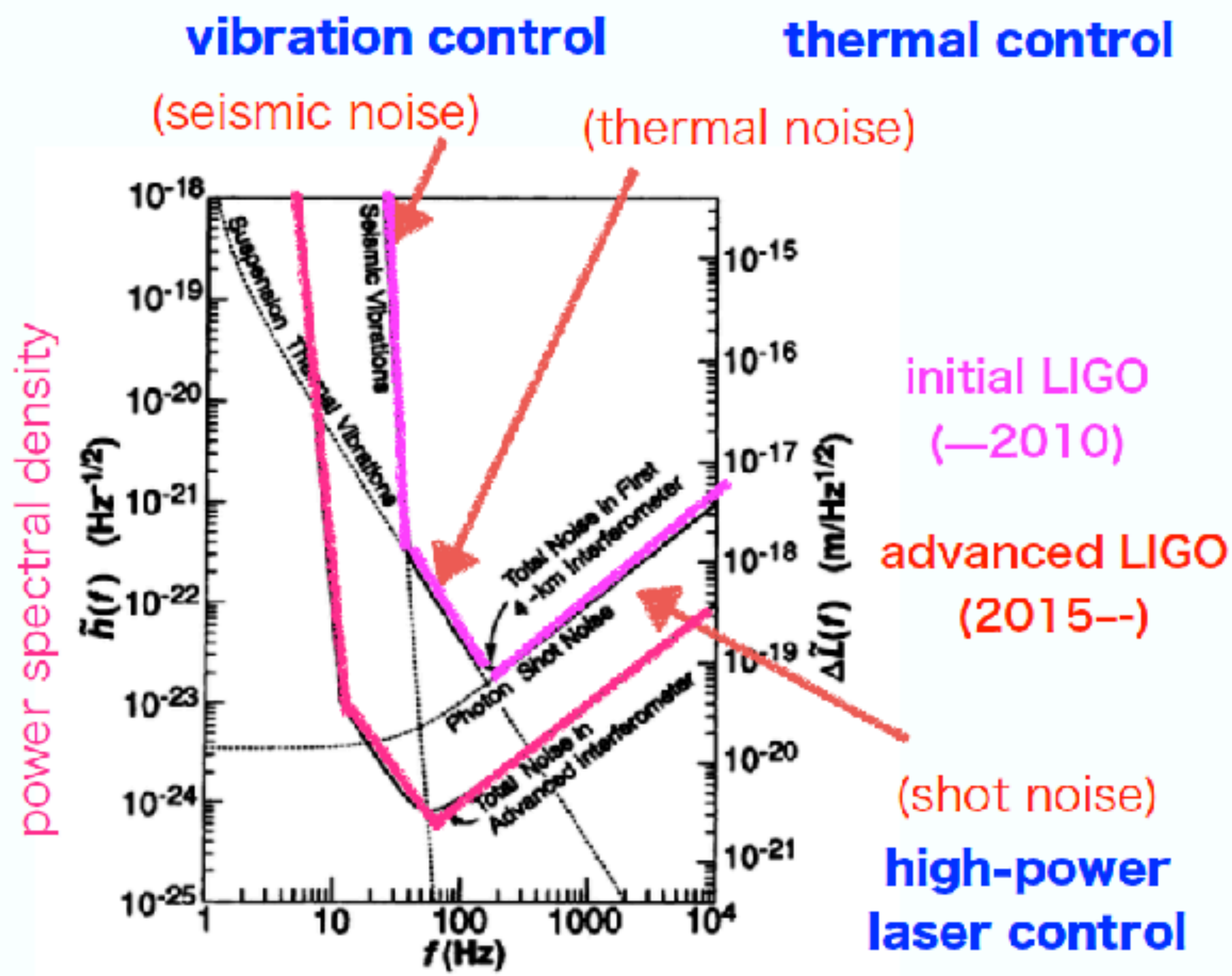
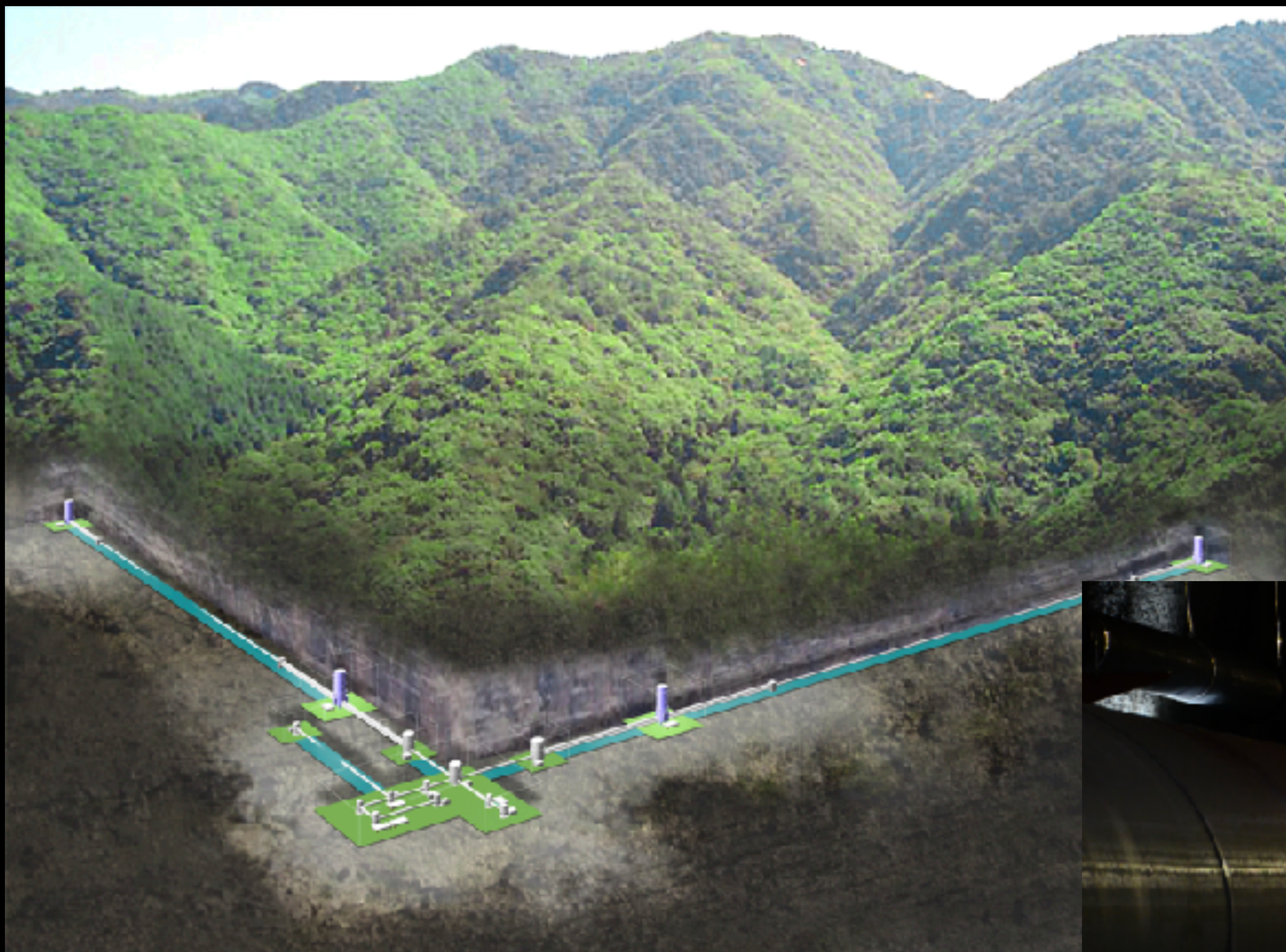


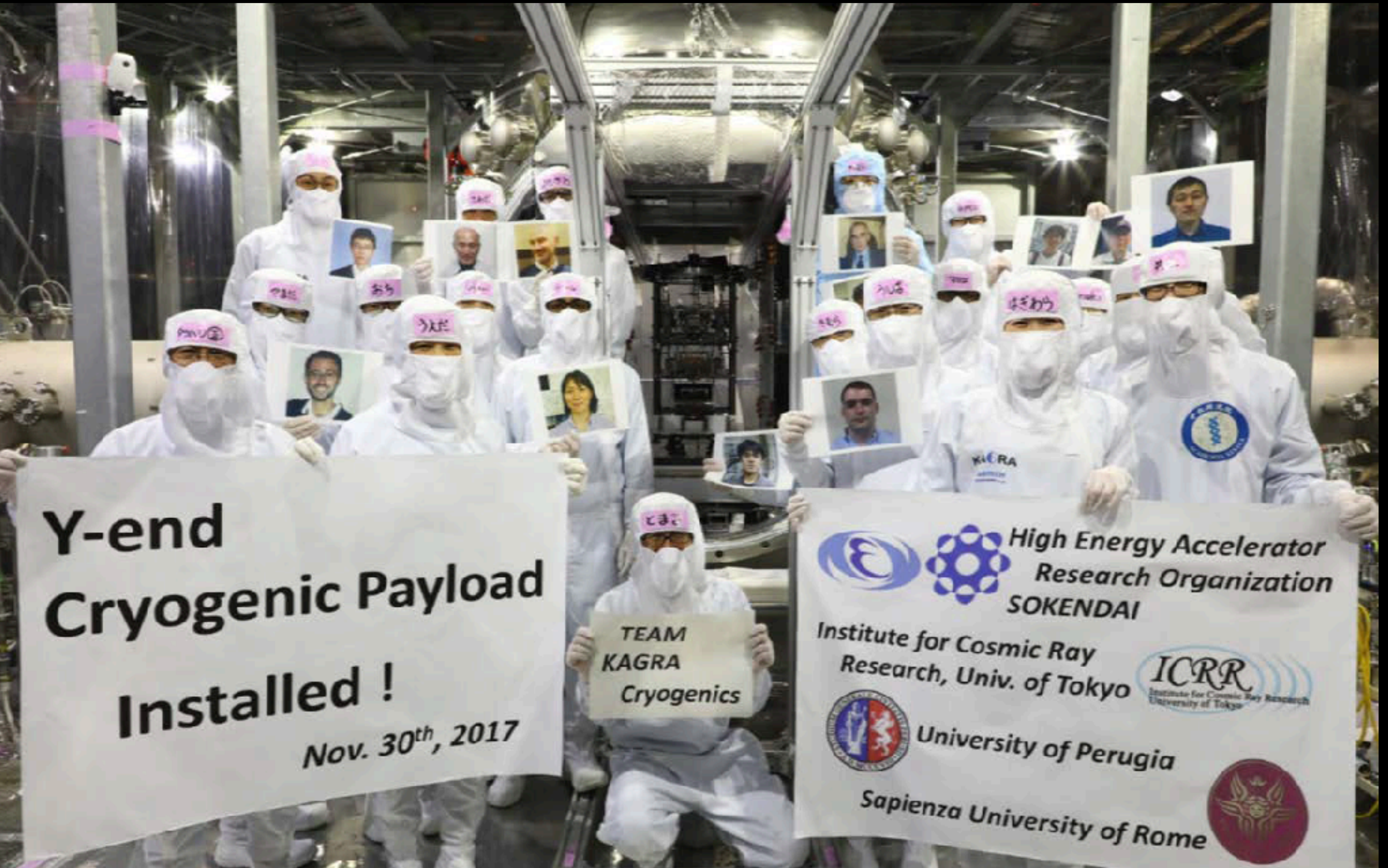
Fig. 7. The expected total noise in each of LIGO's first 4-km interferometers (upper solid curve) and in a more advanced interferometer (lower solid curve). The dashed curves show various contributions to the first interferometer's noise.

Abramovici et al. 1992 Science 256 325

KAGRA

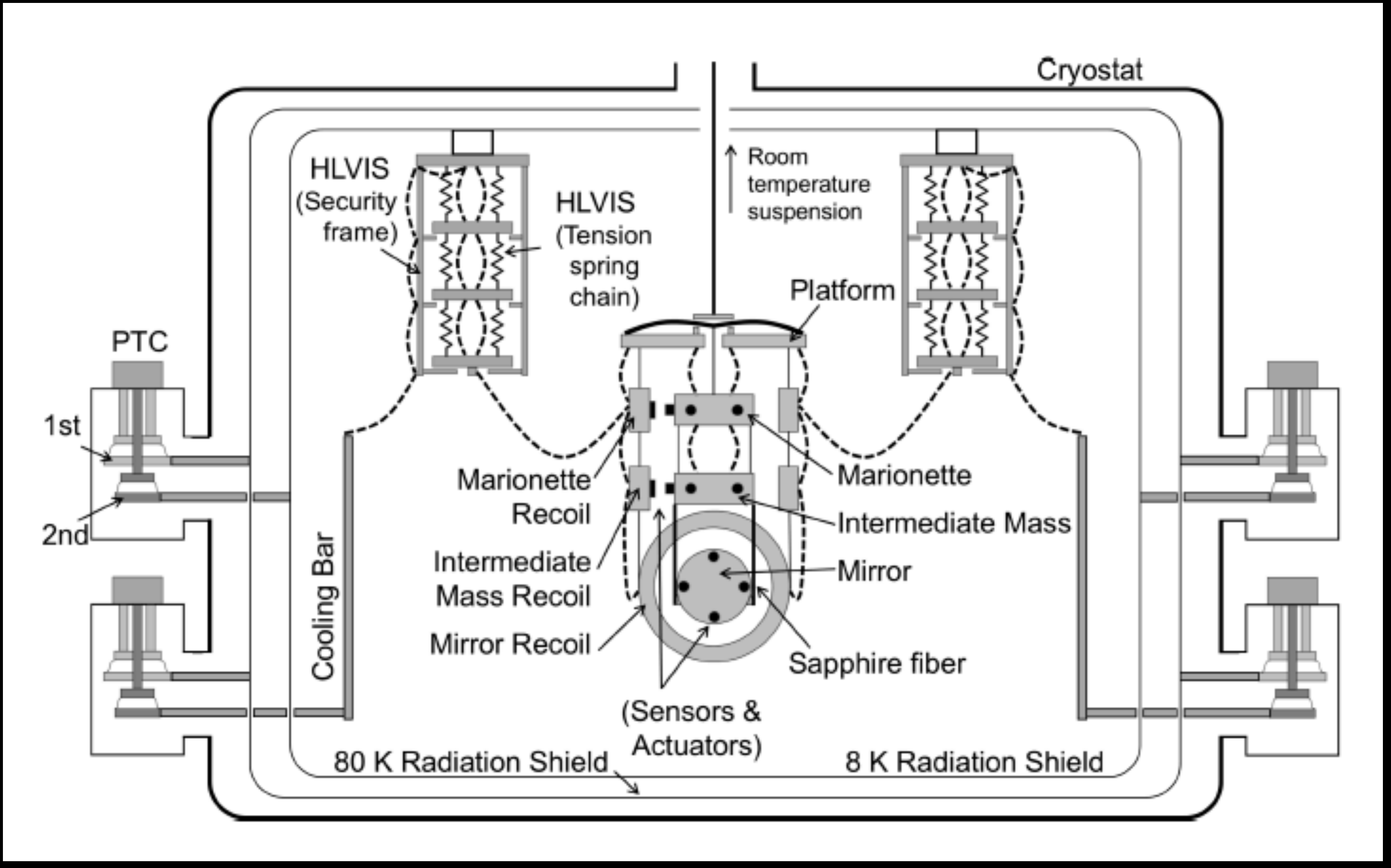
Underground Cryogenic Fabry-Perot Interferometer





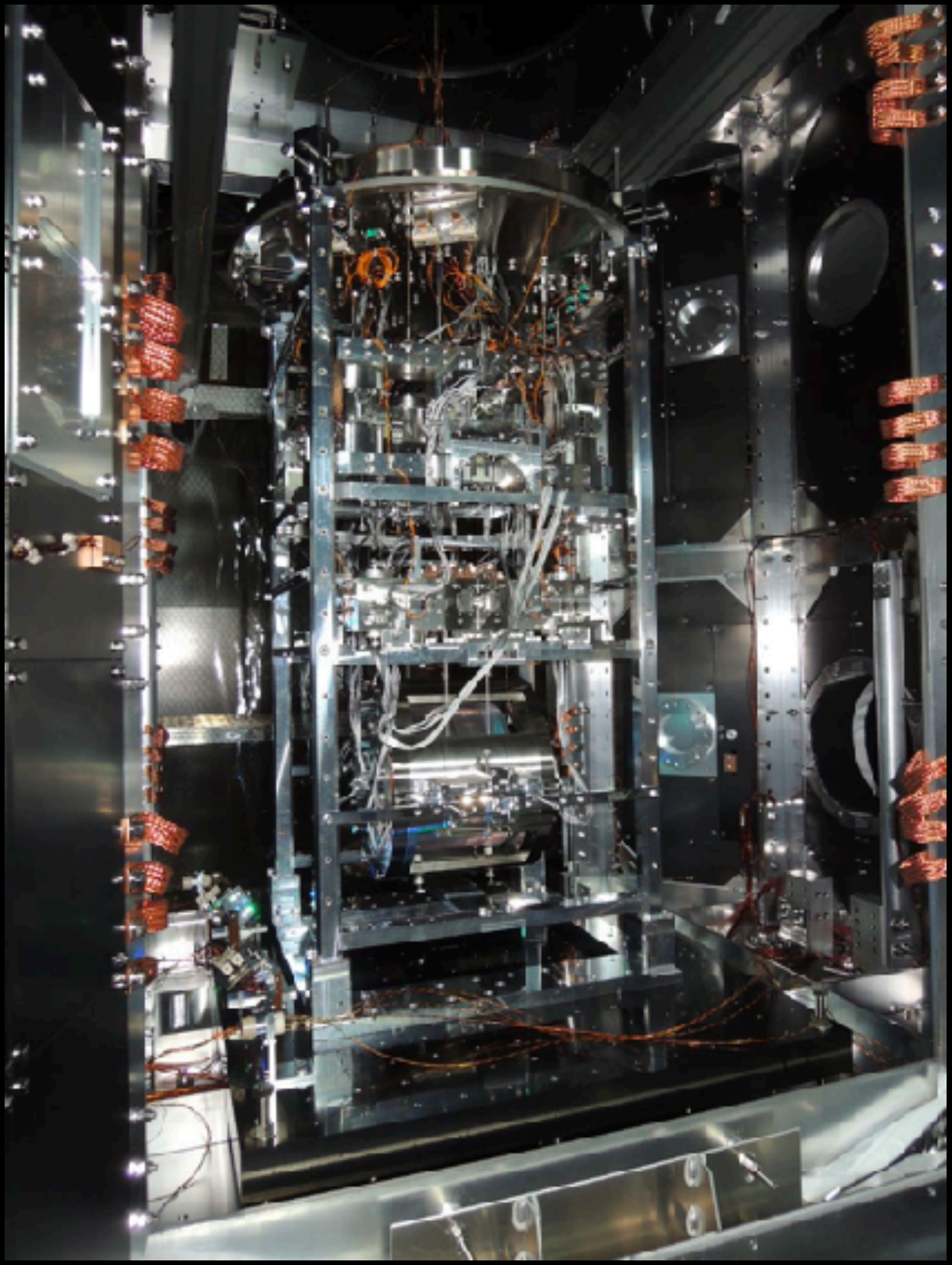
Installation of the 1st cryogenic mirror (2017 Nov 30)

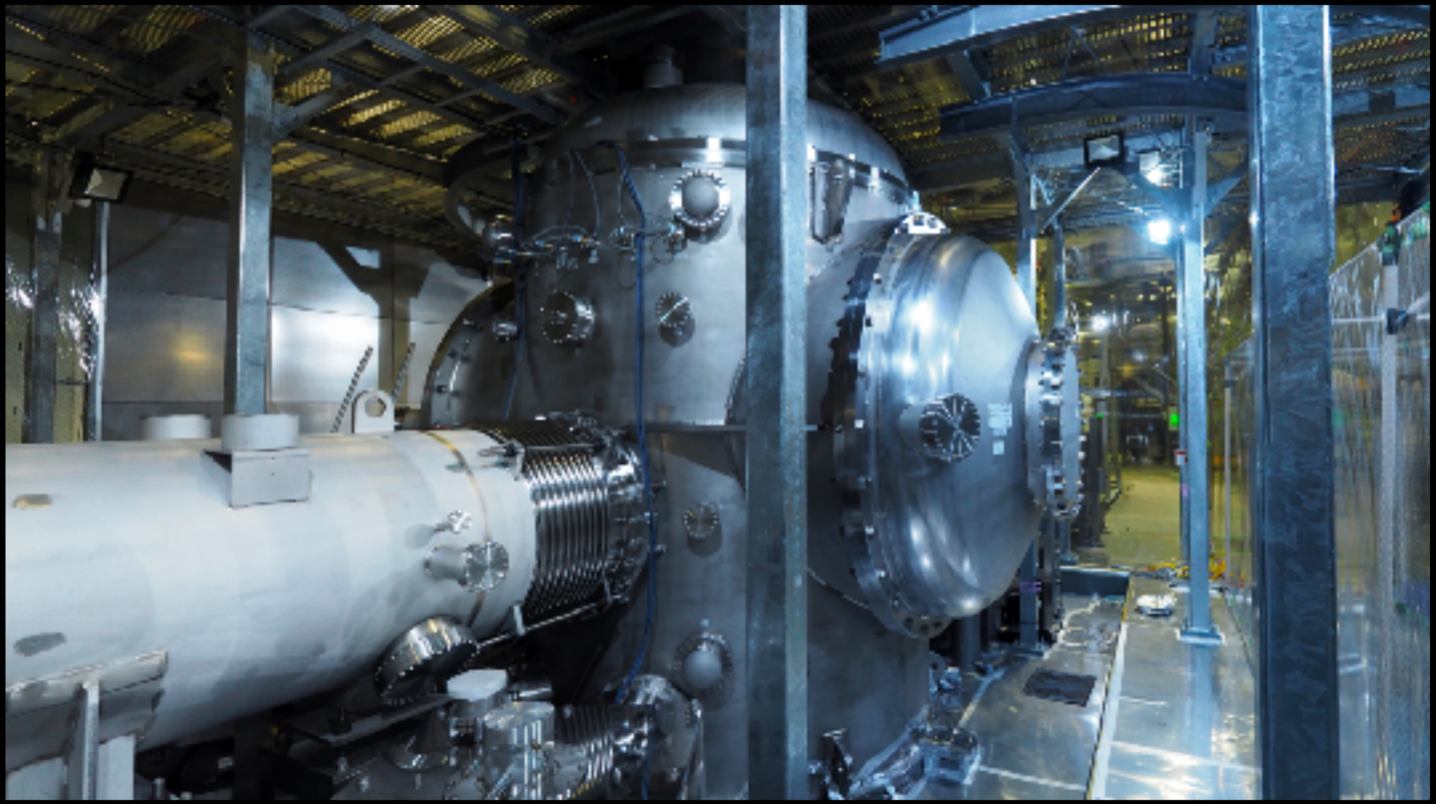
VIBRATION ISOLATION SYSTEMS+ CYROSTAT



Yamada et al. 2020 J.Phy. Conf Ser. I468 012217

VIBRATION ISOLATION SYSTEMS+ CYROSTAT





Cyrostat cools the test mass (Sapphire mirror) to ~20 K

THEORY OF MATCHED FILTERING

- Assume the data stream is $x(t) = h(t) + n(t)$ where $h(t)$ is the signal and $n(t)$ is the **wide-sense stationary noise**, s.t.

$$\langle n(t) \rangle = \langle n(t+\tau) \rangle \quad \langle n(t)n(t') \rangle = \langle n(t+\tau)n(t'+\tau) \rangle \qquad \forall \tau$$

- With the stationarity assumption, we can take $\langle n(t) \rangle = 0$
- Autocorrelation of noise: $\langle n(t)n(t') \rangle = K(t-t')$
- Fourier transform of $K(t-t')$ give power spectral density of noise:

$$\langle n(f)n^*(f') \rangle = S_n(f)\delta(f-f')$$

- Consider an arbitrary filter $F(t)$, s.t.

$$c = \int_{-\infty}^{\infty} F(t)x(t)dt = \int_{-\infty}^{\infty} F^*(f)x(f)df \qquad c \in \mathbb{R}$$

THEORY OF MATCHED FILTERING

$$\mu_c = \langle c \rangle = \langle \int_{-\infty}^{\infty} F^*(f)x(f)df \rangle = \int_{-\infty}^{\infty} F^*(f)h(f)df$$

(By zero-mean stationary noise)

$$\sigma_c^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F^*(f)F^*(f')\langle n(f)n(f')\rangle df df'$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F^*(f)F^*(f')\langle n(f)n(-f')\rangle df df'$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F^*(f)F^*(f')\langle n(f)n^*(f')\rangle df df'$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F^*(f)F^*(f')S_n(f)\delta(f-f') df df' = \int_{-\infty}^{\infty} \|F^*(f)\|^2 S_n(f) df$$

- Define signal-to-noise ratio (SNR):

$$\rho(F) = \frac{\mu_c(F)}{\sigma_c(F)}$$

THEORY OF MATCHED FILTERING

$$\rho(F) = \frac{\int_{-\infty}^{\infty} F^*(f) h(f) df}{\sqrt{\int_{-\infty}^{\infty} \|F^*(f)\|^2 S_n(f) df}} = \frac{\int_{-\infty}^{\infty} F^*(f)\sqrt{S_n(f)}\frac{h(f)}{\sqrt{S_n(f)}} df}{\sqrt{\int_{-\infty}^{\infty} \|F^*(f)\|^2 S_n(f) df}}$$

By Cauchy-Schwarz inequality: $\int_{-\infty}^{\infty} F^*(f)\sqrt{S_n(f)}\frac{h(f)}{\sqrt{S_n(f)}} df \leq \sqrt{\int_{-\infty}^{\infty} \|F^*(f)\|^2 S_n(f) df \int_{-\infty}^{\infty} \frac{\|h(f)\|^2}{S_n(f)} df}$

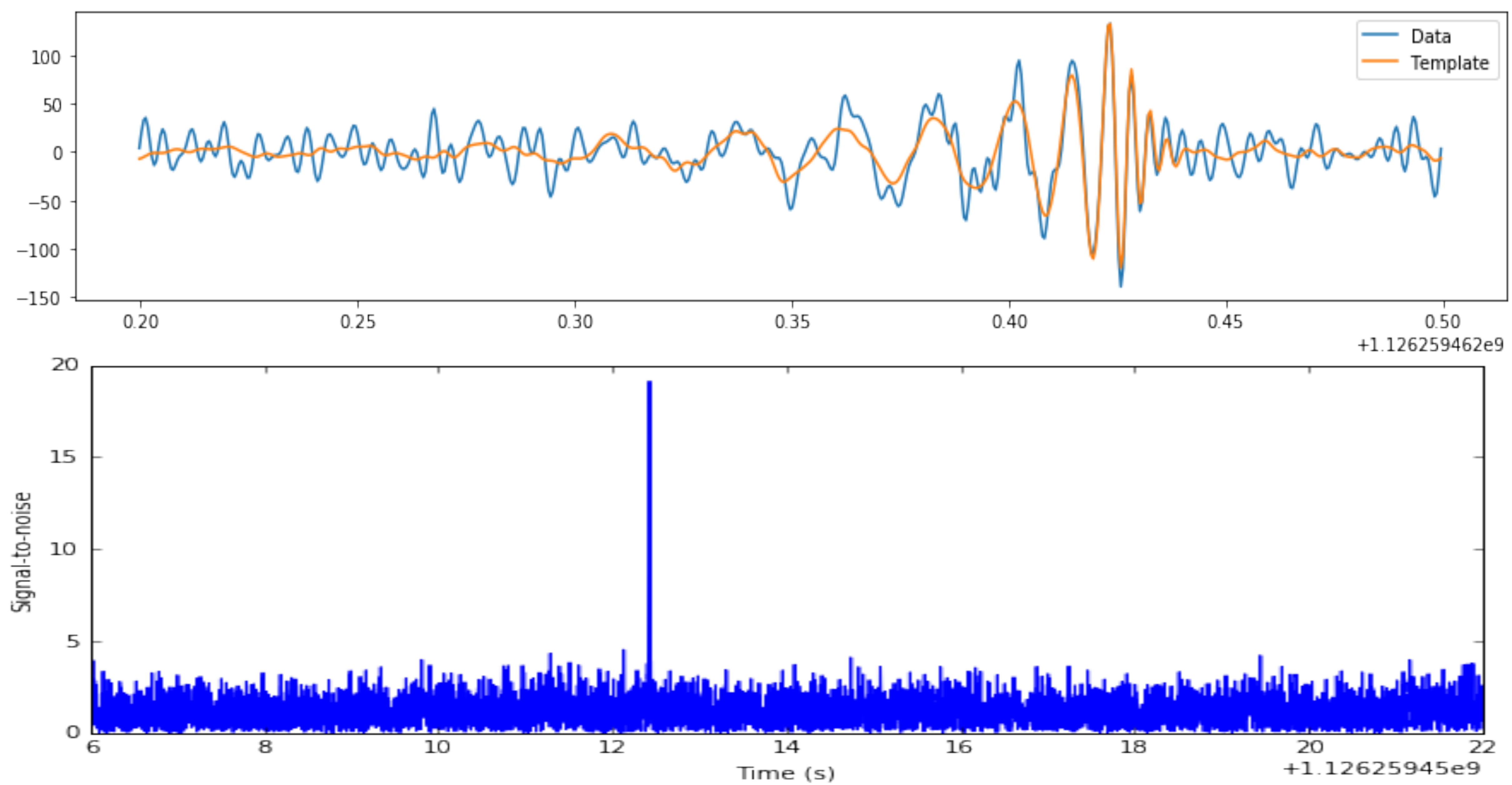
$$\rho(F) \leq \left[\frac{\int_{-\infty}^{\infty} \|F^*(f)\|^2 S_n(f) df \int_{-\infty}^{\infty} \frac{\|h(f)\|^2}{S_n(f)} df}{\int_{-\infty}^{\infty} \|F^*(f)\|^2 S_n(f) df}\right]^{1/2}$$

$$\leq \left[\int_{-\infty}^{\infty} \frac{\|h(f)\|^2}{S_n(f)} df\right]^{1/2}$$

Assuming **white noise** $S_n(f) = \text{constant}$

Matched Filter $F^*(f) = \dfrac{h(f)}{S_n(f)}$ is optimal

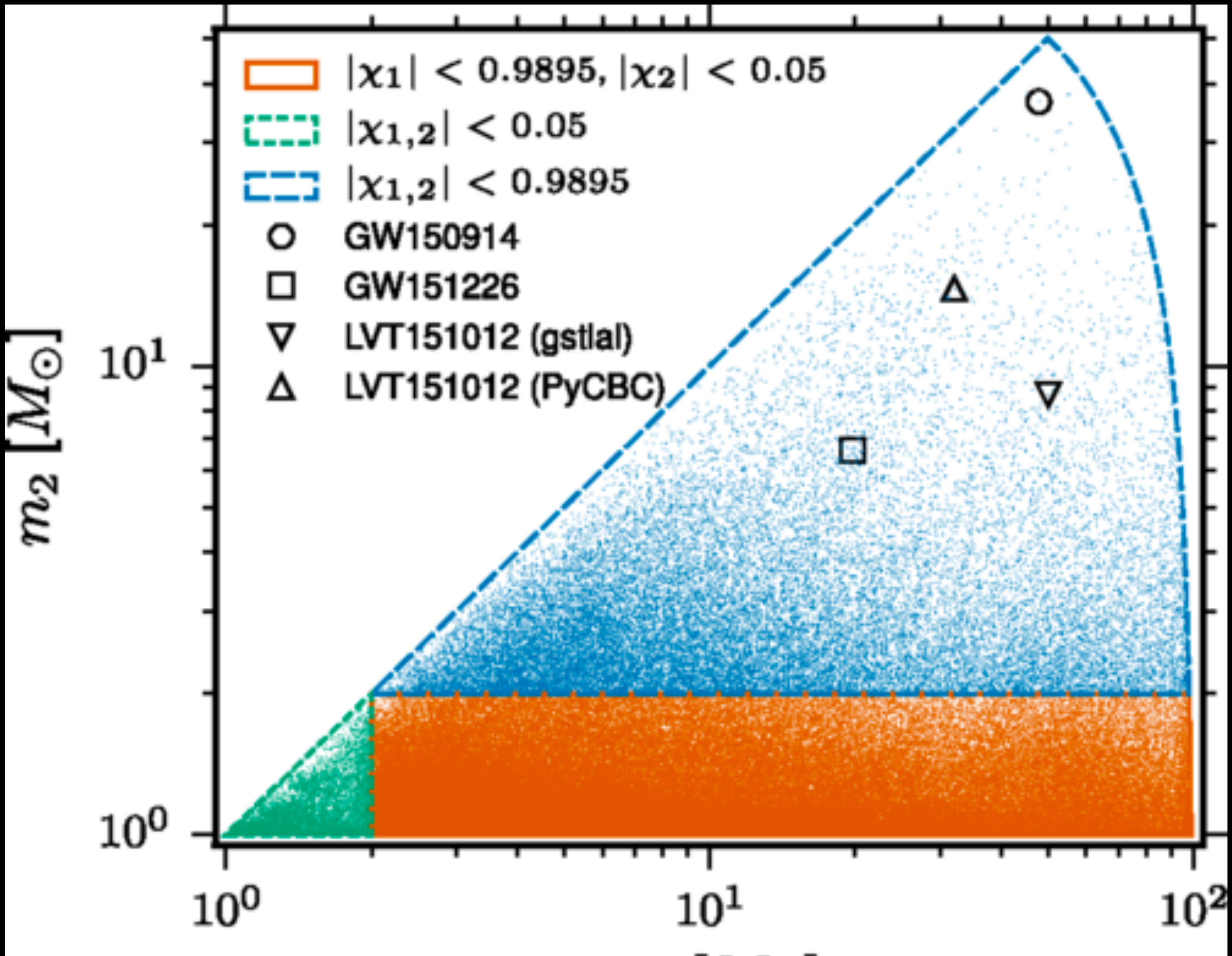
MATCHED FILTERING



A signal at 1126259462.42s with SNR 19.1247029698 is found.

- Matched filtering is an established way for detecting GW signal. Matched filter can be obtain by correlating a **template** $h(t)$ (*with known form*) with an unknown time series to detect the presence of the template in the signal.

MATCHED FILTER



Abbott et al. (2016)

The template bank used in O1 covers a 4D parameter space.

Assume the binaries are circular with the BH spins aligned/anti-aligned with the orbital angular momentum.

~250000 template are generated for the search.

SHORTCOMINGS OF MATCHED FILTERING

- Require an exact form to be known in advance
- Computationally expensive
- Assumptions for its optimal performance might not be fulfilled (i.e. stationary, zero-mean, white noise)
- Do not enable direction visualization of the waveform from the data

1. AUTOREGRESSIVE MODELING

IMPROVE NOISE REDUCTION

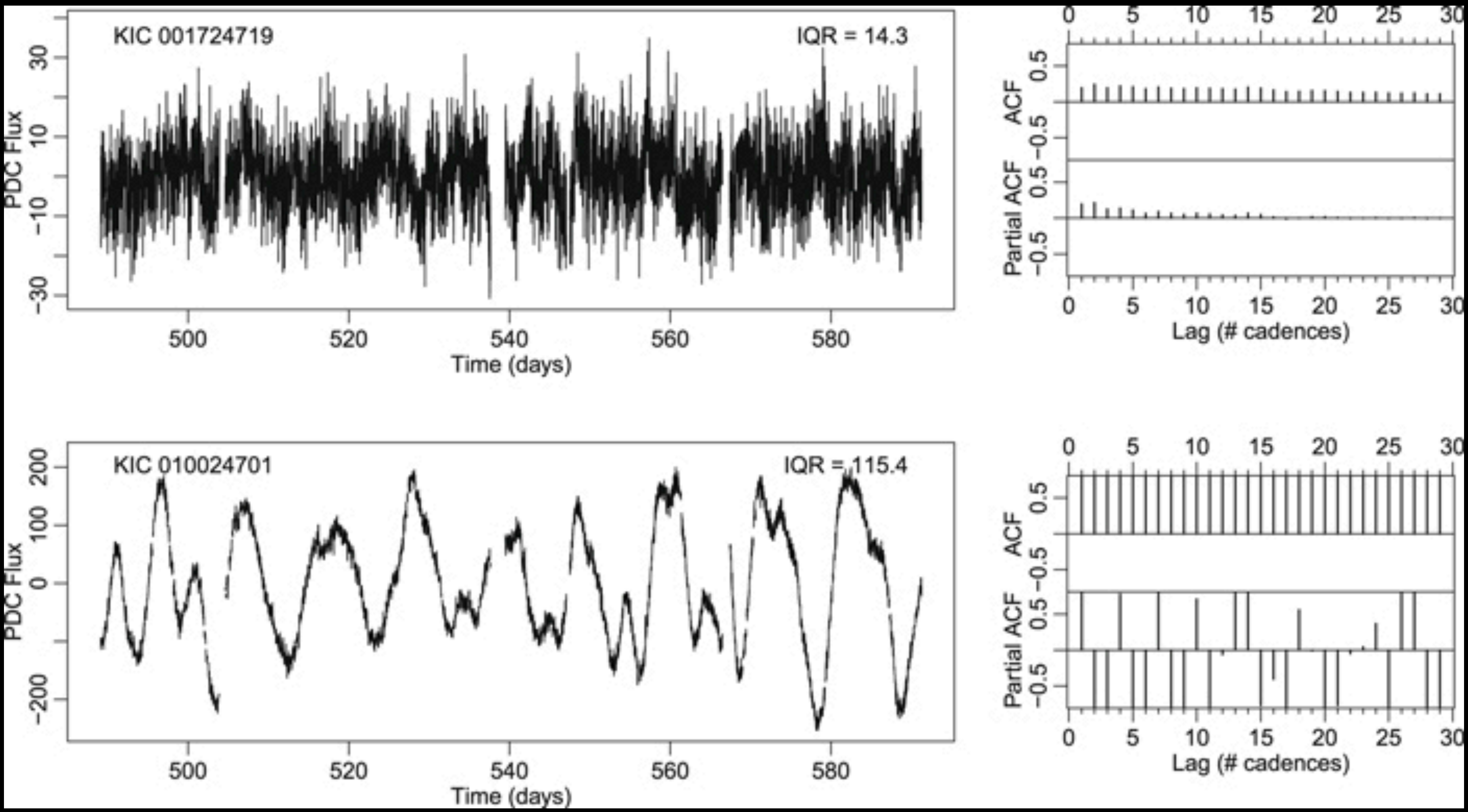
- Enable us to…

1. Search for fainter sources (e.g. continuous wave, CCSNe)
2. More accurate parameter estimation (including localization)
3. Test GR at the limit of strong field

NOISE REDUCTION WITH STOCHASTIC AUTOREGRESSIVE MODELING

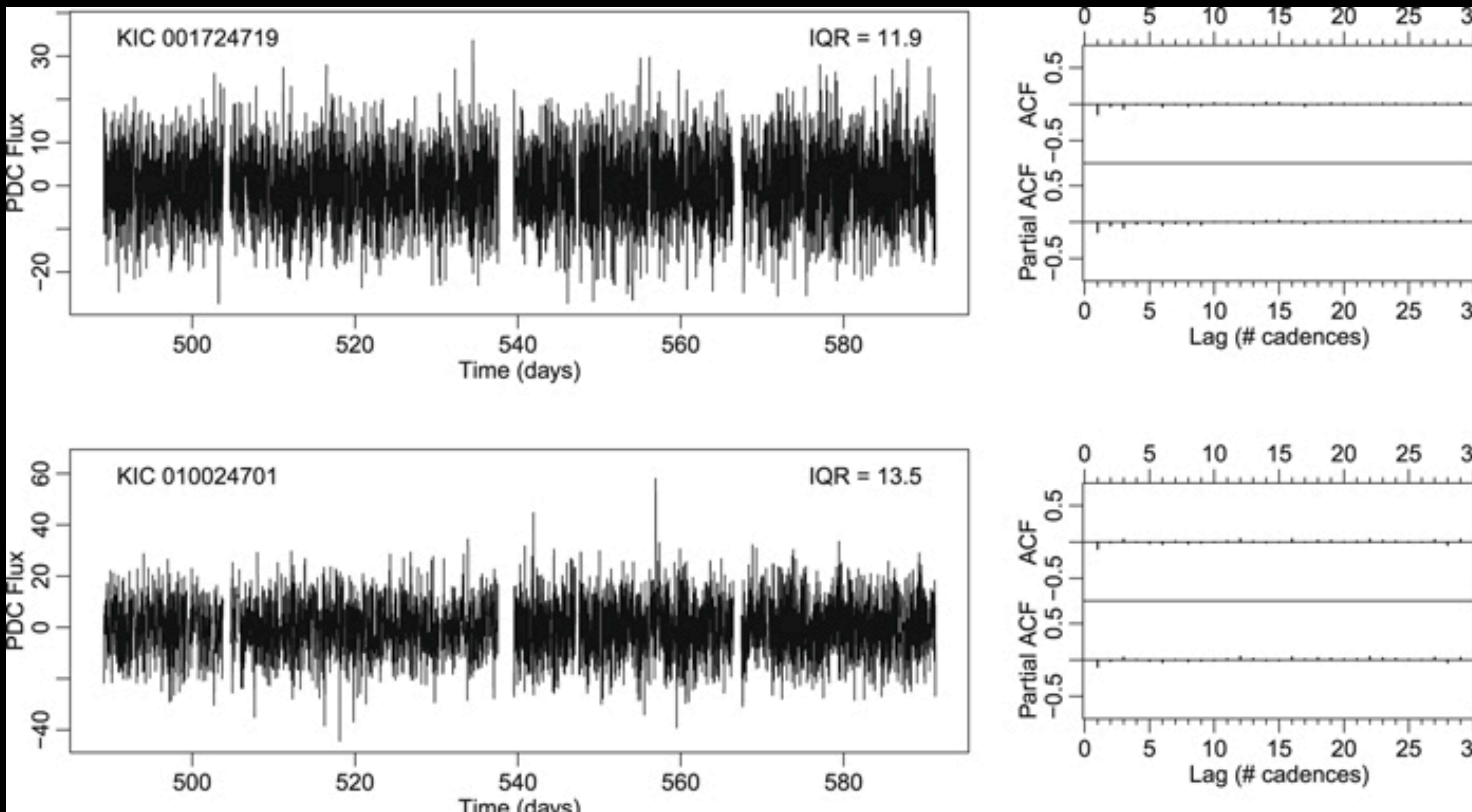
- Computationally simple
- Capable to handle various kinds of noise from non-stationary autocorrelated stochastic processes.
- Many applications in diverse fields (e.g. ECG, econometrics), but not many in astronomy until recently (e.g. exoplanet search, Caceres et al. 2019)

AUTOREGRESSIVE PLANET SEARCH



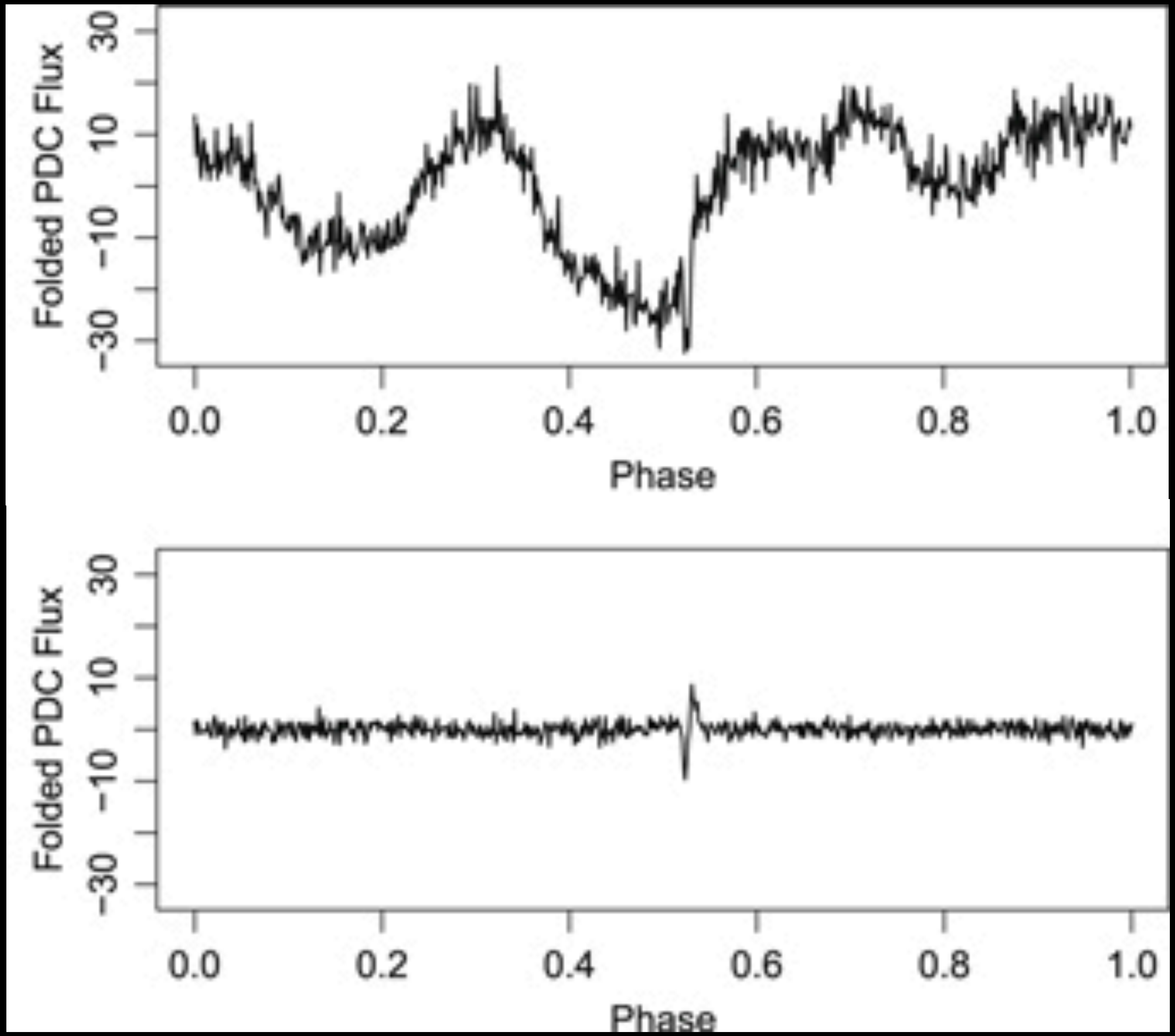
Caceres et al. 2019

AUTOREGRESSIVE PLANET SEARCH



Caceres et al. 2019

AUTOREGRESSIVE PLANET SEARCH



Caceres et al. 2019

AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (**ARIMA**) MODEL

Combining the autoregressive (AR), moving average (MA) and integrated (I) processes together into a single regression procedure, we have ARIMA(*p*,*q*,*d*) model:

$$(1 - B)^d x_t = \sum_{i=1}^{p} a_i x_{t-i} + \sum_{j=1}^{q} b_j \epsilon_{t-j} + \epsilon_t + c$$

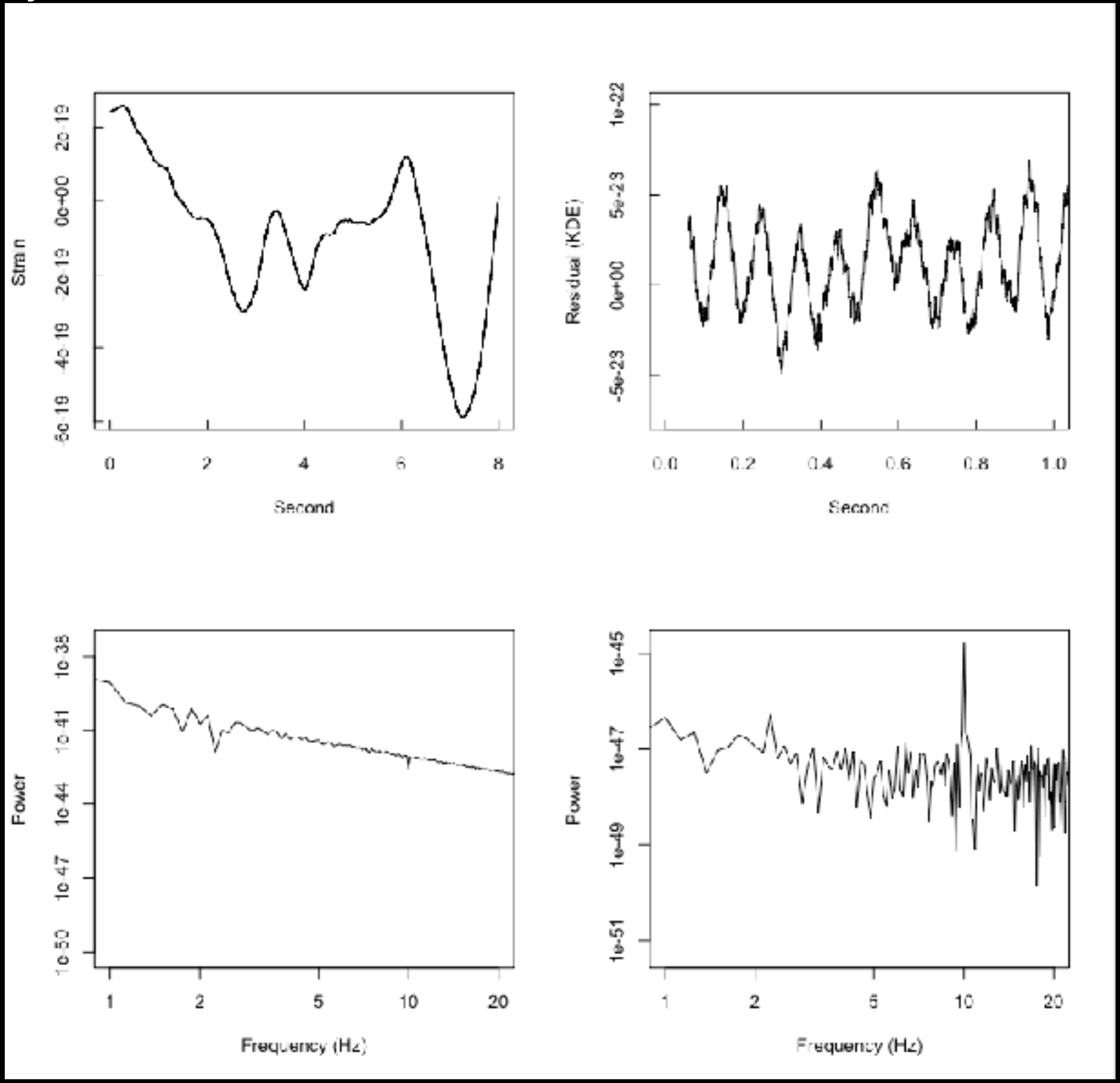
The model parameters can be determined by maximum likelihood estimation.

MERITS OF ARIMA MODEL

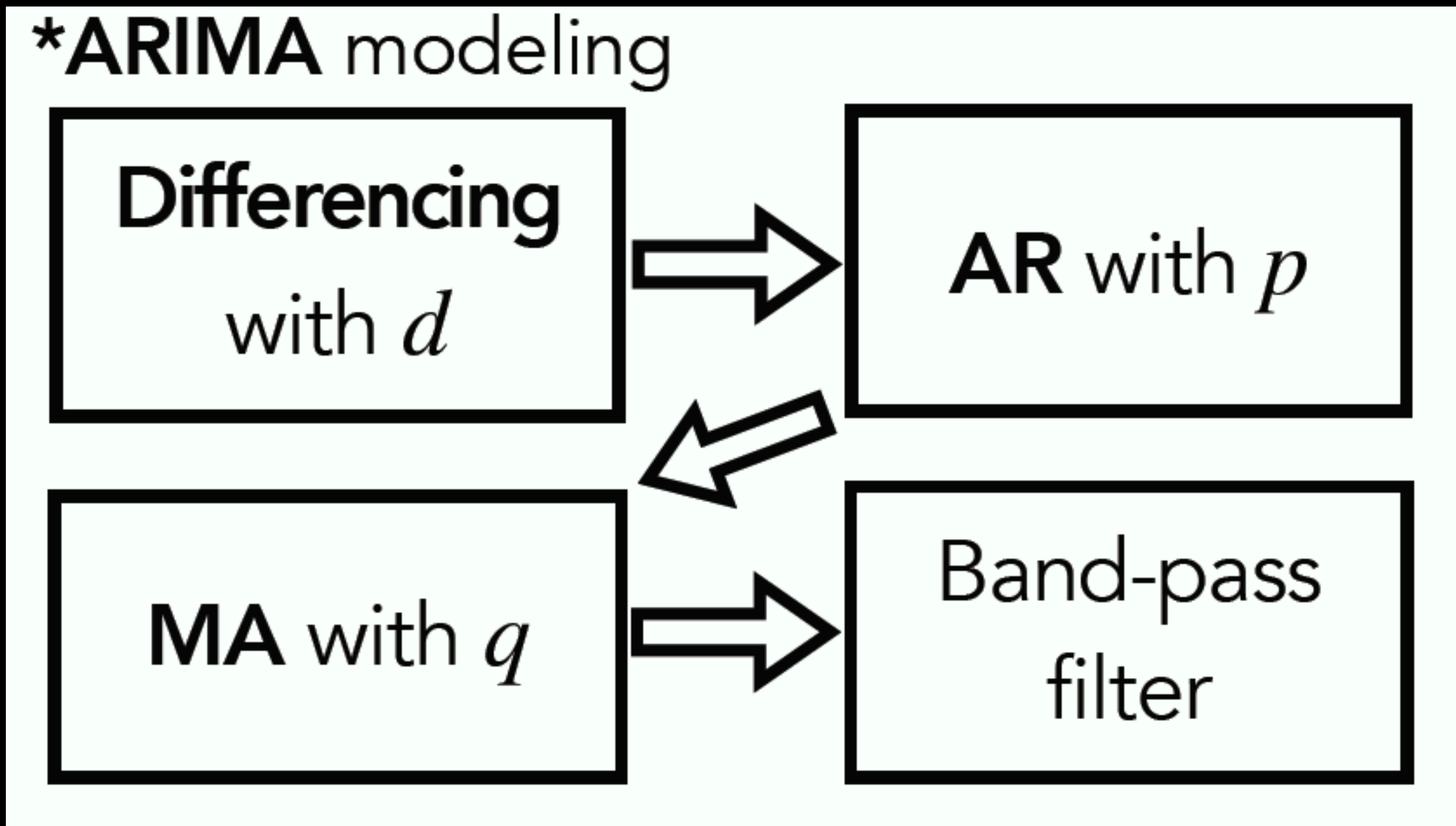
- Computational simple
- Flexibility in modeling various form of noises
- Efficient in de-trending and whitening

PROOF-OF-CONCEPT

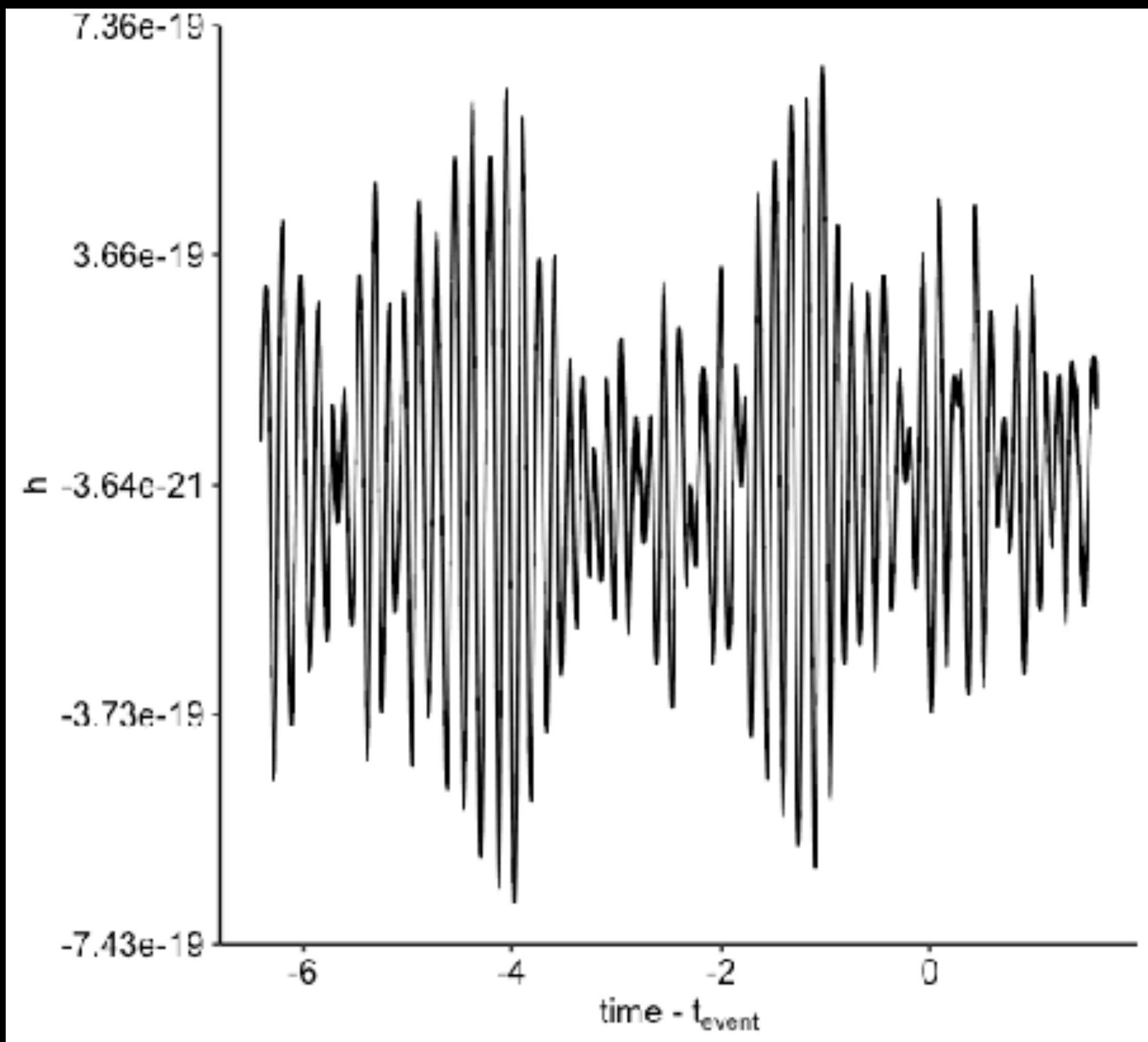
- The simulated LIGO noise strain series with a constant 10 Hz sinusoidal signal of $h \sim 10^{-21}$ injected.
-



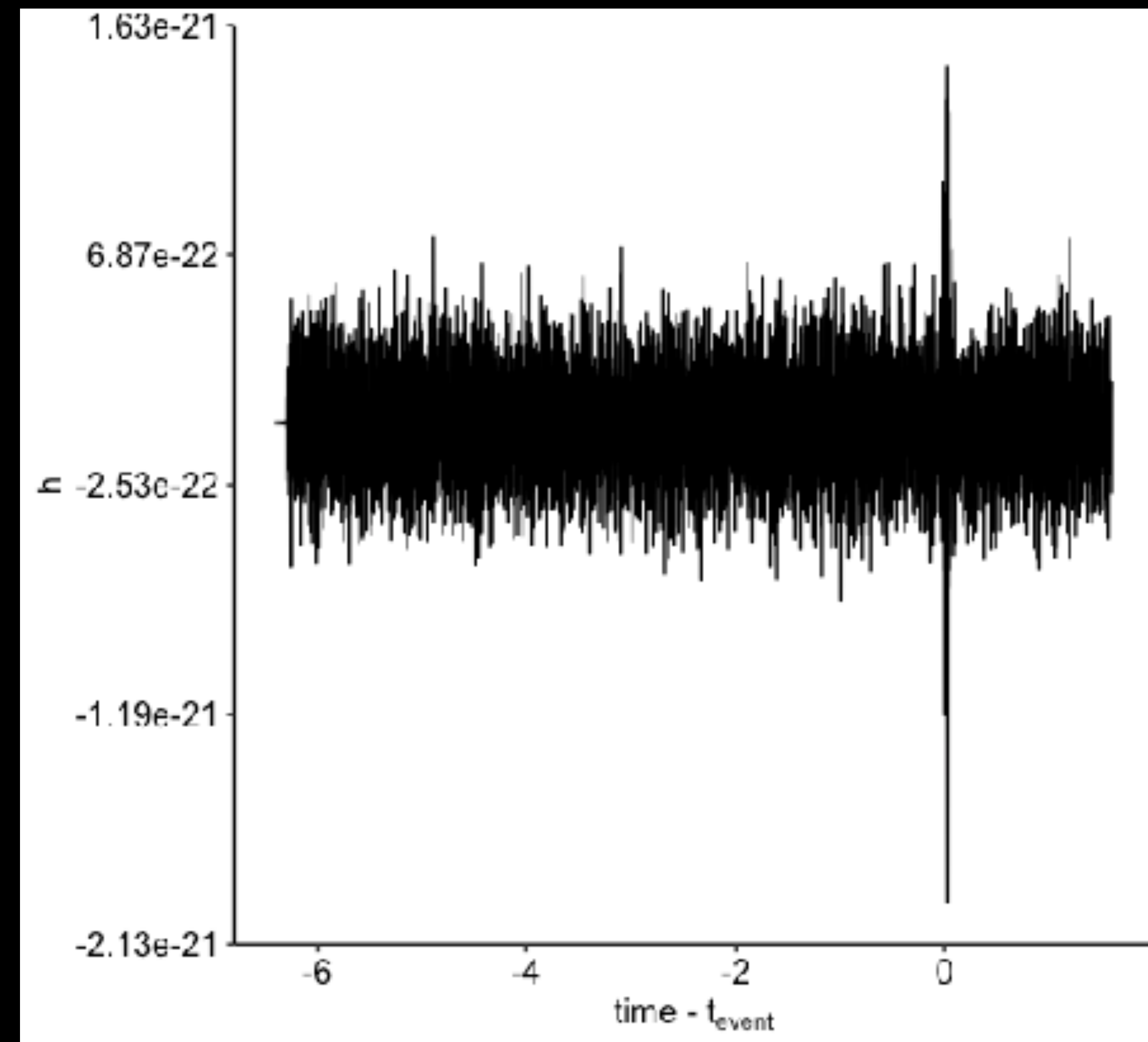
AUTOREGRESSIVE GW SEARCH



ARIMA MODELING OF GW150914

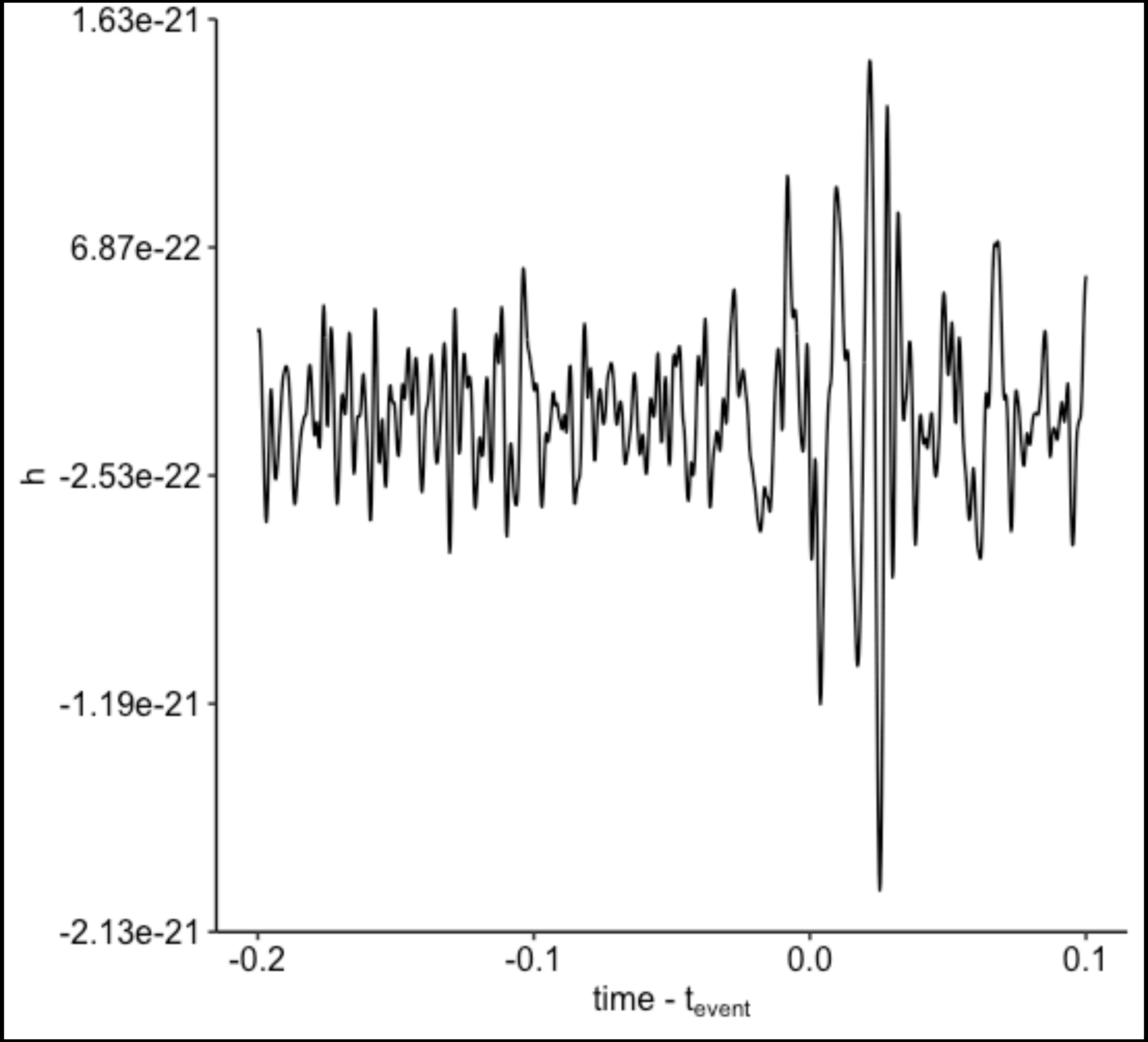


Raw data



Residuals

ARIMA MODELING OF GW150914



Residuals (zoom-in)

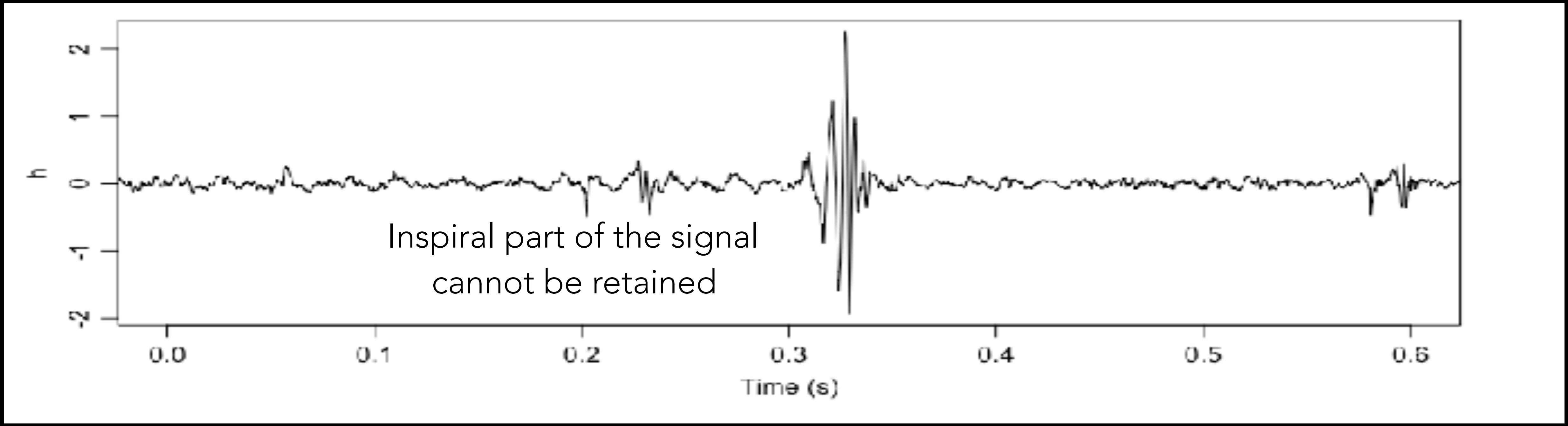
EFFECTIVENESS OF ARIMA MODEL

- A diversity of noises are subsumed into ARIMA model without any fine-tuning and a priori knowledge of the noise nature
- A desirable feature of this method is that the transient GW signals were NOT absorbed by ARIMA model.
- ARIMA is a maximum likelihood estimation procedure which weight all data points equally. As the transient signal is a small fraction in a given window, their data points are essentially ignored in the model.

A CONCERN IN OPTIMAL MODELING

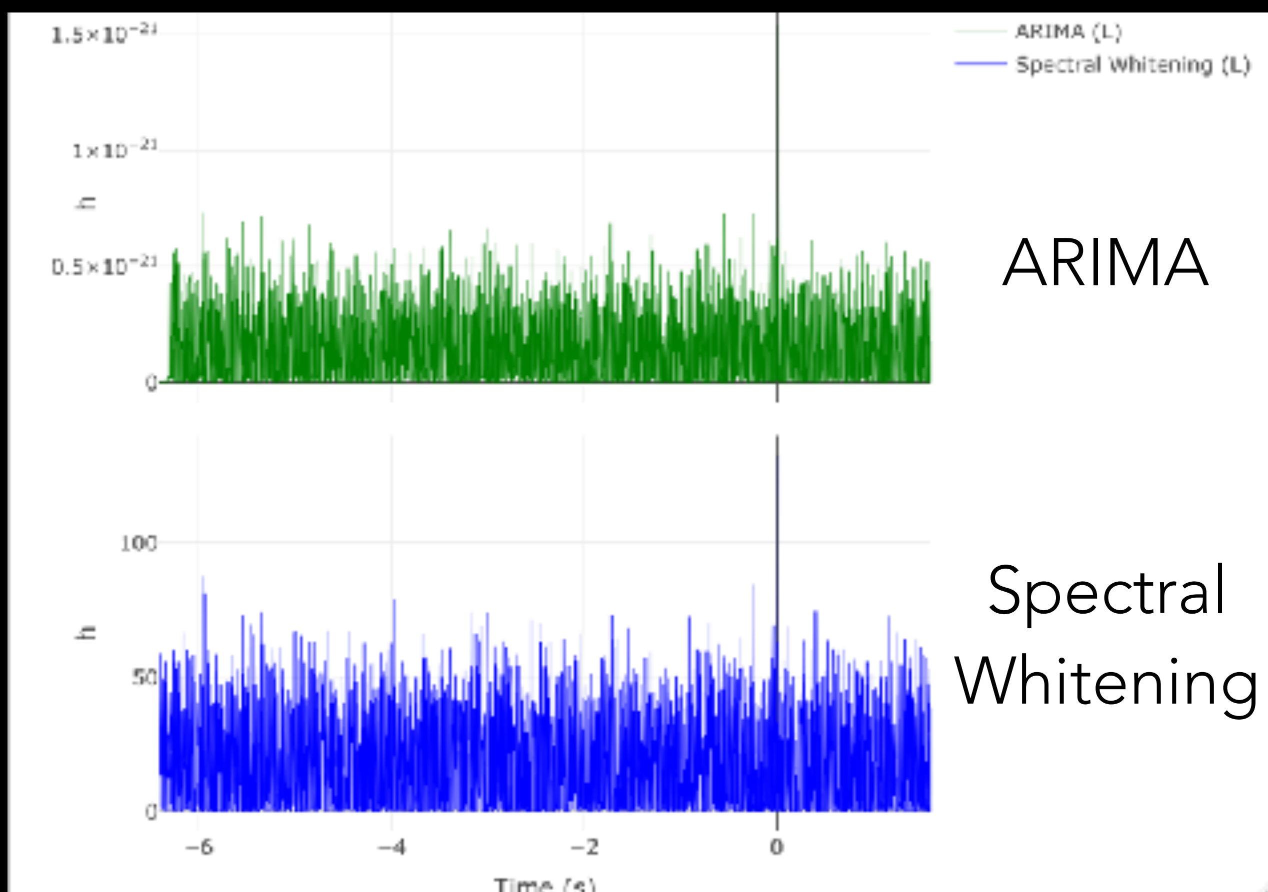
- Conventionally, ARIMA model is obtained by selecting the orders of the model (i.e. p, q, d) based on certain information criterion (e.g. AIC).

- Although the optimal model resulted from this procedure can minimize the noise, it might over-subtract the data, including the potential signal (or part of the potential signal).

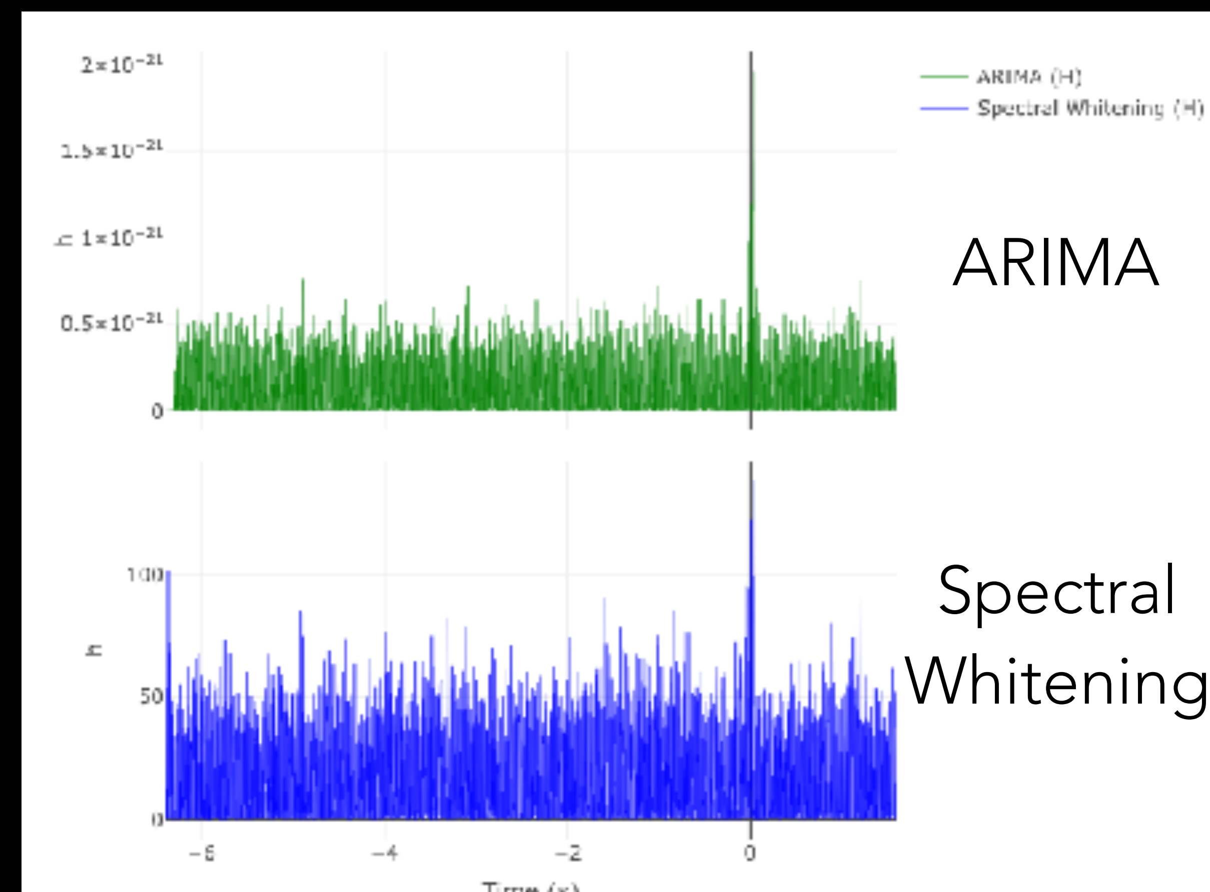


ARIMA VS SPECTRAL WHITENING

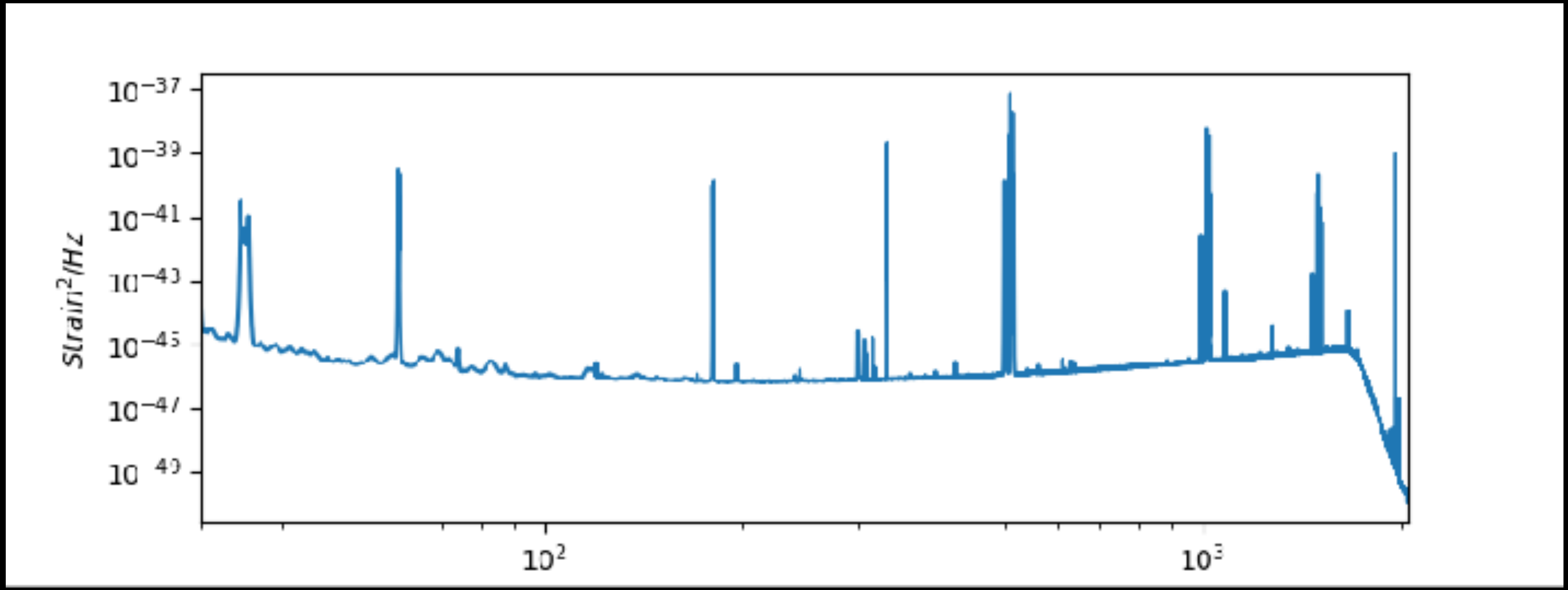
33.2% improve in SNR



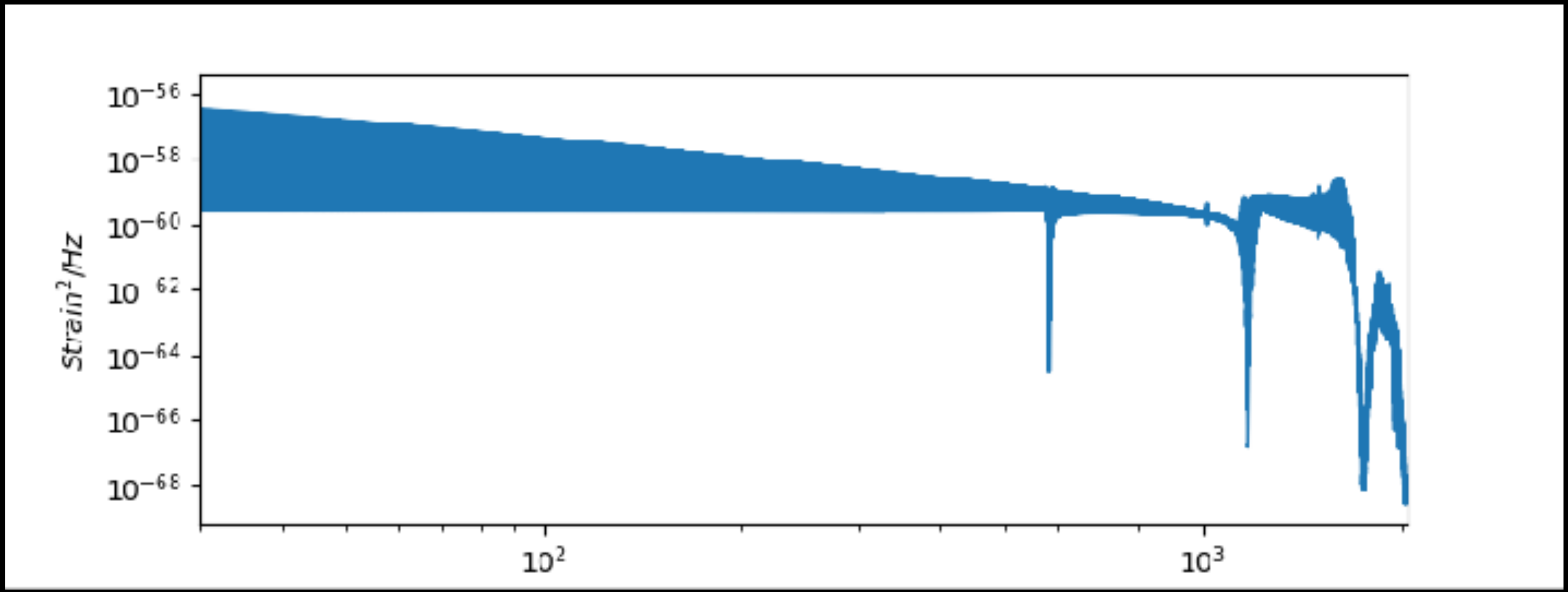
31.5% improve in SNR



WHITENING & LINE-REMOVAL BY ARIMA

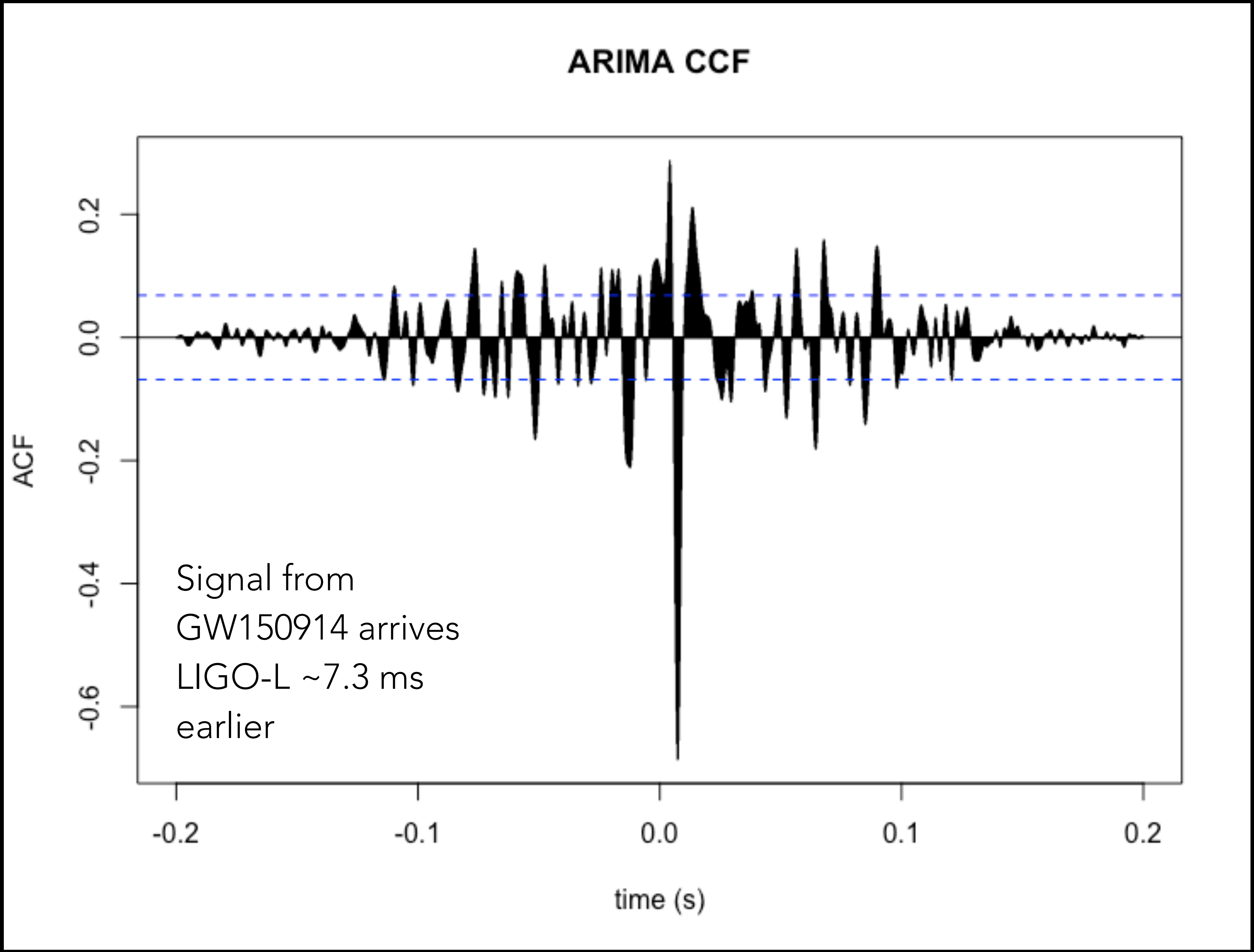


Raw Data

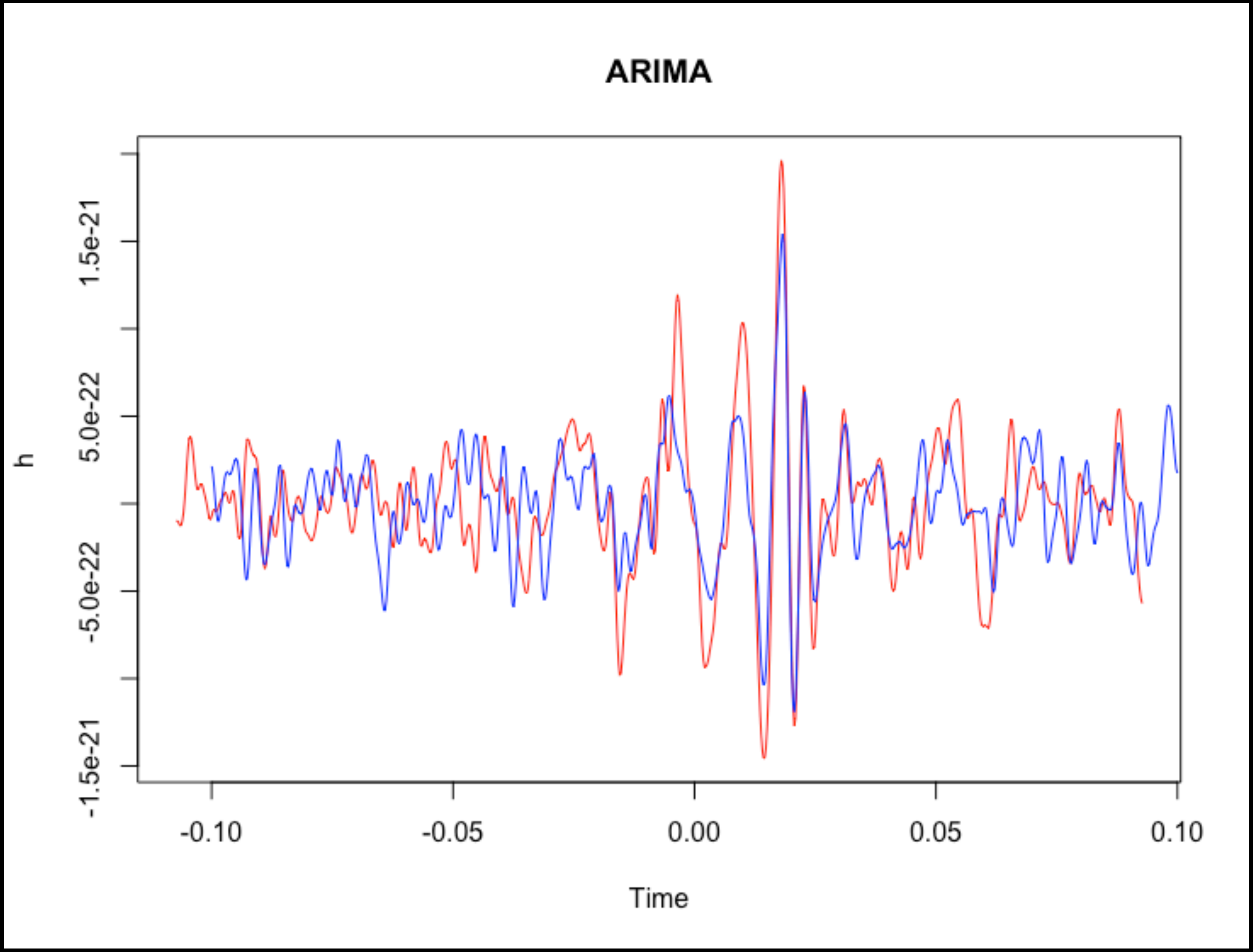


ARIMA

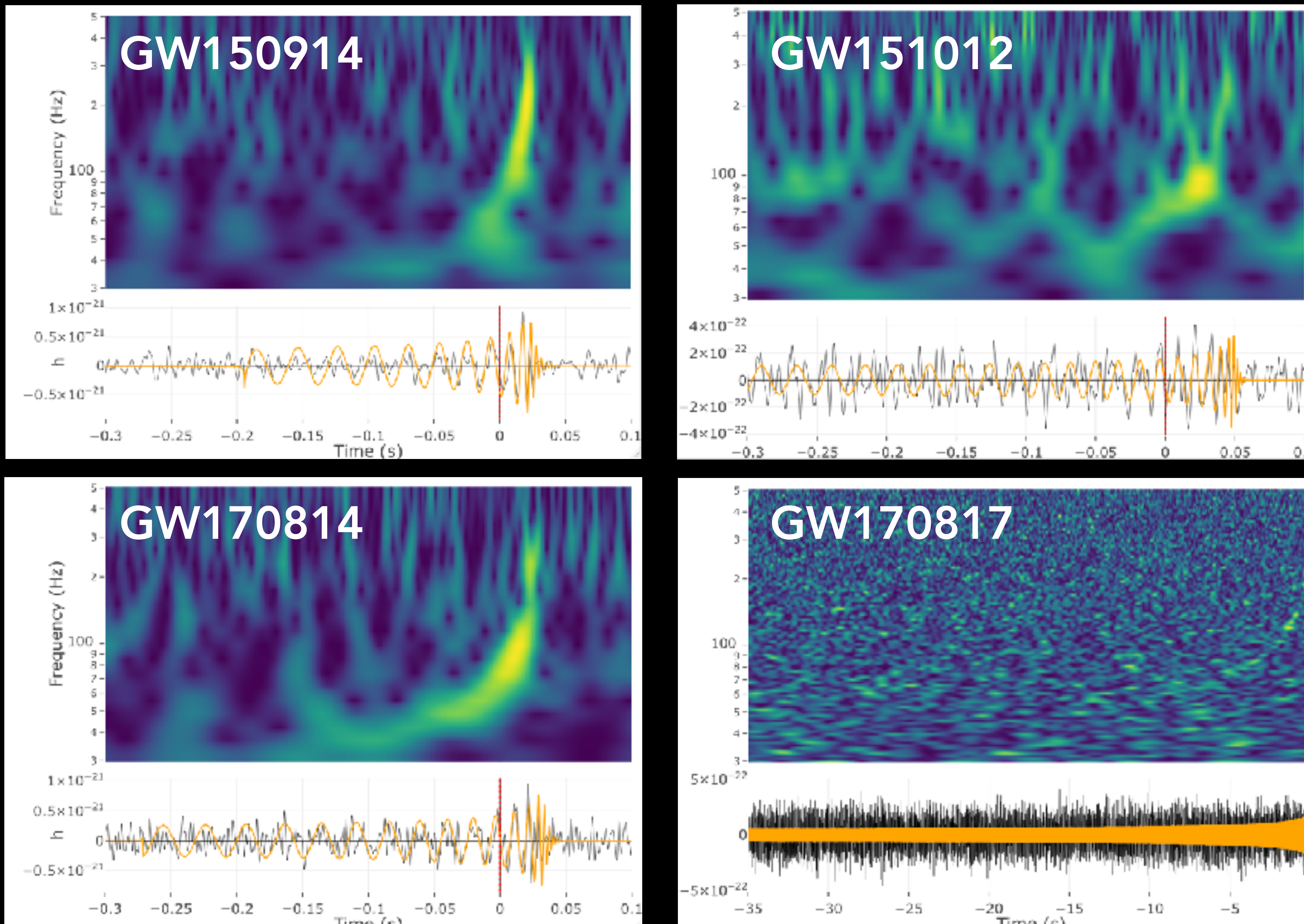
CROSS-CORRELATION BETWEEN DATA FROM DIFFERENT DETECTORS



CROSS-CORRELATION BETWEEN DATA FROM DIFFERENT DETECTORS



RECOVERING GWTC-1



ON-GOING/FUTURE WORKS

1. Automatic search for the hyper-parameters

- Since model selection based on information criterion can lead to over-subtraction, we are investigating the way to choose the order of the model.

- In the moment, we limit the order of AR part (i.e. p) to be 10% of the data length.

- For q and d, the ranges are at the order of unity. In the current stage, we set them manually. We are investigating a more efficient way for obtaining the optimal model.

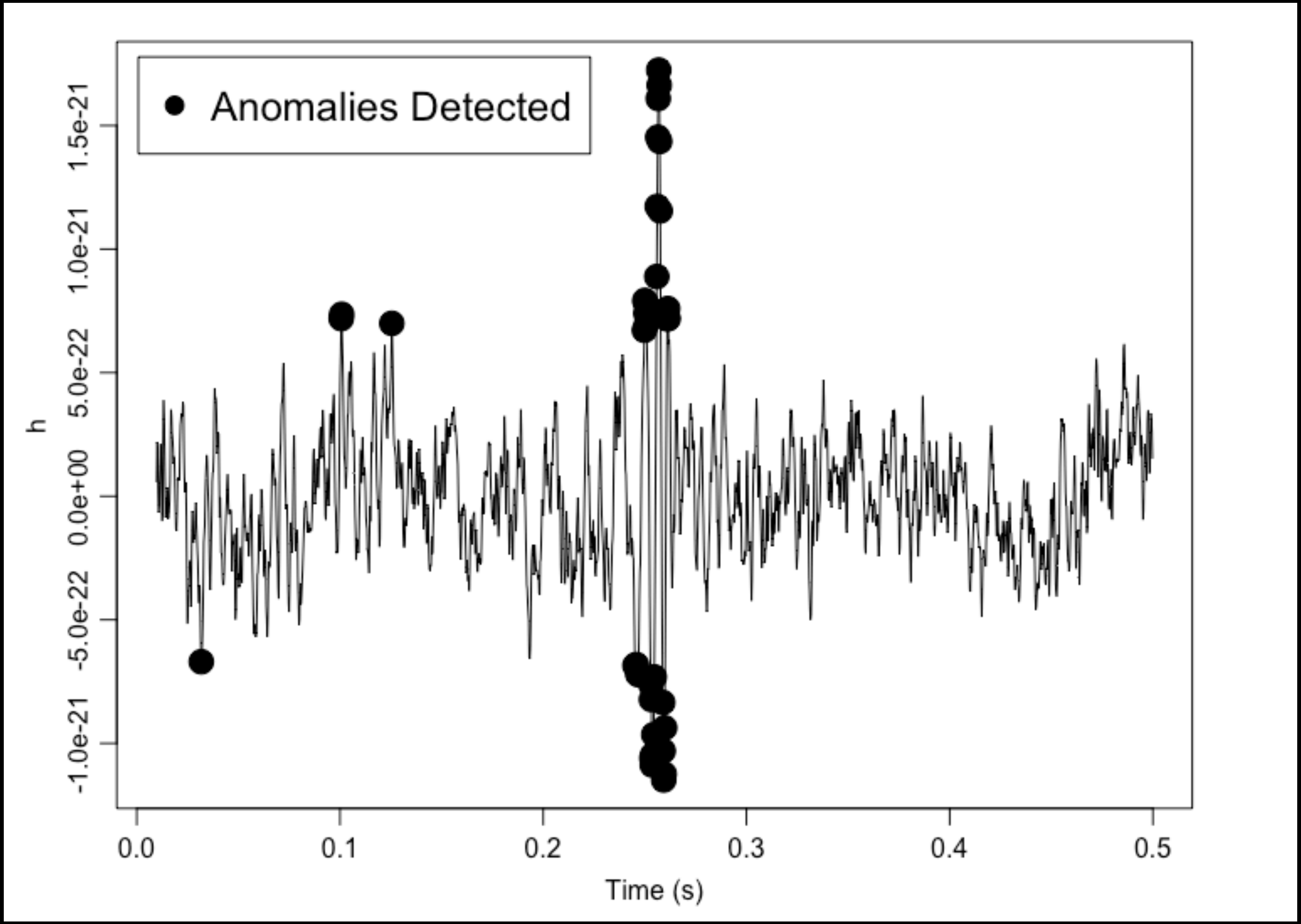
ON-GOING/FUTURE WORKS

2. Anomaly Detections

- After the noise subtraction, events candidates can be identified as anomalies, which differ from normal instances significantly.
- If the duration of the signal is significantly longer than the sampling interval, a cluster of anomalies is expected.
- Anomalies detected from different detectors (LIGO-H, LIGO-L,KAGRA,VIRGO) can be cross-correlated and analysed with clustering technique.
- The shortlisted anomalies can be taken as event candidates for further analysis.

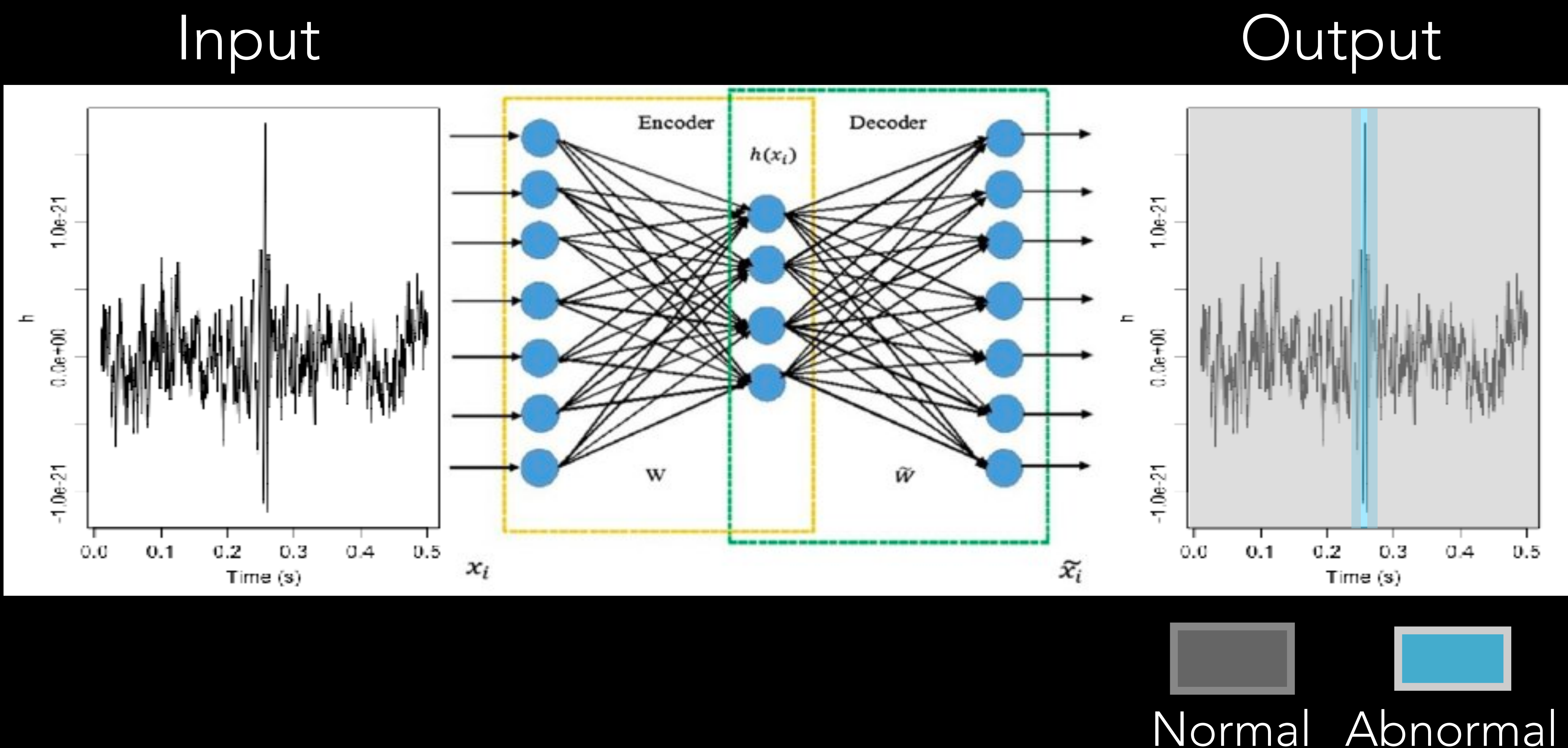
ON-GOING/FUTURE WORKS

2. Anomaly Detections



ON-GOING/FUTURE WORKS

2. Further work will be devoted to improve the performance of anomaly detection with machine learning techniques (e.g. autoencoder).



ON-GOING/FUTURE WORKS

3. Template-free Parameter Estimation

- Cleaned signal can be fitted with an 2nd stage AR model.
- Signal can be reconstructed from the best-fit model.
- Characteristic equation can be obtained from $\{a_j\}$ and the order p.
- $$F(z) = 1 - \sum_{j=1}^{p} a_j z^j = 0$$
- QNM frequency/Damping can be obtained from the complex roots.

$$z_k = \exp(i2\pi f_k \Delta t)$$

$\mathrm{Re}(f_k) \longrightarrow$ Frequency

$\mathrm{Im}(f_k) \longrightarrow$ Damping

It has been shown that this can extract the ring-down freq./damping timescale from GW150914 (Shinkai 2018,2019).

2. IMPROVING DEEP LEARNING CLASSIFICATION WITH GENERATIVE ADVERSARIAL NETWORK (GAN)

GLITCHES

- Understanding non-Gaussian transient noises (i.e. glitches) is critical for detector characterization.
- Because their high rate, they can mimic or obscure the astrophysical signal and potentially leads to false detection.
- Separating these environmental/instrumental noises from the astrophysical signal is an important issue.

GRAVITY SPY DATASET

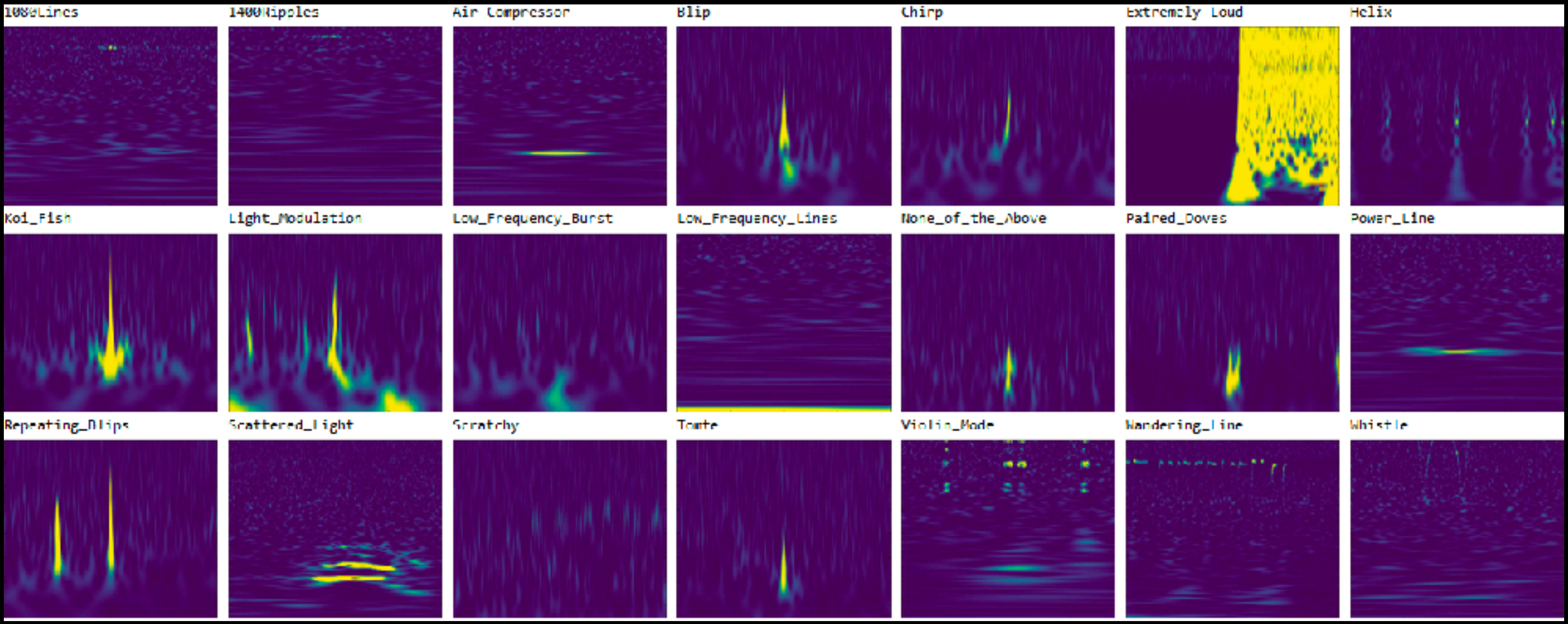
- Combination of crowd-sourcing with machine learning in classifying transients recorded by the gravitational wave detectors by using spectrograms.

- Classes in the glitch datasets are firstly tagged by citizen scientists, which provide seeds for machine learning.

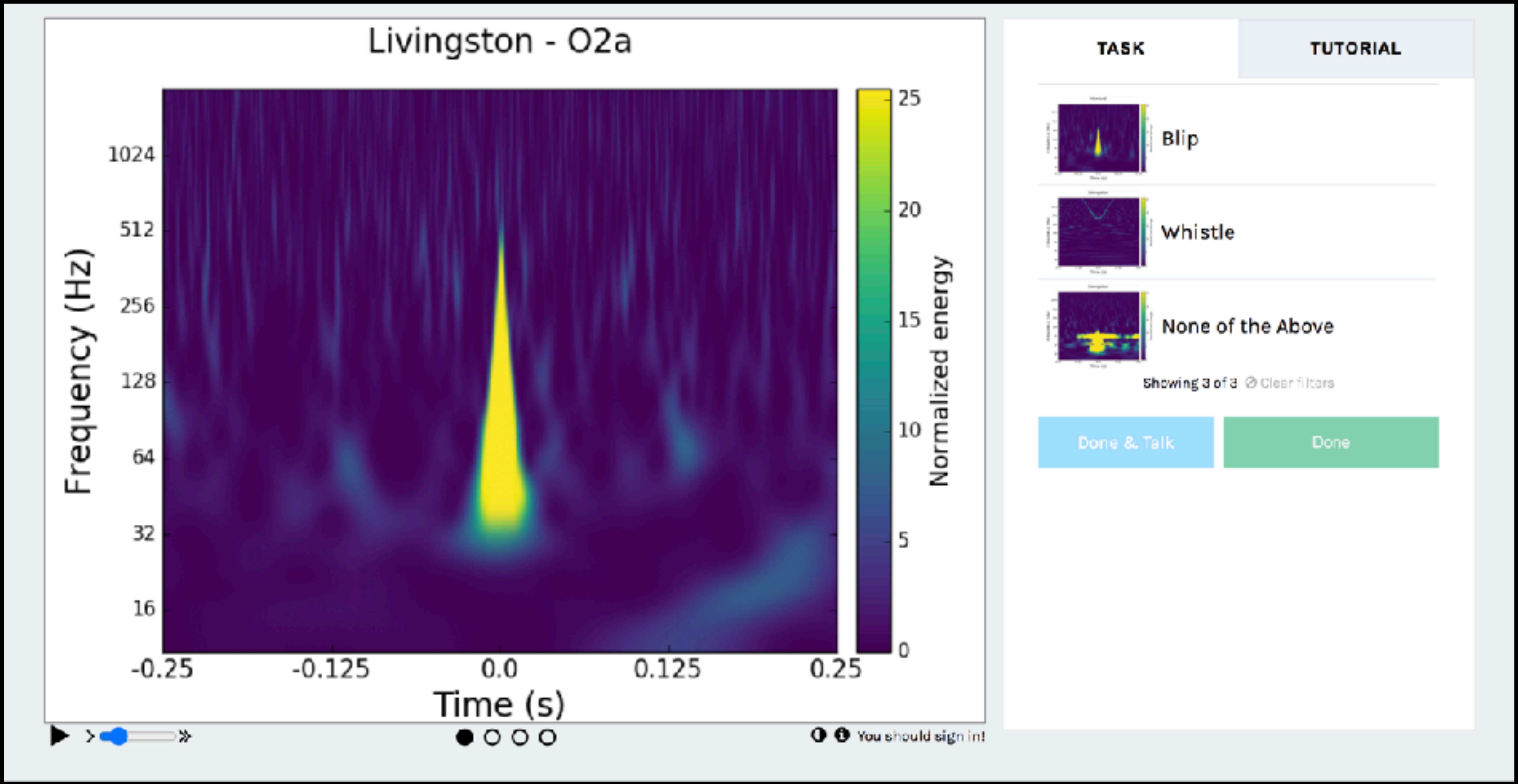
- Trained classifier can sift more data and sent the interesting/abnormal glitches back to citizen scientists.

- Providing a test ground for classification algorithm.

GRAVITY SPY DATASET



GRAVITY SPY AS CITIZEN SCIENCE



https://www.zooniverse.org/projects/zooniverse/gravity-spy/classify

CLASSIFYING WITH LOGISTIC REGRESSION

True Class \ Predicted Class	1080Lines	1400Ripples	Air Compressor	Blip	Chirp	Extremely Loud	Helix	Koi Fish	Light Modulation	Low Frequency Burst	Low Frequency Line	None of the Above	No Glitch	Paired Doves	Power Line	Repeating Blips	Scattered Light	Scratchy	Tomte	Violin Mode	Wandering Line	Whistle
1080Lines	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1400Ripples	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Air Compressor	0	0	0.89	0	0	0	0	0	0	0	0	0	0	0	0.11	0	0	0	0	0	0	0
Blip	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chirp	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Extremely Loud	0	0	0	0	0	0.94	0	0.01	0	0	0	0	0	0	0	0	0.01	0	0	0	0.03	0
Helix	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Koi Fish	0	0	0	0.01	0	0	0	0.99	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Light Modulation	0	0	0	0.01	0	0	0	0.01	0.97	0.01	0	0	0	0	0	0	0	0	0	0	0	0
Low Frequency Burst	0	0	0	0.01	0	0	0	0	0.02	0.95	0.02	0	0	0	0	0	0	0	0	0	0	0
Low Frequency Line	0	0	0	0	0	0	0	0	0	0	0.97	0	0.03	0	0	0	0	0	0	0	0	0
None of the Above	0	0	0	0	0	0	0	0	0.23	0	0.08	0.38	0	0	0	0	0	0.31	0	0	0	0
No Glitch	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0
Paired Doves	0	0	0	0	0	0	0	0	0	0	0	0	0.50	0.50	0	0	0	0	0	0	0	0
Power Line	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0
Repeating Blips	0	0	0	0.38	0	0	0	0	0.02	0	0	0	0	0	0	0.60	0	0	0	0	0	0
Scattered Light	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.01	0	0.99	0	0	0	0	0
Scratchy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0
Tomte	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0
Violin Mode	0	0	0	0	0	0	0	0	0	0	0.01	0	0.01	0	0	0	0	0	0	0.97	0	0
Wandering Line	0.29	0	0	0	0	0	0	0	0	0	0	0	0.14	0	0	0.14	0	0.14	0	0	0.14	0.14
Whistle	0	0	0	0.02	0	0	0	0	0	0	0.02	0	0.04	0	0	0	0	0	0	0	0	0.91
Recall	0.96	1.00	1.00	0.93	1.00	1.00	1.00	0.98	0.93	0.99	0.93	1.00	0.77	1.00	0.97	0.96	0.99	0.91	1.00	1.00	0.33	0.98
Precision	1.00	1.00	0.89	1.00	1.00	0.94	1.00	0.99	0.97	0.95	0.97	0.38	1.00	0.50	1.00	0.60	0.99	1.00	1.00	0.97	0.14	0.91

Predicted Class

True Class

1.0 0.8 0.6 0.4 0.2 0.0

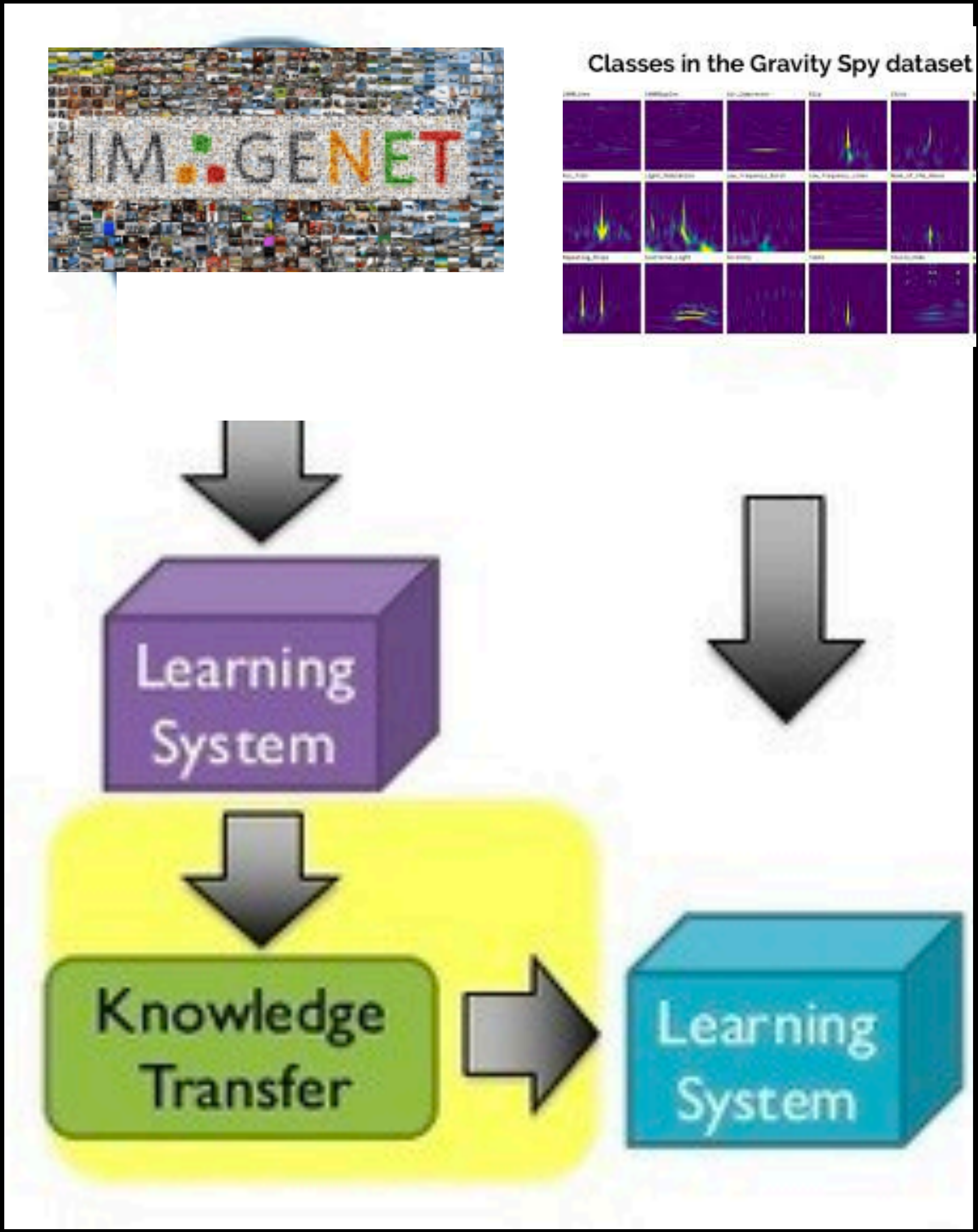
Bahaadini et al. (2018)

CHALLENGE OF DEEP LEARNING IN CLASSIFYING GRAVITY SPY DATA

Class Name	Version 1.0	Version 1.1	Blank (V1.1)
1080lines	1312	1312	0
1400ripples	928	324	0
Air_compressor	232	232	0
Blip	7476	7288	32
Chirp	264	240	0
Extremely_loud	1816	1788	536
Helix	1116	1116	0
Koi_fish	3320	2824	8
Light_modulation	2292	2048	8
Low_frequency_burst	2628	2484	8
Low_frequency_lines	1812	1788	0
No_glitch	724	600	52
Paired_doves	108	108	36
Power_line	1812	1796	4
Repeating_blips	1140	1052	8
Scattered_light	1836	1772	32
Scratchy	1416	1348	0
Tomte	464	412	0
Violin_mode	1888	1648	0
Wandering_line	176	168	0
Whistle	1220	1196	4
None_of_the_above	352	324	8

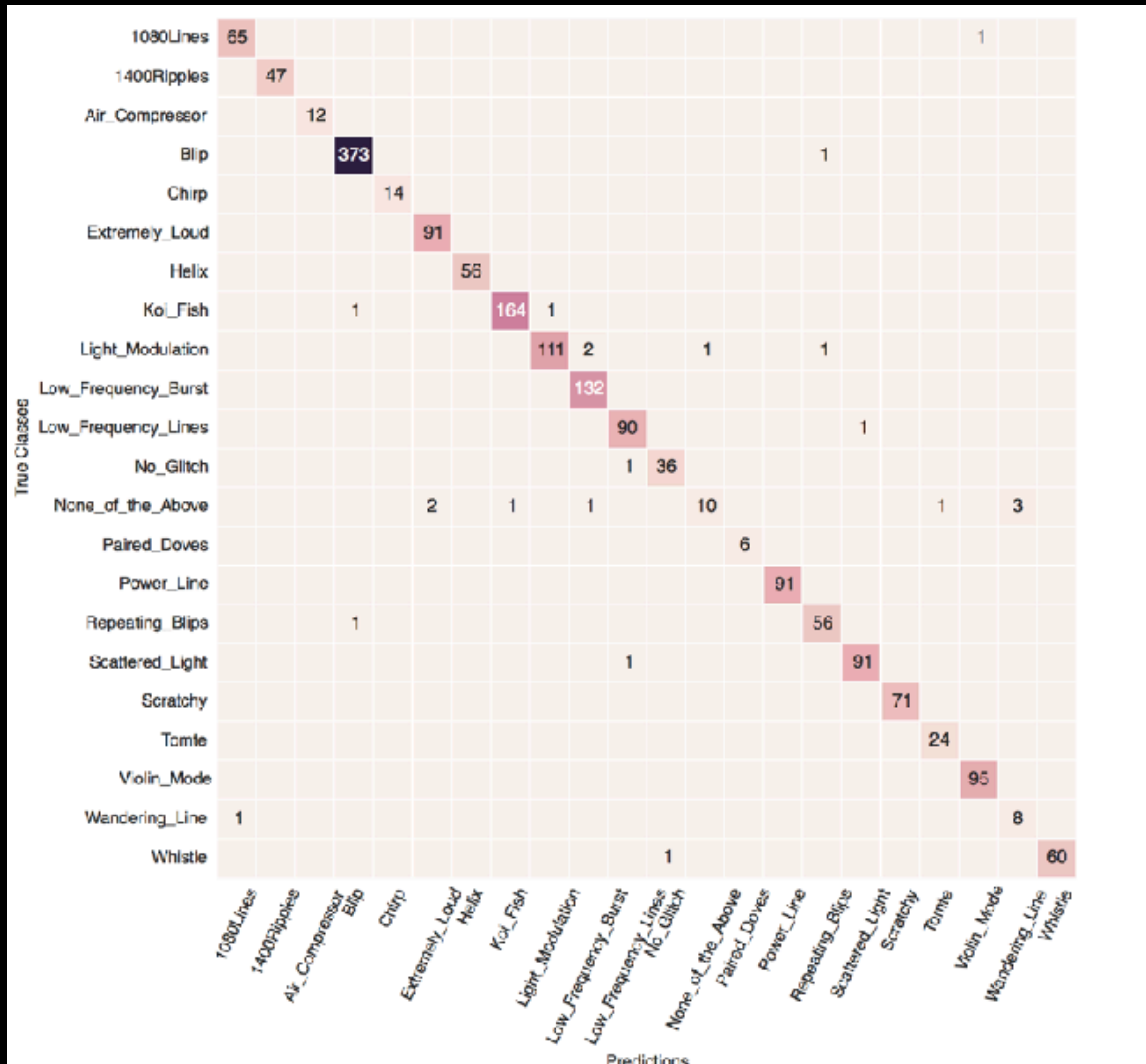
1. Small sample
2. Imbalanced data

TACKLING WITH TRANSFER LEARNING



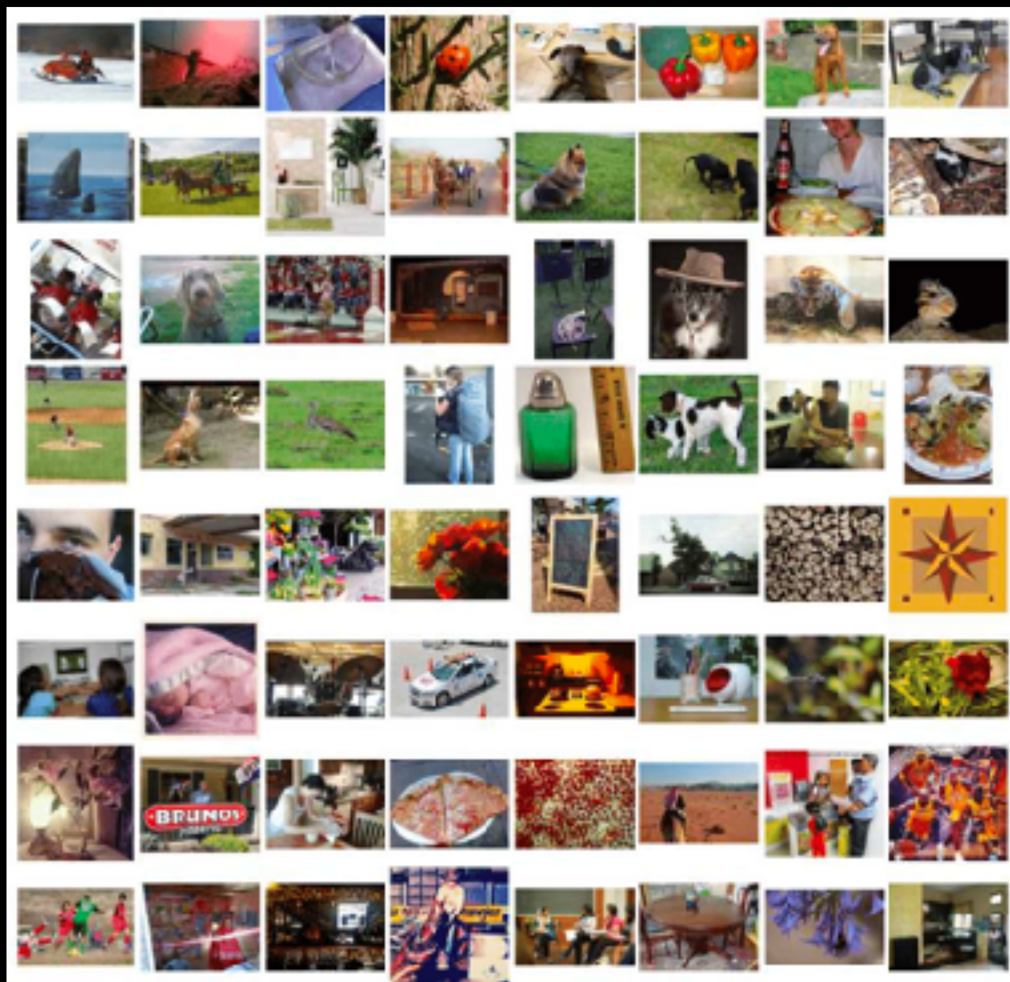
- Using the large data set of images with wide diversity, e.g. ImageNet, to train the network. Then modifying the final layer based on the number of required classes and then fine-tuning the parameters on the original data

TACKLING WITH TRANSFER LEARNING

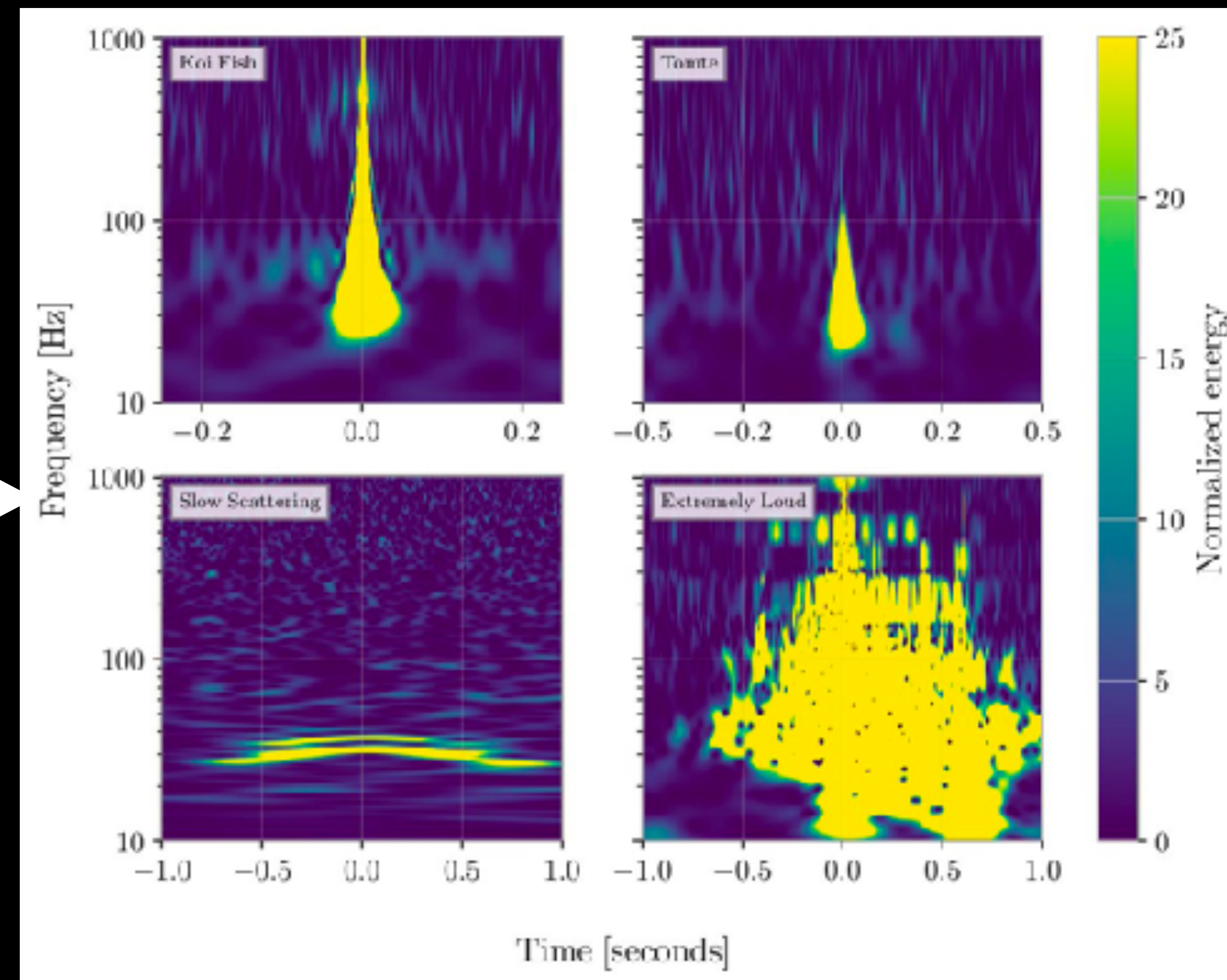


George et al. (2018)

POSSIBLE PROBLEM OF PRE-TRAINING WITH IMAGENET



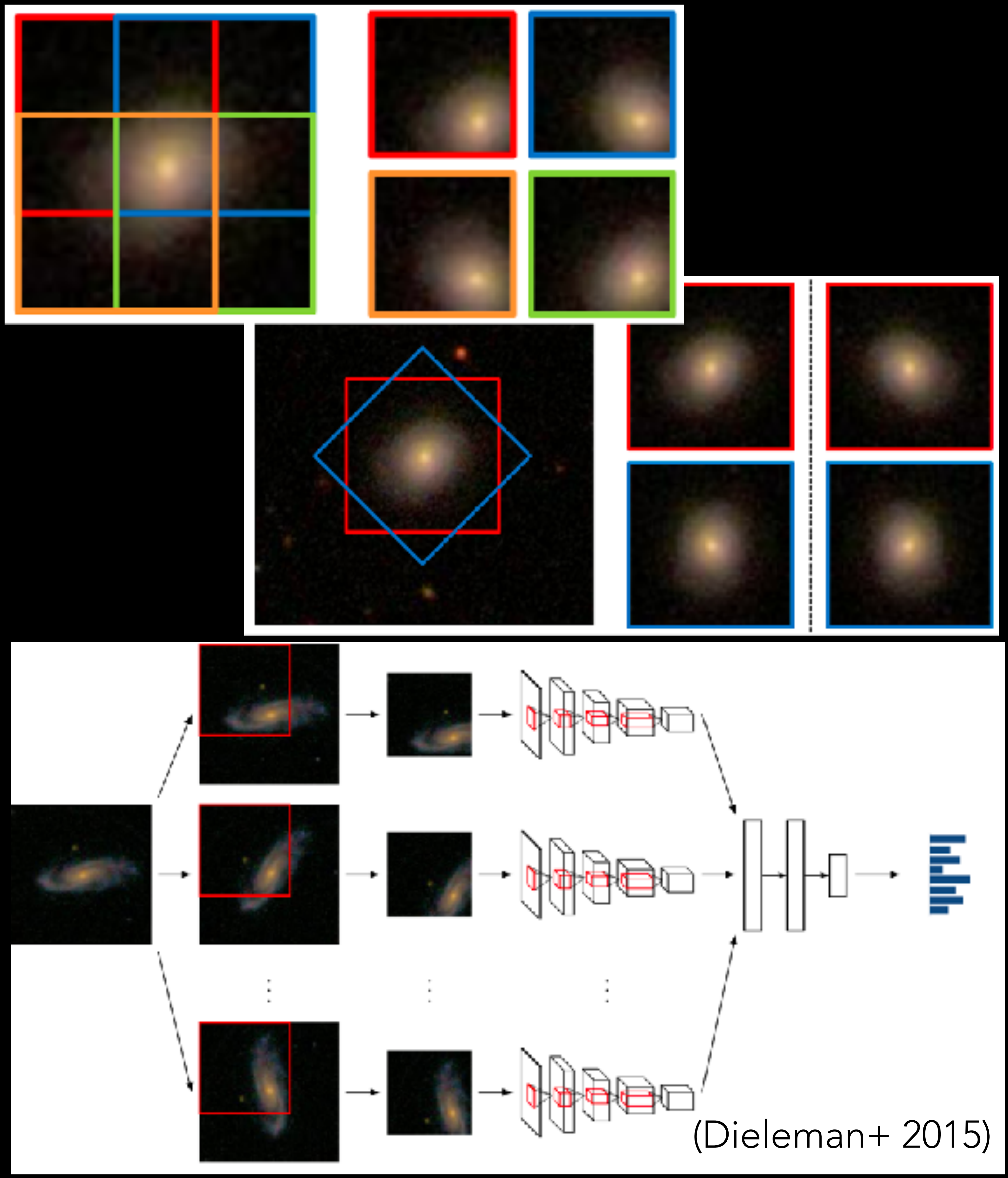
?



- The underlying assumption of transfer-learning is that the pre-training data set is sufficiently similar to the data in your problem.
- However, there is a fundamental difference between normal images and spectrograms. - Lack of symmetry

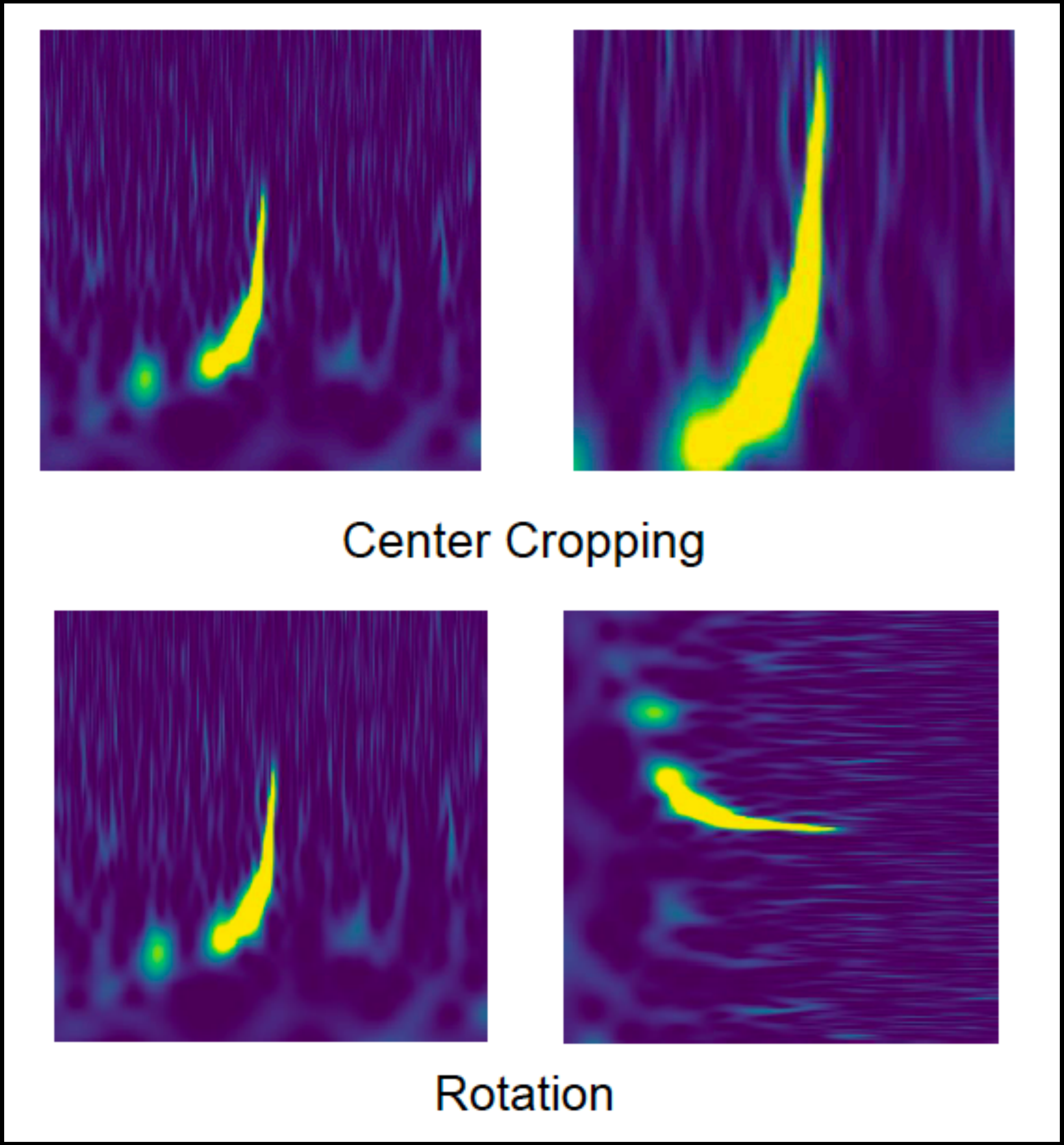
DATA AUGMENTATION

- For images of galaxies, rotating images should not affect its classification (i.e. rotation invariant).They are also scale, reflectional and translational invariant.
- All these invariances can be exploited to data augmentation - creating new training data by perturbing existing data points.
- The topological structures of the perturbed images are consistent with he original data.
- The increased diversity can result in less overfitting and better generalization.



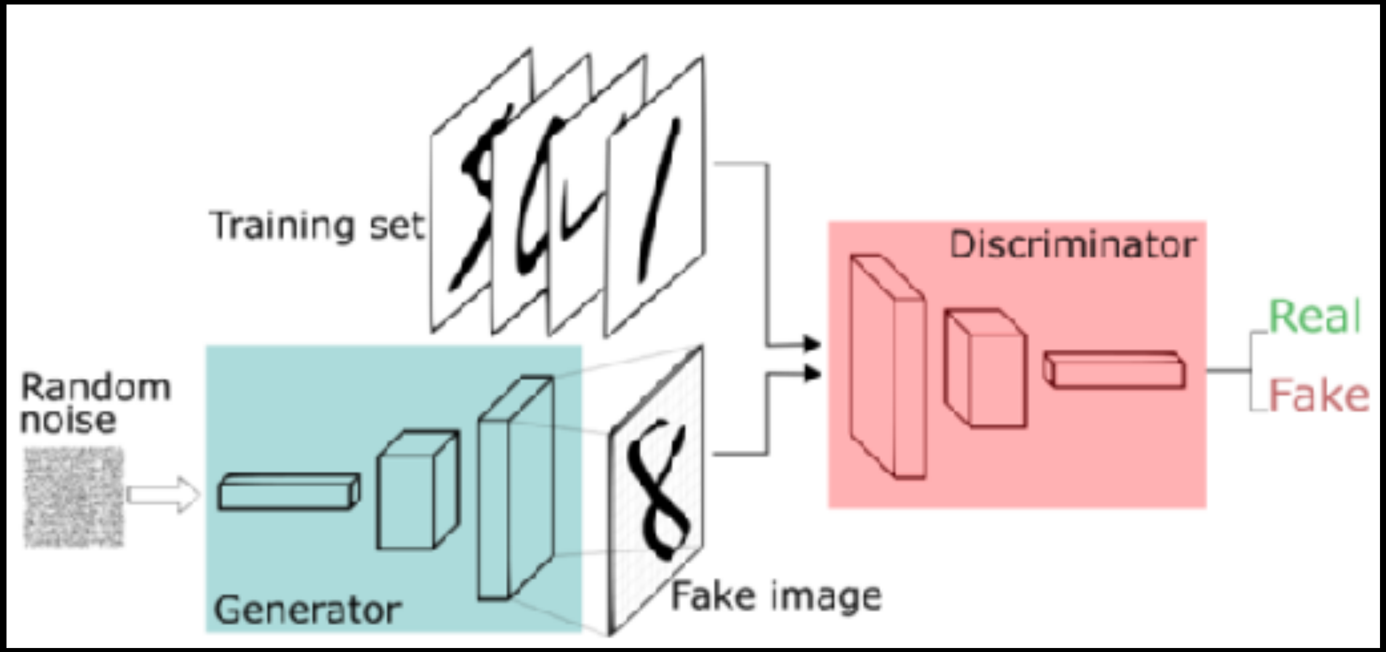
(Dieleman+ 2015)

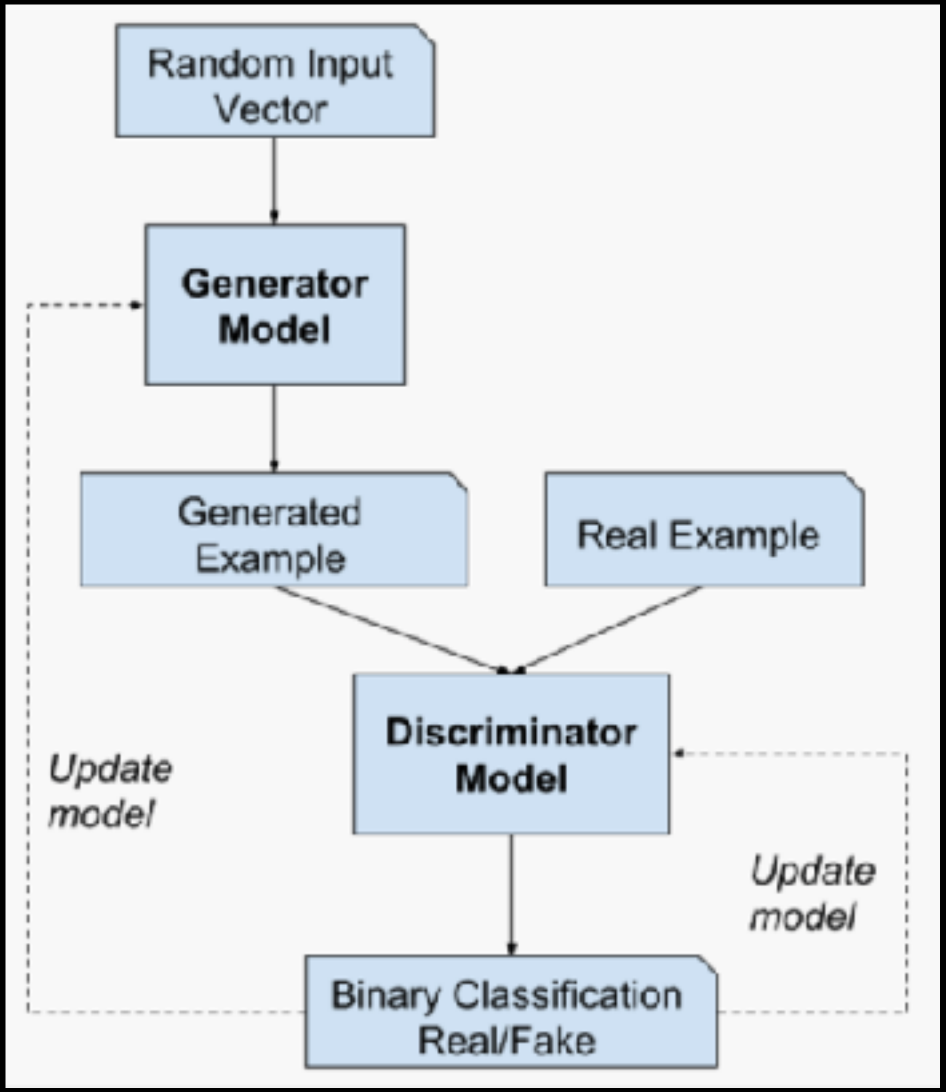
DATA AUGMENTATION



- Different from images, each dimension of spectrogram represents different quantities
- Lack of symmetries
- Conventional data augmentation algorithms for images CANNOT be applied to spectrogram

GENERATIVE ADVERSARIAL NETWORK (GAN)





- Conceived by Goodfellow in 2014.
- GAN is built based on game theory played by two networks.
- Generator and Discriminator are put together in a two-player zero-sum game.

USE GAN IN COLOURISING B&W IMAGES

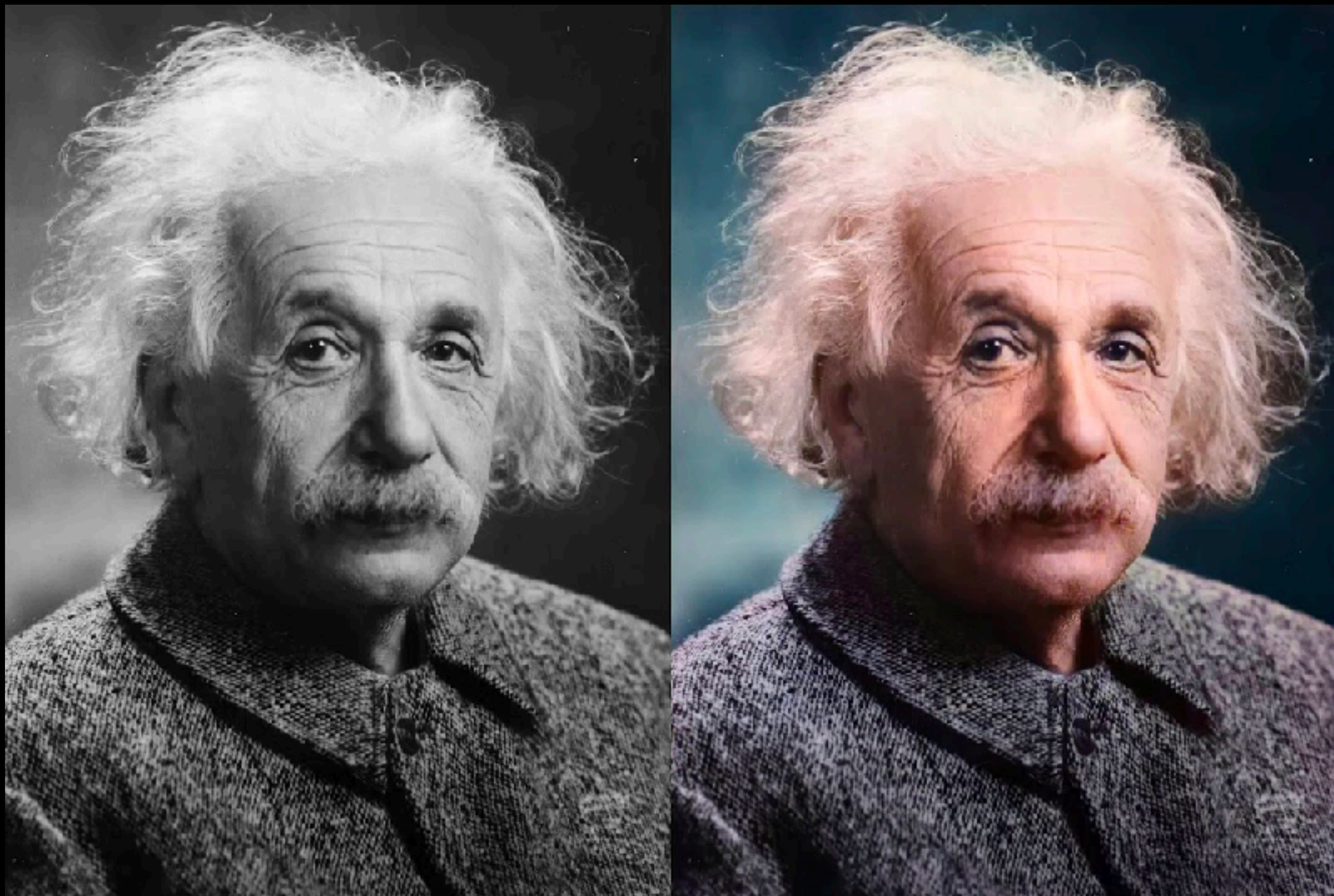
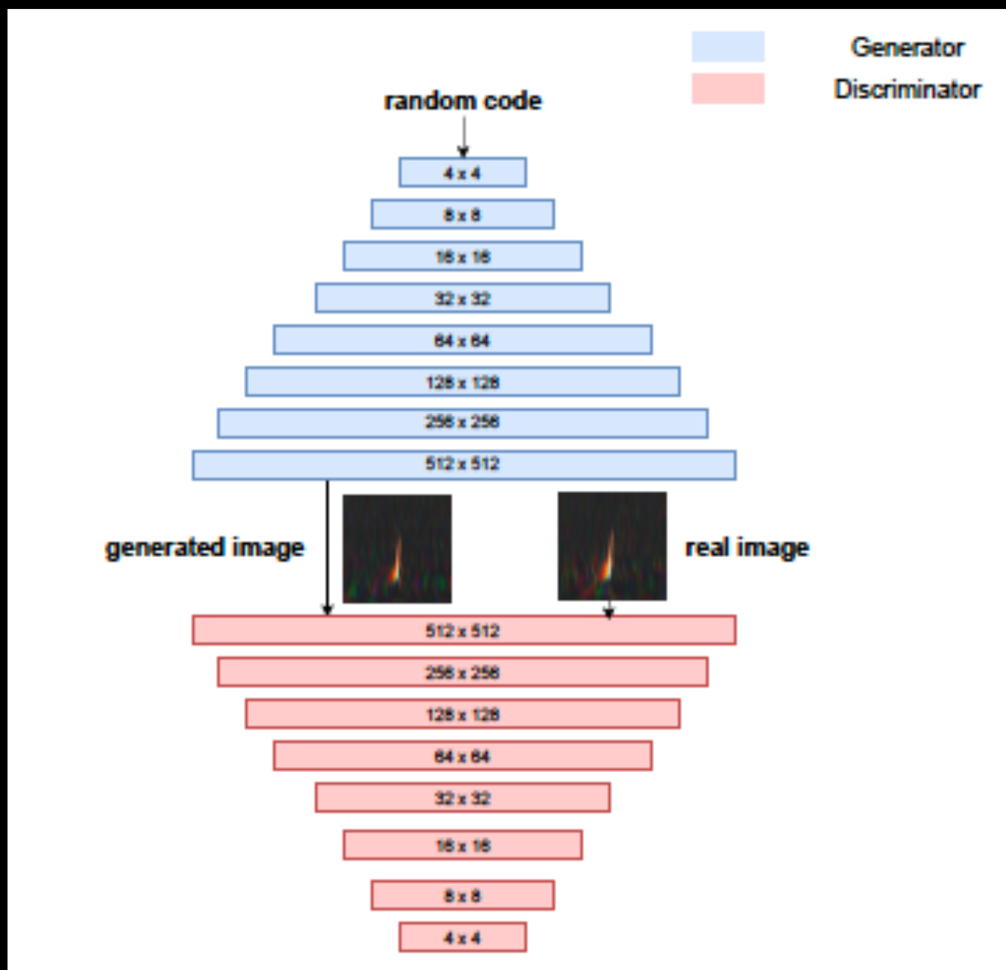
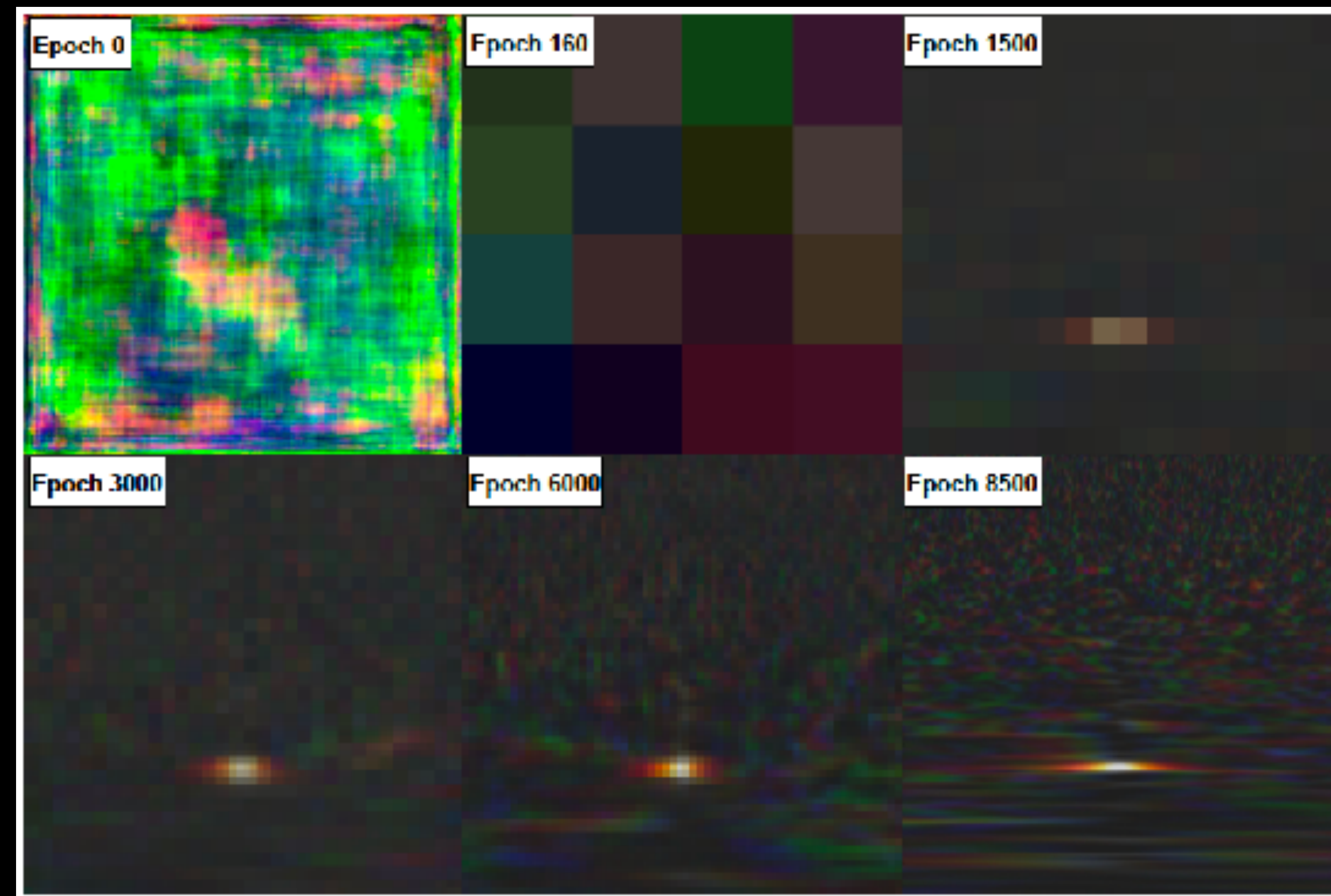


Image Courtesy: DeOldify

PROGRESSIVE GENERATION OF SPECTROGRAM

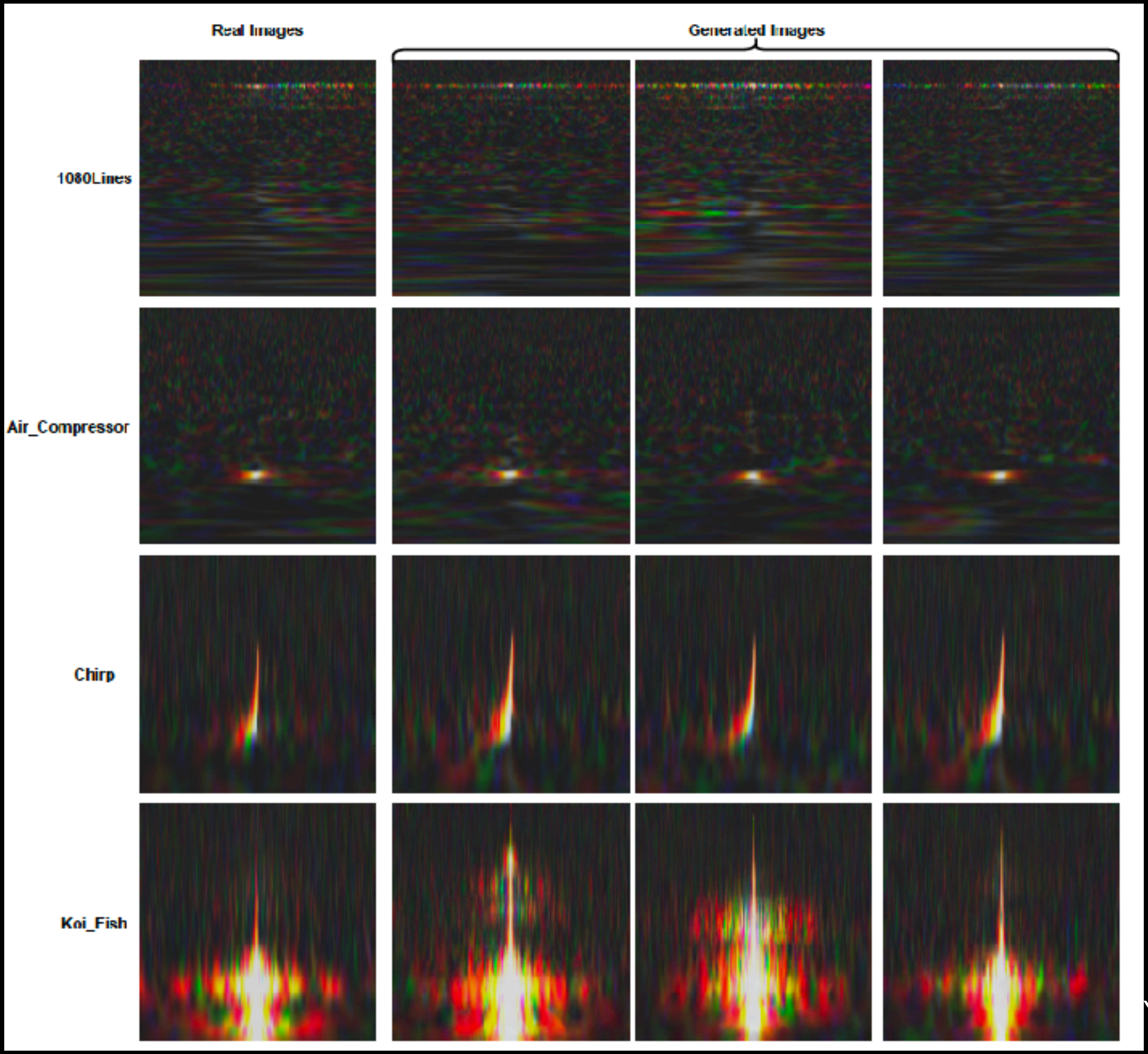




Yan et al. submitted

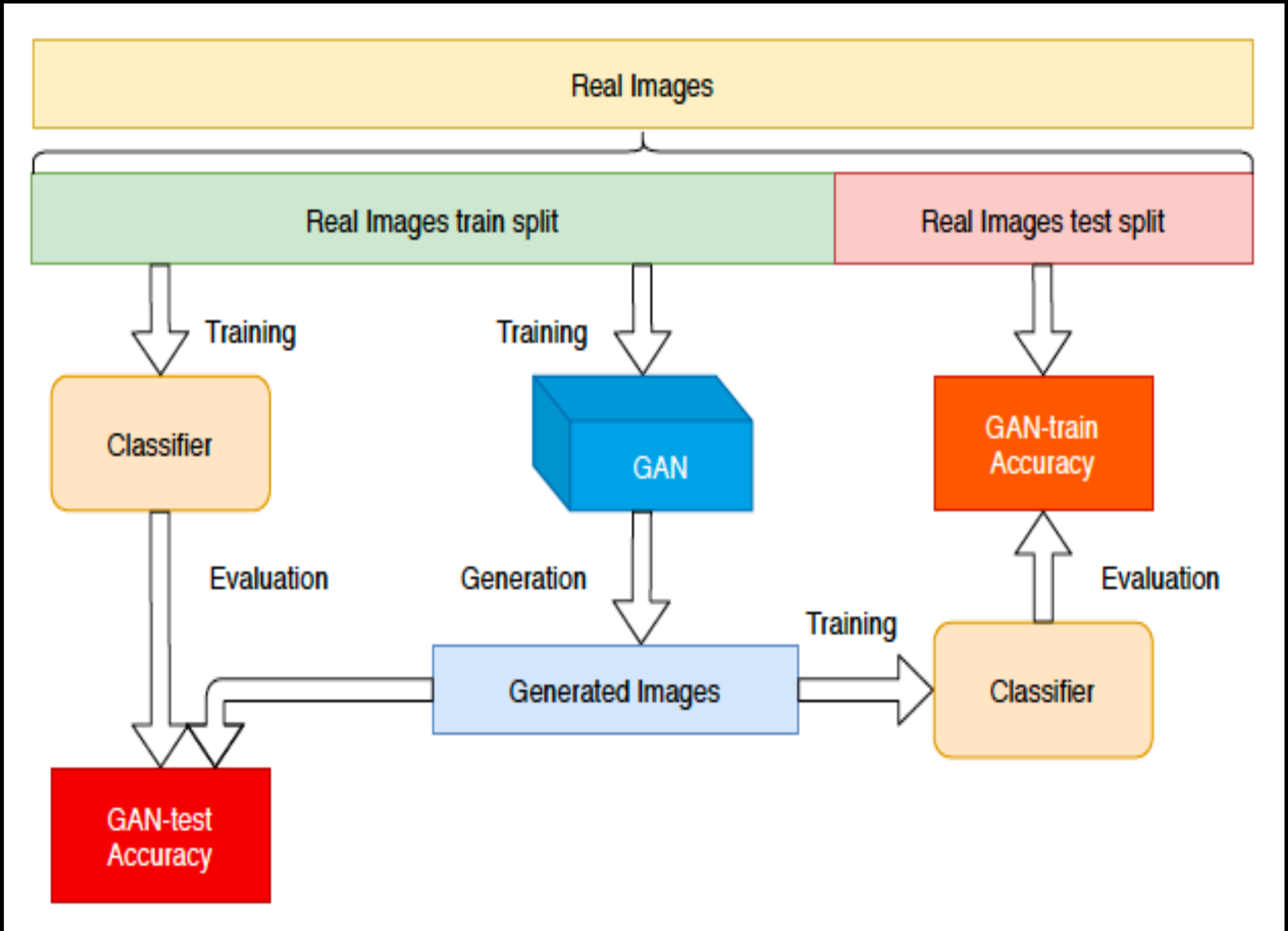
- Progressive growing of GAN (ProGAN) gradually grows layers with increasing resolution.

EXAMPLES OF GENERATED SPECTROGRAMS



Yan et al. submitted

MEASURING PERFORMANCE OF GAN



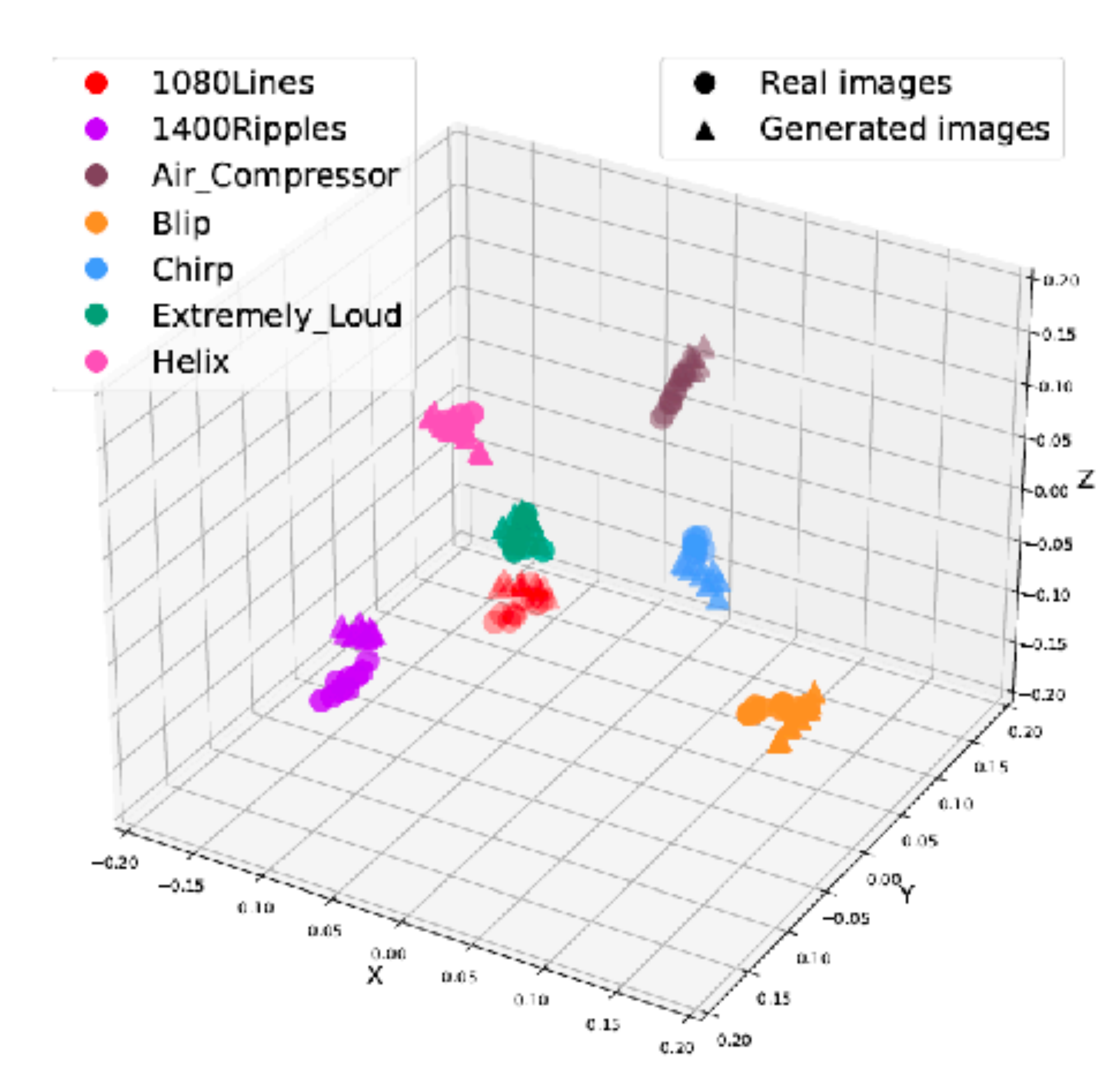
- GAN-test score measures quality of images.
- GAN-train score measures diversity

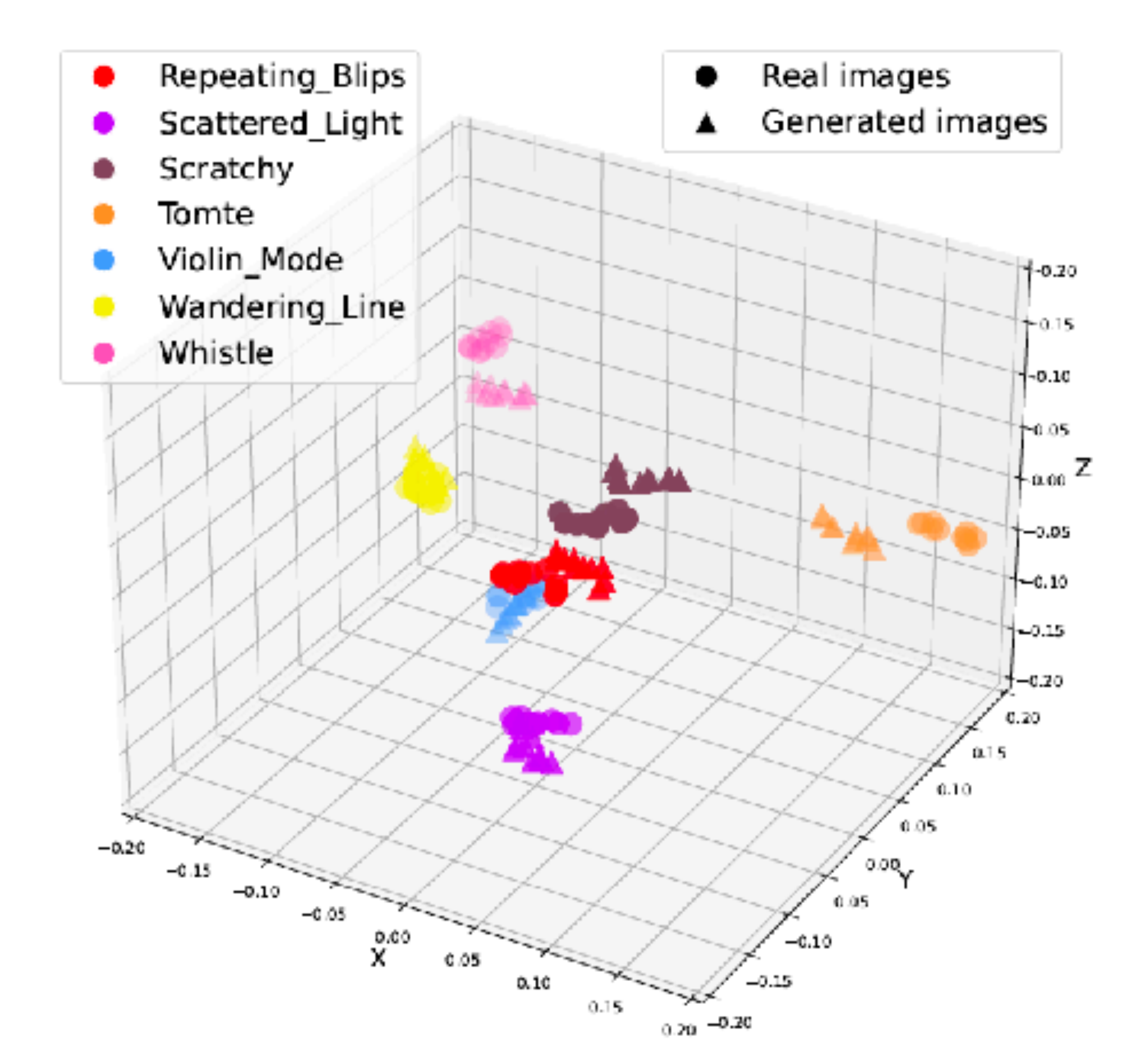
		ResNet50 (244×244)	ResNet101 (244×244)	Inception-V3 (299×299)	GoogLeNet (244×244)
single-duration spectrograms (1.0s)	GAN-train	81.00(%)	**91.55**(%)	73.40(%)	66.05(%)
	GAN-test	97.57(%)	97.18(%)	91.83(%)	**98.68**(%)
multi-duration RGB spectrograms	GAN-train	97.88(%)	97.64(%)	**98.43**(%)	98.35(%)
	GAN-test	99.31(%)	99.35(%)	95.87(%)	**99.73**(%)

Yan et al. submitted

MEASURING PERFORMANCE OF GAN

Visualizing the clustering properties in low-dimensional space by t-SNE.





Yan et al. submitted

RESOLVING THE IMBALANCE DATA PROBLEM

Algorithm 1: Data Augmentation Algorithm

Input: the number of real images in one class, N_{real}

Output: the number of generated images in one class, N_{aug}

Set the total number of images in one class, N_{total} , to 500, 1000, 2000, 3000 and 5000, respectively;

For each class, ;

if $N_{real} < N_{total}$ **then**

$N_{aug} = N_{total} - N_{real}$;

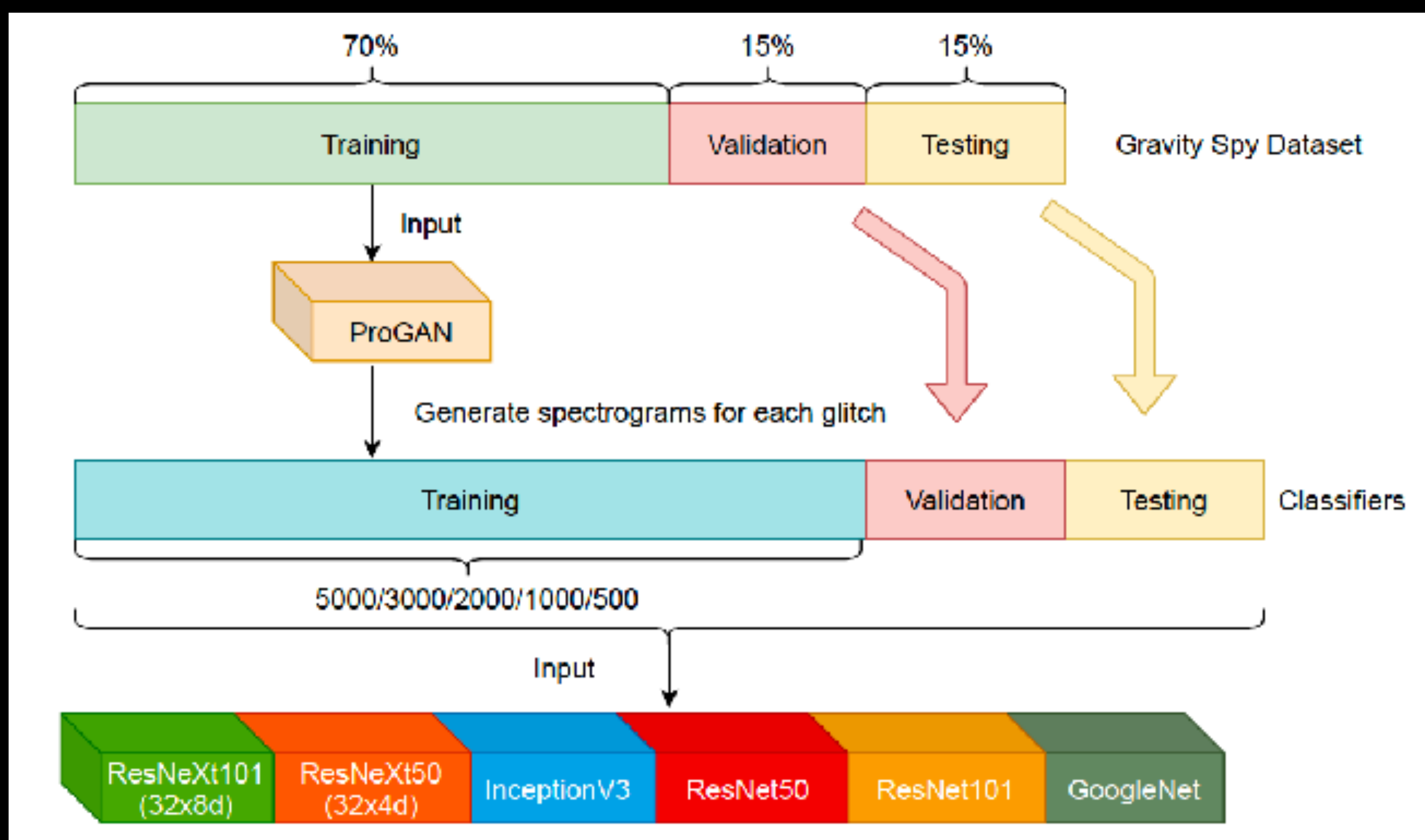
else

$N_{aug} = 0$;

end

Return the number of generated images, N_{aug};

Yan et al. submitted

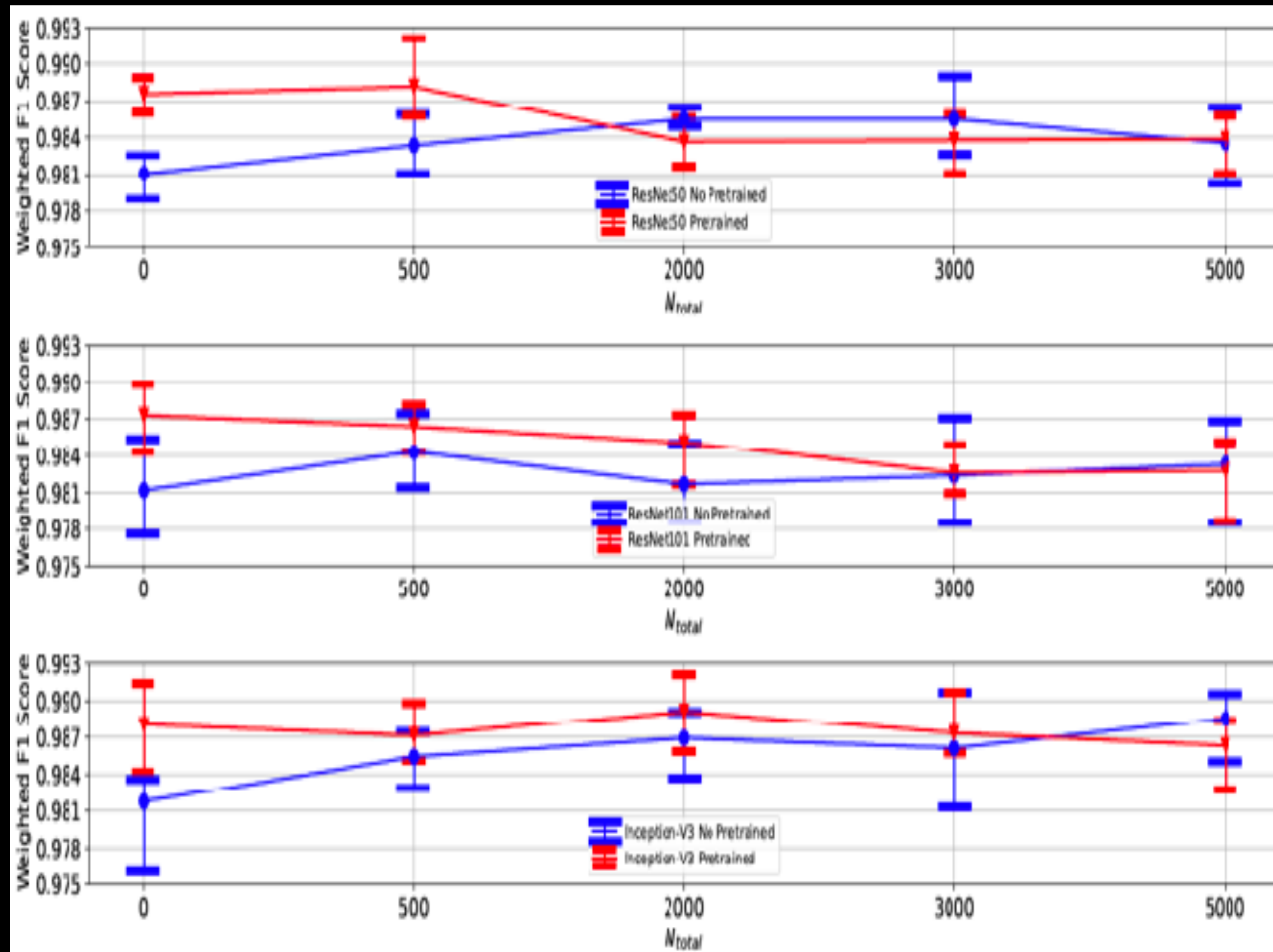


EXPERIMENTAL RESULTS

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$



Yan et al. submitted

EXPERIMENTAL RESULTS

Without Pre-train $N_{\text{total}} = 0$

	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5
Air Compressor	1	0.9412	1	1	1
Paired Doves	0.6667	0.6667	1	0.6667	0.6667
Wandering Lines	0.8333	0.7273	1	0.7143	0.9231

With Pre-train $N_{\text{total}} = 0$

	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5
Air Compressor	1	1	1	1	0.9412
Paired Doves	1	1	1	1	1
Wandering Lines	1	1	1	0.9231	1

Without Pre-train $N_{\text{total}} = 5000$

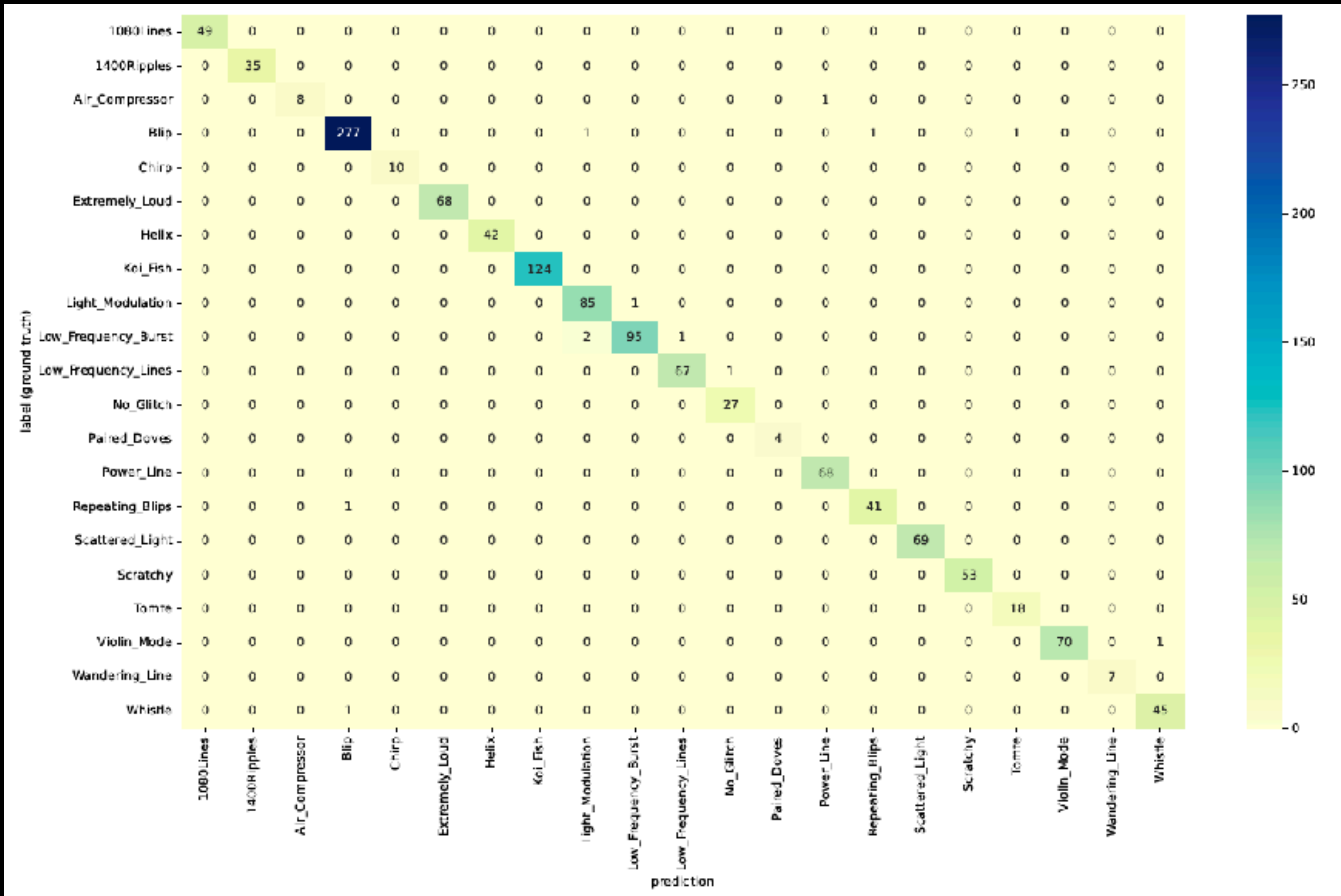
	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5
Air Compressor	1	0.9412	0.9412	0.9412	0.9412
Paired Doves	1	1	1	1	1
Wandering Lines	1	1	1	1	0.9333

Yan et al. submitted

EXPERIMENTAL RESULTS

Yan et al. submitted

Without Pre-train $N_{\text{total}} = 2000$



- Achieve perfect precision and recall on 7 classes.

MERITS OF GAN IN GW ASTRONOMY

1. Provide a framework for data augmentation

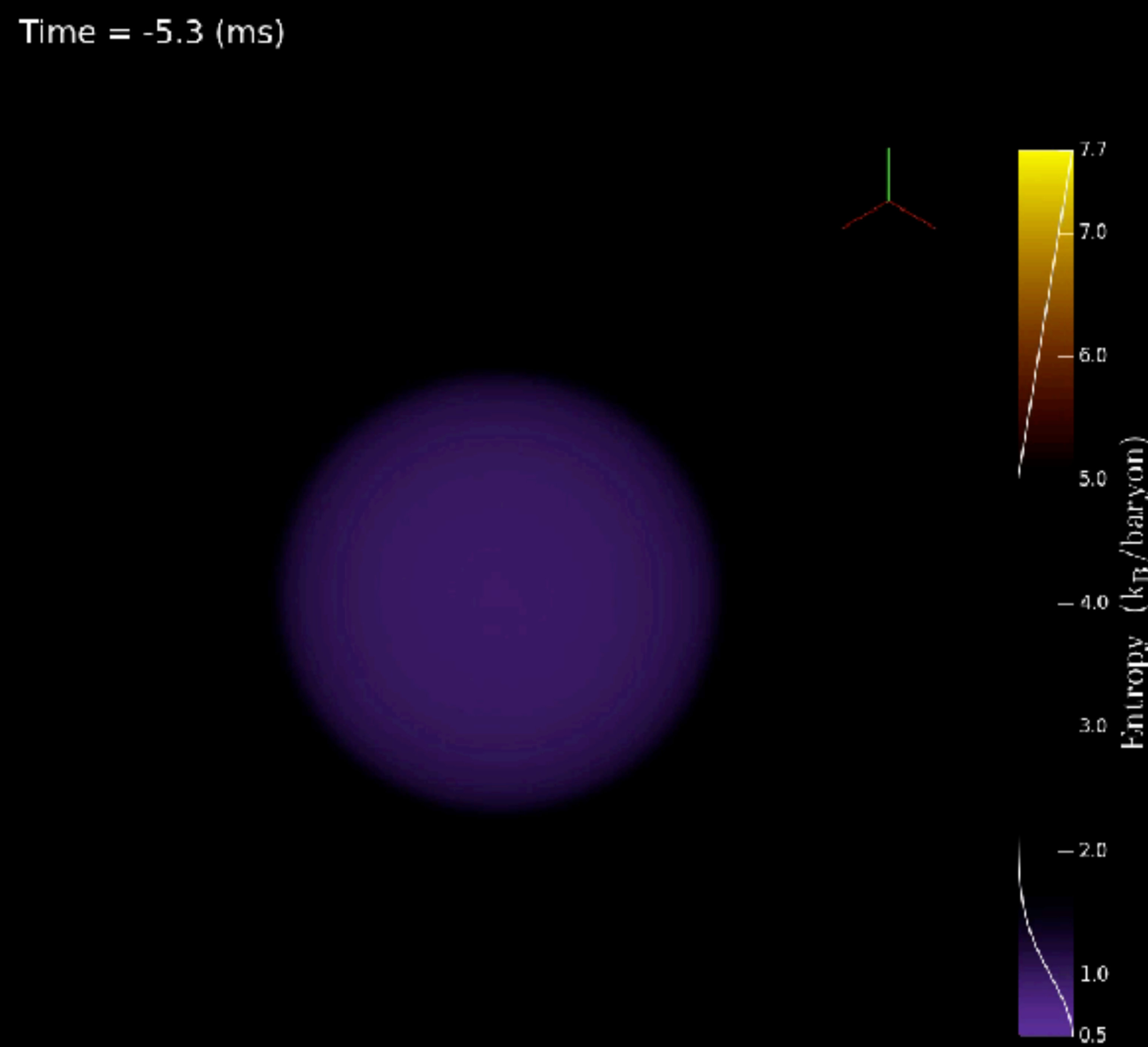
2. Enable an alternative means for employing deep learning without pre-training.

3. Improve the classification performance particularly for the classes with small samples.

FUTURE PLAN - GW FROM CCSN

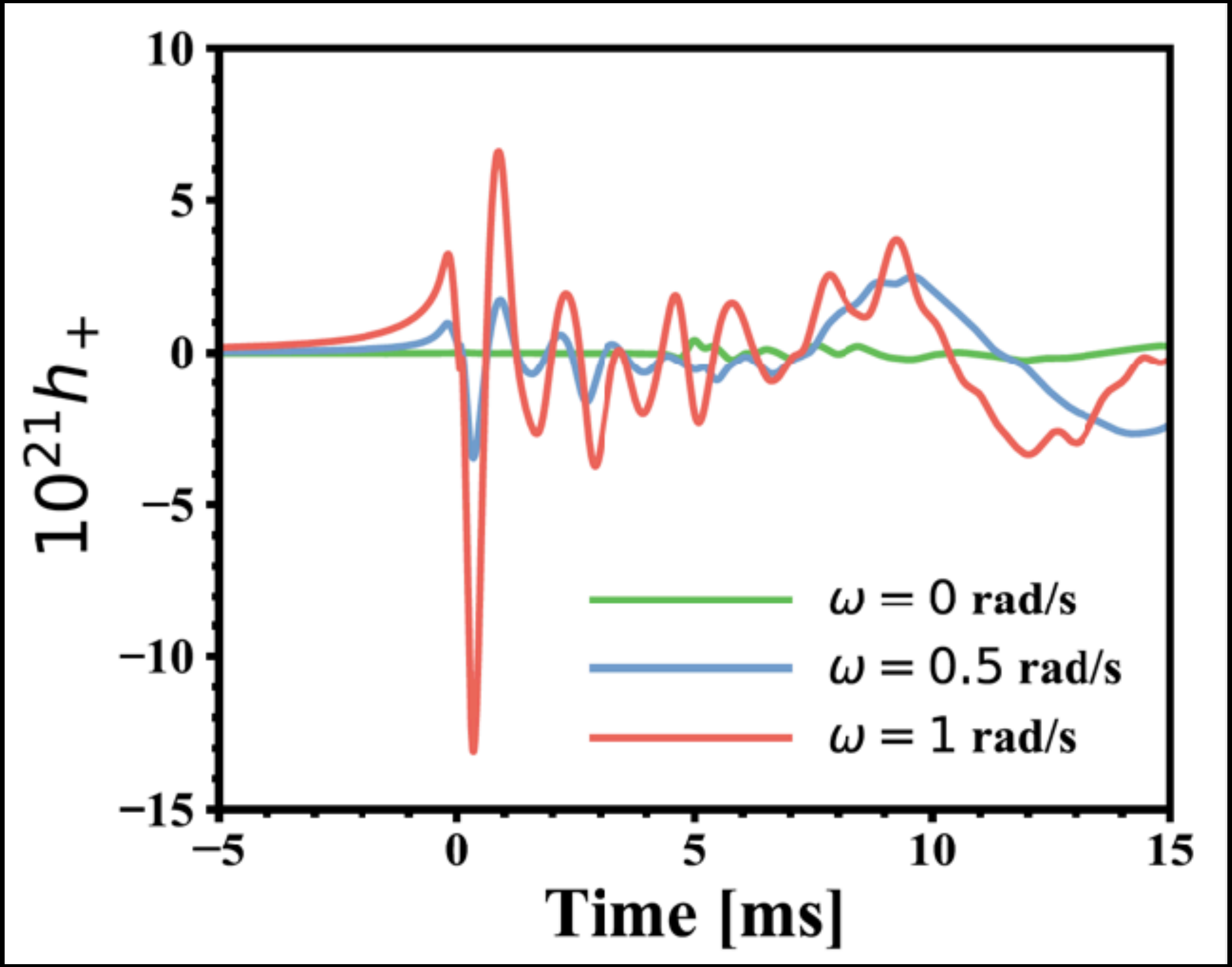
- Detection of GW from core-collapsed SNe will be the next milestone.
- i) Neutrino-driven, ii) Magnetorotational mechanism
- Majority of observed CCSNe can be explained by neutrino-driven mechanism
- This can cause convection and standing accretion shock instability (SASI)

NEUTRINO-DRIVEN CCSN

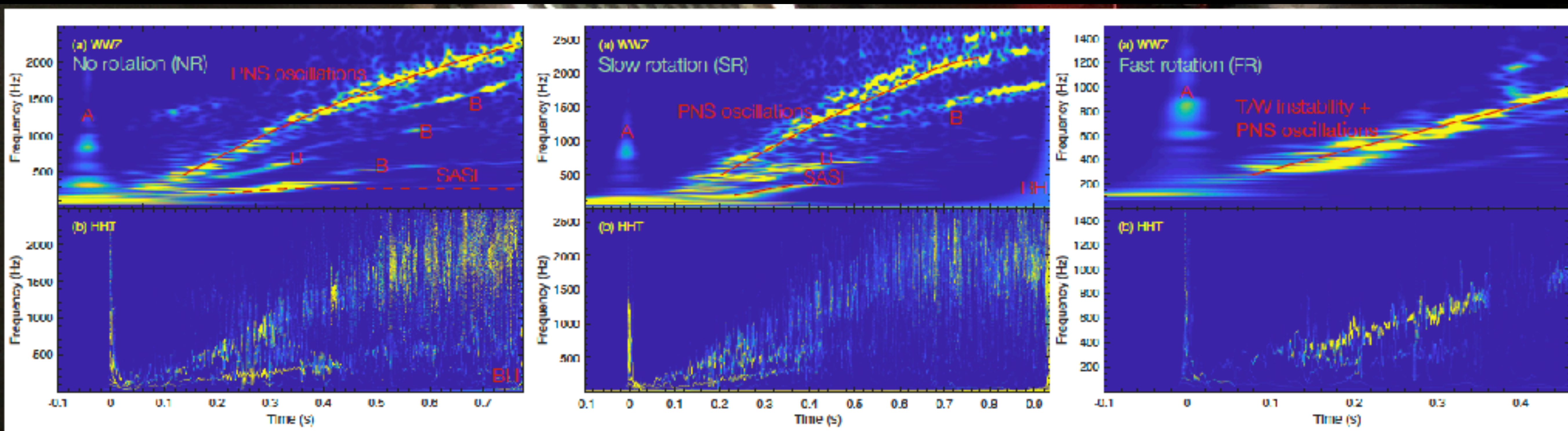


Simulation Courtesy: Kuo-Chuan Pan (NTHU)

GW FROM CORE-COLLAPSED SN



GW FROM CORE-COLLAPSED SN



CHALLENGE IN SEARCHING GW FROM CCSN

- Since the GW signal produced by CCSNe is affected by turbulence, it is expected to be stochastic in nature.
- The signal evolution cannot be deterministically predicted.
- This prevents matched filtering from being applied to CCSNe.
- Deep learning techniques have been explored to detected and classify the possible GW signals from CCSNe.

SEARCH GW FROM CCSN BY DEEP LEARNING

However, the training sample (simulated) is limited, unbalanced and can't cover a wide parameter space.

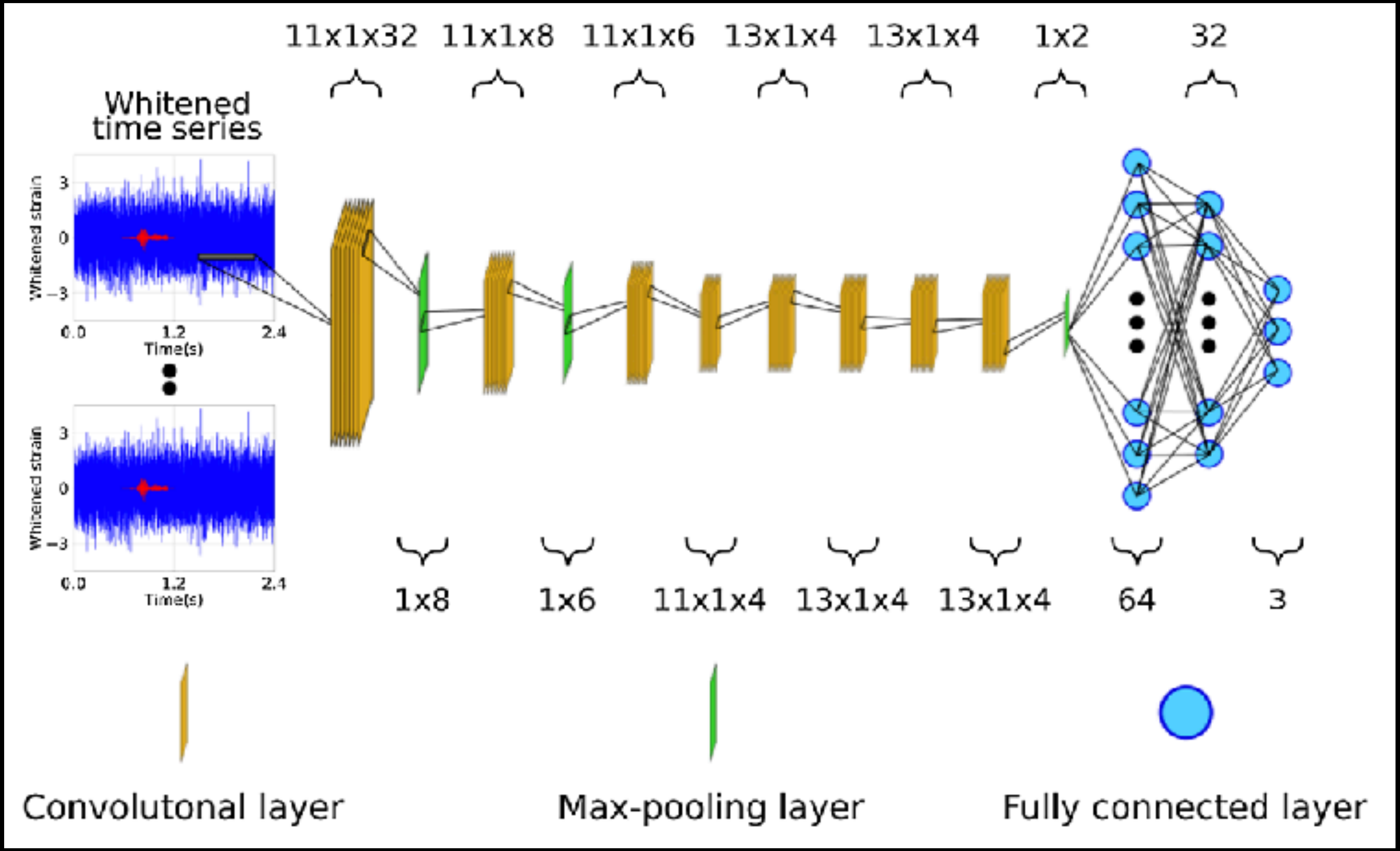
	Mechanism	Mass ($M_\odot$)	No.
Abdikamalov [86]	M	12.0	92
Dimmelmeier [87]	M	11.2,15.0,20.0,40.0	136
Richers [88]	M	12.0	1824
Andresen [89]	N	15.0	6
Kuroda [90]	N	11.2,15.0	2
Muller [91]	N	15.0,20.0	6
Murphy [96]	N	12.0,15.0,20.0,40.0	16
Ott_1 [43]	N	15.0	2
Ott_2 [97]	N	27.0	8
Powell [92]	N	3.5^*,18.0	2
Radice [93]	N	9,10,11,12,13,19,25,60	8
$\mathrm{Yakunin}_1$ [94]	N	12,15,20,25	4
$\mathrm{Yakunin}_2$ [95]	N	15	1

*Mass of a star in a binary system with an initial helium mass of 3.5 $M_\odot$

Chan, Heng & Messenger (2020)

SEARCH GW FROM CCSN BY DEEP LEARNING

True alarm rates < 66% are obtained from this experiment
(depends on distance & waveform)



Chan, Heng & Messenger (2020)

IMPROVE THE PERFORMANCE WITH GAN

- The trained network can perform better when presented with waveforms of unexpected features, which is more likely in reality.
- This can be achieved by training with a sample of larger diversity.
- GAN has the potential to enhance the detectability of CCSN GW signal by deep learning, altogether with relieving the problems of small and imbalanced training sample.

SUMMARY

- We have been searching the novel methods for improving the GW signal processing.
- ARIMA model provides an efficient and flexible way to denoise the data, while retaining the physically interesting signal.
- GAN can enhance the deep learning performance by uplifting the problems of small and imbalanced training sample. Also, it can improve the generalization by providing the data with larger diversity
- In our future projects, these methods can be combined in searching for GW signals from various sources (e.g. continuous wave from NS, bursts from CCSNe).

THANK YOU VERY MUCH