中山大學天琴中心
TIANQIN CENTER FOR GRAVITATIONAL PHYSICS, SYSU

Detecting gravitational waves from extreme mass ratio inspirals using convolutional neural networks

Xue-Ting Zhang[1], Chris Messenger[2], Natalia Korsakova[3], Man Leong Chan[4], Yi-Ming Hu[1], Jian-dong Zhang[1]

1 TianQin Research Center for Gravitational Physics & School of Physics and Astronomy, Sun Yat-sen University

2 SUPA, School of Physics and Astronomy, University of Glasgow

3 Artemis, Observatoire de la Côte d’Azur, Boulevard de l’Observatoire

4 Department of Applied Physics, Fukuoka University

Phys. Rev. D 105, 123027(2022)

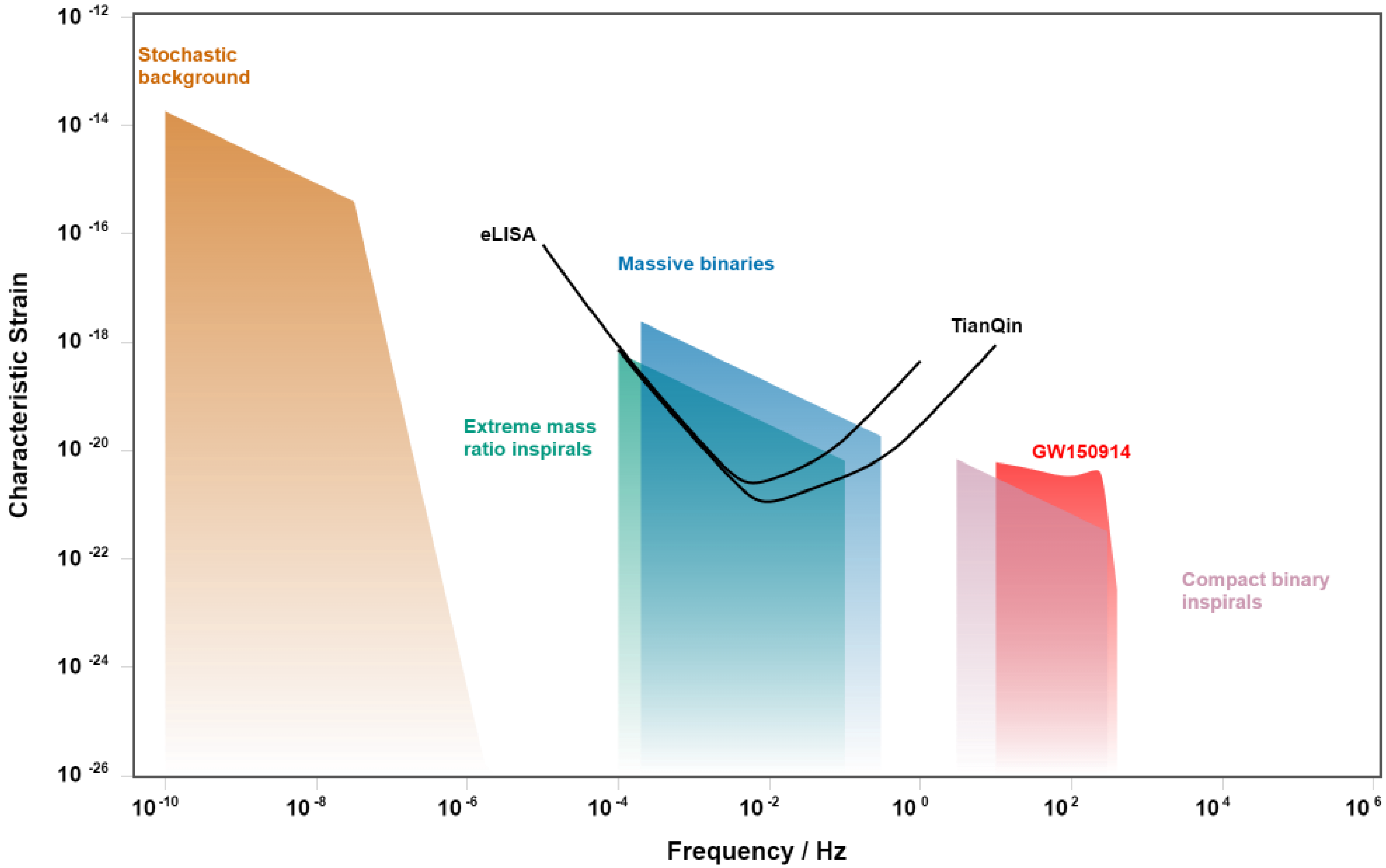
LISA Data Analysis & Machine Learning Workshop, 2022.

Outline

1. Background
2. Methods & Results
3. Conclusion

Background

◻ GW sources and Space-borne detectors

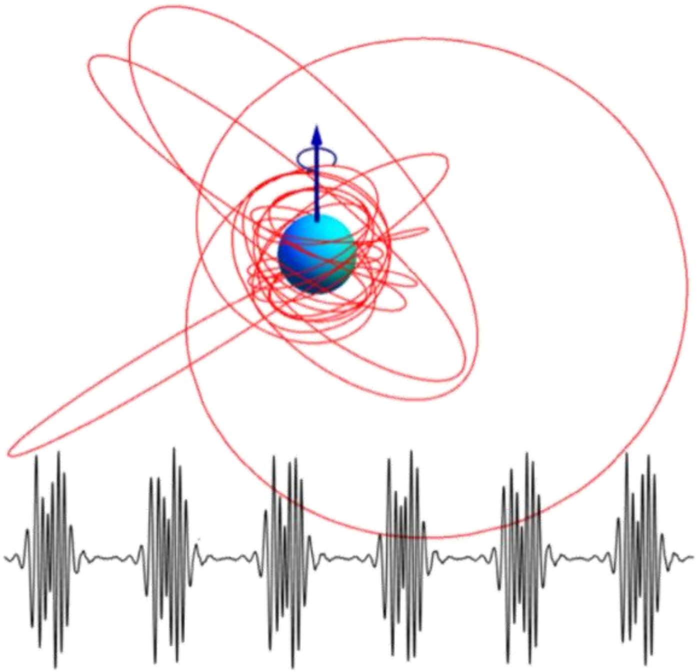


Christopher Cole et al, [** Gravitational Wave Sensitivity Curve Plotter ** (gwplotter.com)]

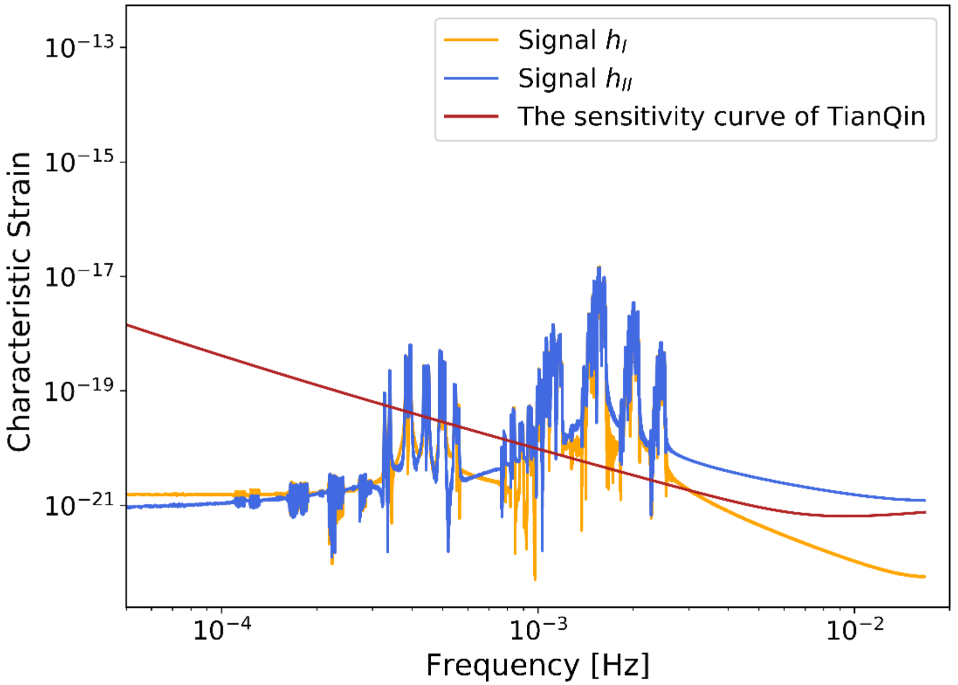
Extreme Mass Ratio Inspiral (EMRI)

◻ CO-MBH system[1]

- TianQin can observe O(1)-O(100) GW events [2]。
- ideal laboratories to study gravity in a strong regime.



V. Cardoso et al., arXiv:1908.11390



Responded signals from TianQin, 3 months long

[1] Amaro-Seoane, P. LRR. 2018, 21, 4.

[2] Hui-Min Fan et al. PRD. 2020, 102: 063016.

Challenges to EMRI signal detection

◻ Waveform modeling

- requirement： accurate, efficient, extensive

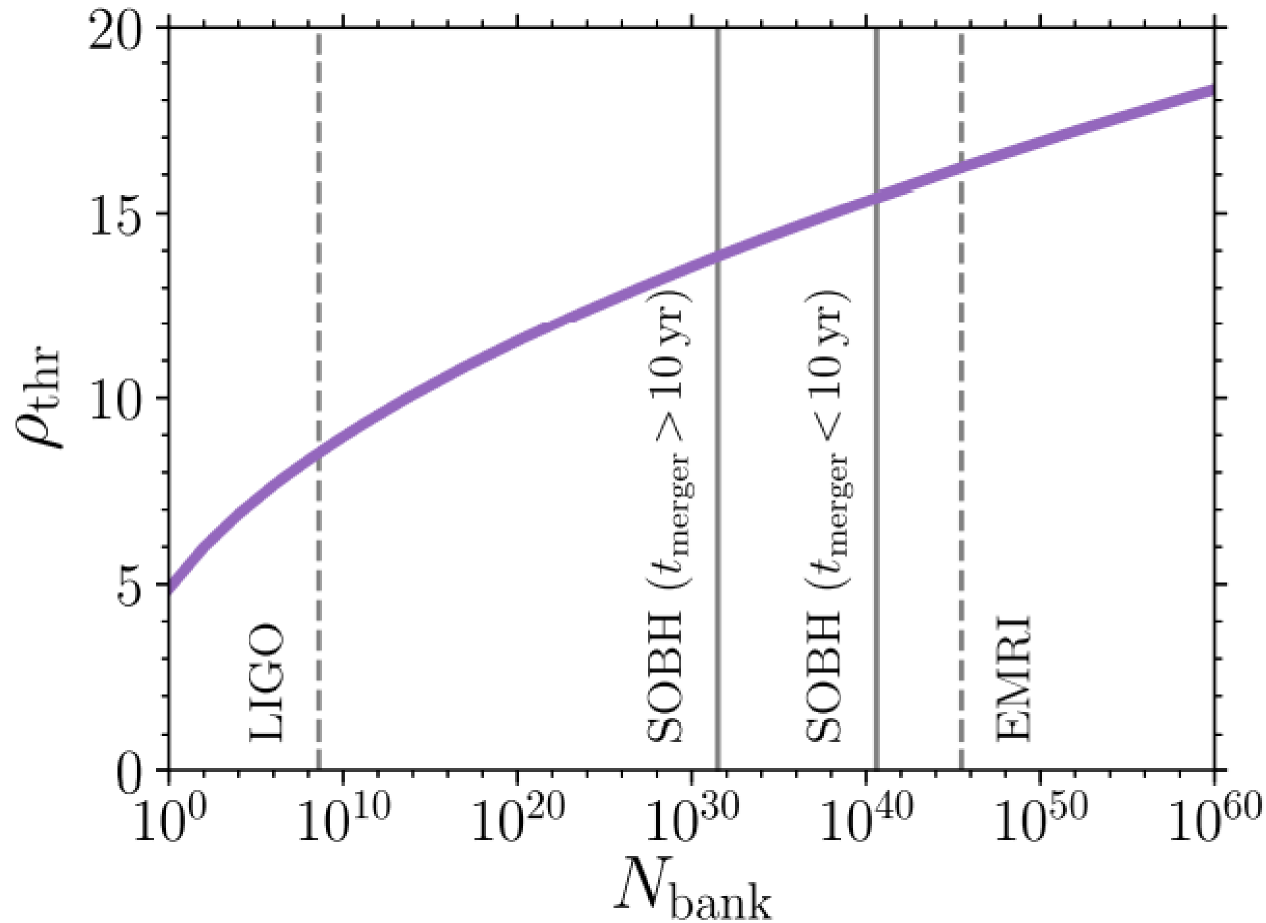
waveform	paper	difficulties
Kludge waveform	AK: Leor Barack and Curt Cutler. PRD 69.8,; NK: [2] Stanislav Babak et al. PRD 75, 024005; AAK: Alvin J K Chua et al. CQG 32(2015) 232002;	Most of the widely used waveform models are expected to quickly dephase from the physical waveform
Self-force	[1] Poisson, E. LRR. (2004) 7: 6 [2] A. Pound, et al arxiv:1908.07419 [3] L. Steve Drasco et al. PRD 73, 024027; [4] L. Barack, CQG 26, 213001 (2009). [5] M. Van De Meent, PRD 97, 104033 (2018). [6] J. Miller, et al. PRD 103, 064048 (2021) [7] S. A. Hughes, et al. PRD 103, 104014 (2021) [8] J. McCart, et al. PRD 104, 084050 (2021),	
others	PW: [1] Yan Wang, et al. PRD, 2012, 86: 104050. FEW: [2] Michael L. Katz, et al PRD 104, 064047	

- An ideal EMRI search method should be versatile enough, so that even though it was tuned under kludge waveforms, it can still be effective for a real signal.

Challenges to EMRI signal detection

Signal Detection – matched filtering

- **The template bank is huge. Both template-based algorithms and template-free methods have been proposed to detect the EMRI signals.**



Christopher J. Moore,2019

CNNs detect GW signals

paper	content
George (2018)	Using CNN to detect real BBH signals
Gabbard (2018)	Using simulated BBH signals, they compare the performance between matched-filtering and CNN.
Schäfer (2020), Chan(2020), Bayley(2020)…	Using CNN to detects more complex and long-lived GW signals, like BNS, continuous GW.

Can we detect a EMRI signal by CNN-like machine learning algorithms?

Goal

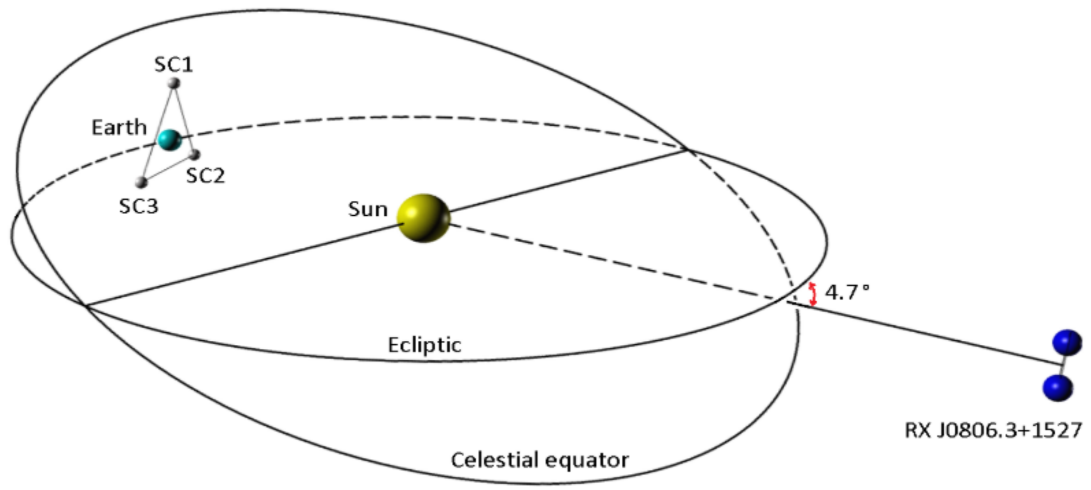
- ☐ **Detecting one EMRI signal buried on noise by using a CNN**

Decomposition	Contents
1.Data preparation	(1) Noise simulation (2) Signal simulation (3) Input sample
2. Signal Detection	(1) Training a CNN by given training data (2) Testing a trained CNN by different testing data

Outline

1. Background
2. Methods & Results
3. Conclusion

Detector Configuration



Geocentric orbit, orientated to J0806.3+1527

Mission lifetime： 5 years

Arm length: $\sim 10^5$ **km**

Sensitive curve:

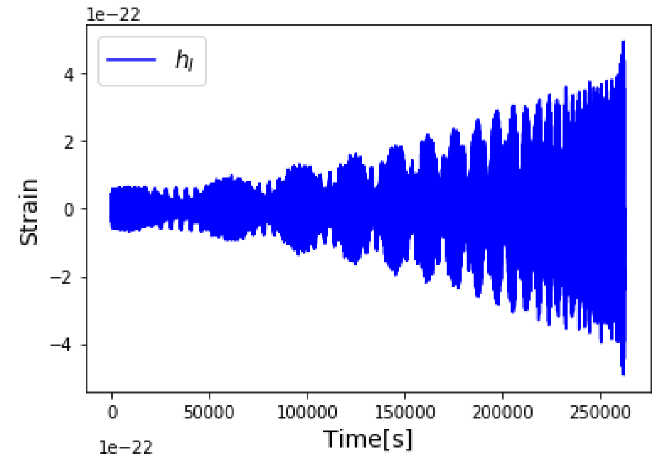
$$S_n(f) = \frac{1}{L_{arm}^2}\left[\frac{4S_a}{(2\pi f)^4}\left(\frac{1+10^{-4}}{f}\right)+S_x\right]\left[1+0.6\left(\frac{f}{f_*}\right)^2\right]. \quad \langle \tilde{\mathrm{n}}^*(f) * \tilde{\mathrm{n}}(f')\rangle = \frac{1}{2}S_{\mathrm{n}}(f) * \delta(f-f')$$

[1] Jun Luo et al 2016 Class. Quantum Grav. 33 035010

Signal Simulation

□ lower frequency approximation response

- **analytic kludge(AK) Waveform[1] :** $(M, \mu, \alpha, e_{lso}, \phi_0, \alpha_0, \lambda, \gamma_0, \nu_{lso}, \theta_S, \phi_S, \theta_K, \phi_K, t_c, D_L)$
- **Responded signals**[2]: $h_{I,II}(t) = F^{+}_{I,II} h_{+}(t) + F^{\times}_{I,II} h_{\times}(t)$



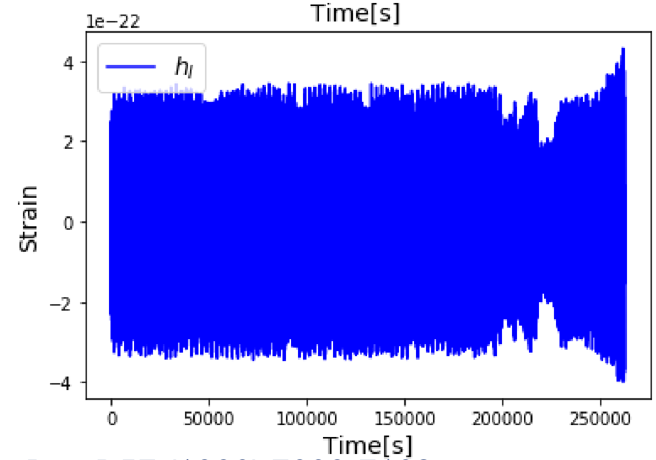
Antenna Pattern functions:

$$F_I^{+} = \frac{1}{2}\left(1 + \cos^2\theta\right)\cos 2\phi \cos 2\psi - \cos\theta \sin 2\phi \sin 2\psi$$

$$F_I^{\times} = \frac{1}{2}\left(1 + \cos^2\theta\right)\cos 2\phi \sin 2\psi + \cos\theta \sin 2\phi \cos 2\psi$$

$$F_{II}^{+} = \frac{1}{2}\left(1 + \cos^2\theta\right)\sin 2\phi \cos 2\psi - \cos\theta \sin 2\phi \sin 2\psi$$

$$F_{II}^{2} = \frac{1}{2}\left(1 + \cos^2\theta\right)\sin 2\phi \sin 2\psi + \cos\theta \sin 2\phi \cos 2\psi$$



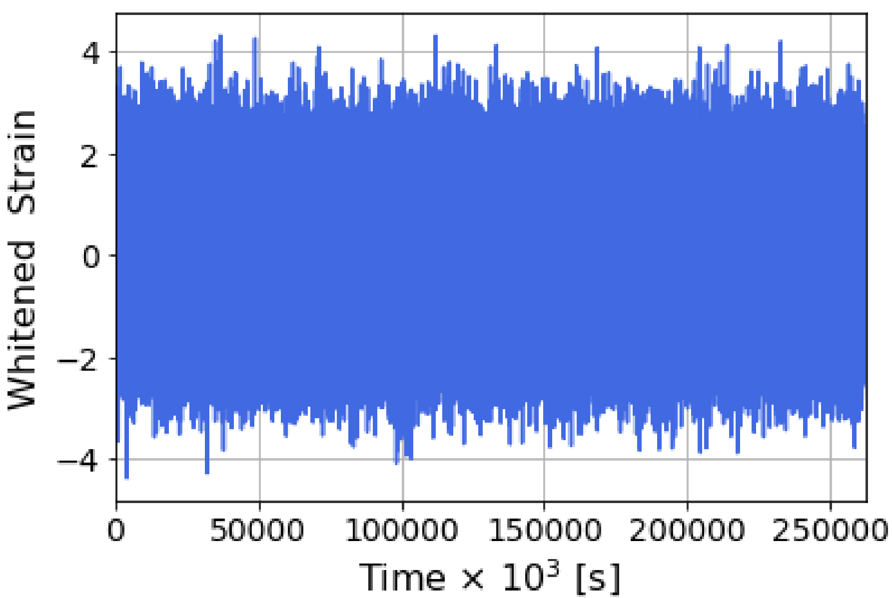
[1] L. Barack and C. Cutler, Phys. Rev. D 69, 082005 (2004) [2] Curt Cutler, Phys. Rev. D57 (1998) 7089-7102

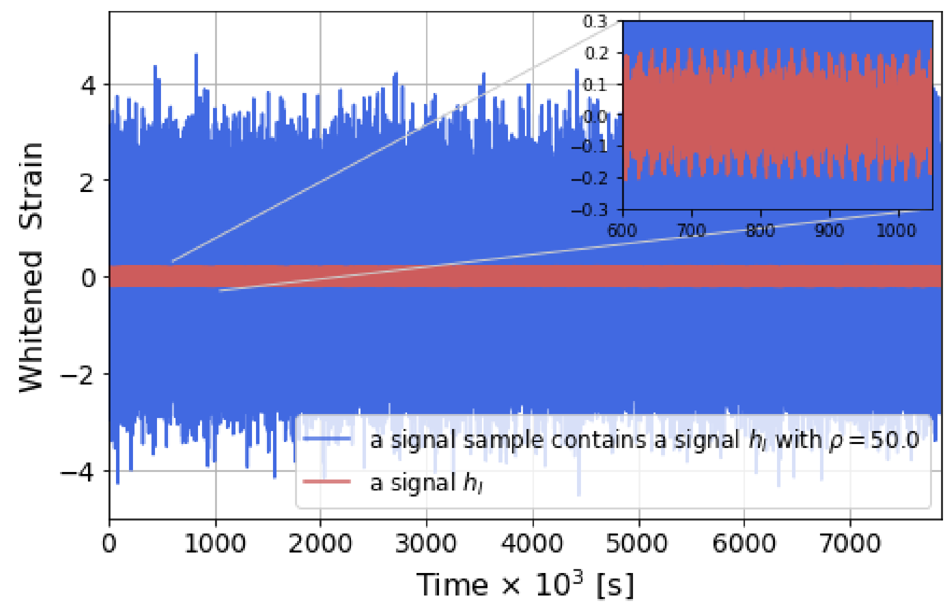
[3] K A Arnaud et al 2007 CQG 24 S551 [4]Cornish N J and Rubbo L J 2003 PRD 67 022001

Input samples

☐ noise-only sample and signal-plus-noise sample

- **Duration: 7864320 seconds**
- **Sample rate: 1/30 Hz**
- **A input sample shape: (2, 262144)**



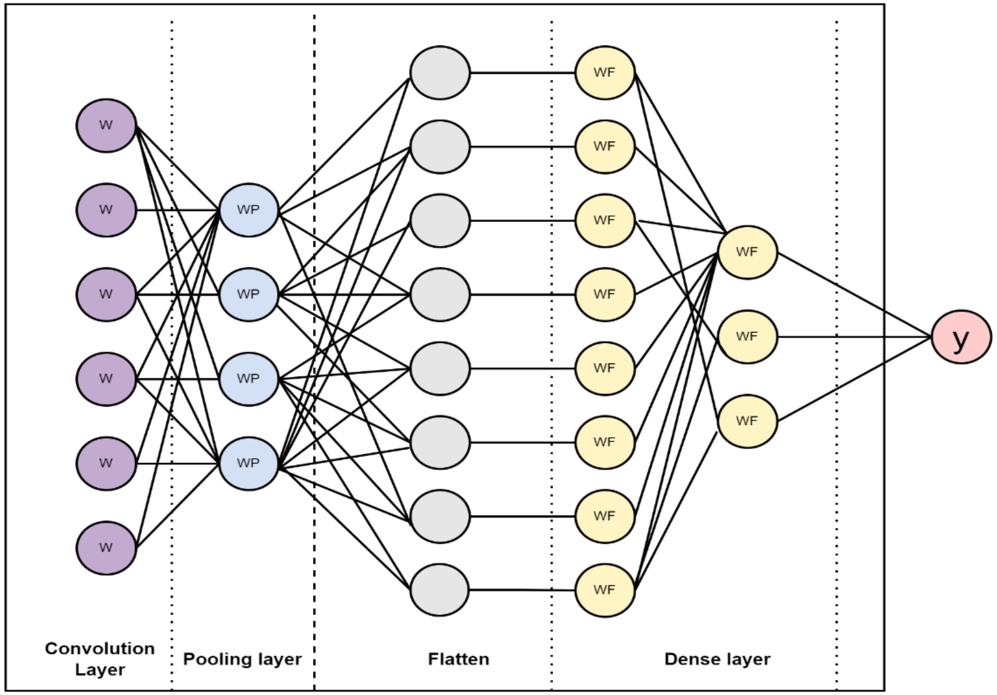


Detection method

convolutional neural network (CNN)

- a highly nonlinear function that maps the input space of the data to the output space: $y = f_w(d)$
- Binary classifier:

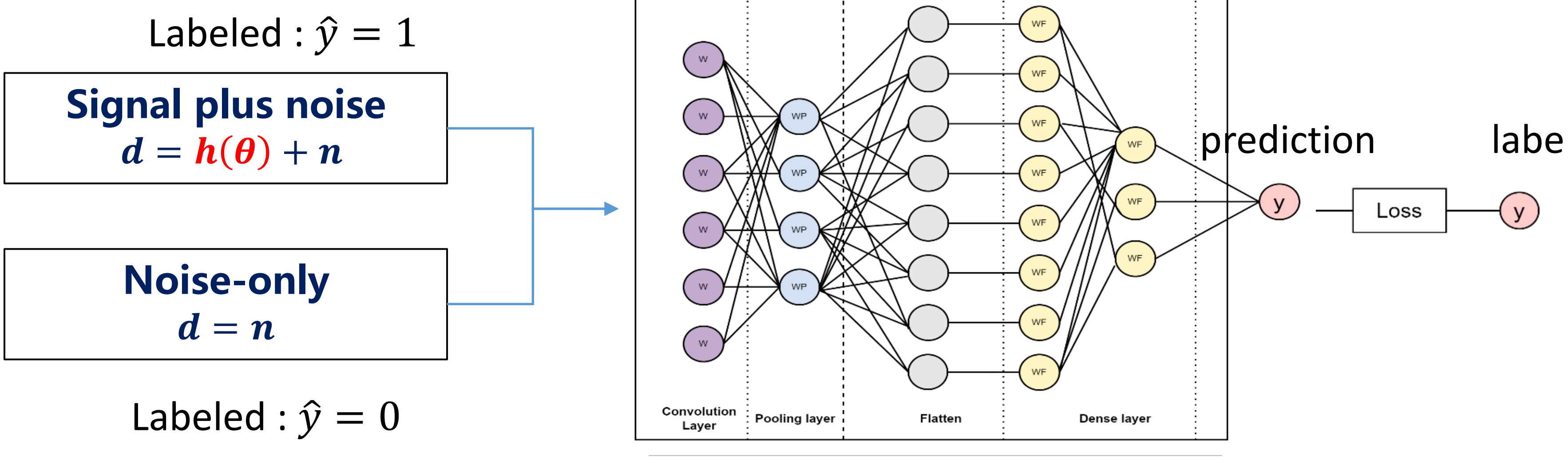
$$y = P(H_1|d) = CNN(d), \qquad y^0 = P(H_0|d) = 1 - y$$



Detection method

□ Training phase

- Training data contains signals with SNR U[50, 120] by rescaling the D_L
- Loss function: $\text{Loss} = \frac{1}{N}\sum_i -[\hat{y}_i \, log(y_i(Wd)) + (1-\hat{y}_i) log(1-y_i(Wd)].$



- Ntrain=500 000, Nval= 50 000, Nepoch=300, Nbatch=56
- Trained time : 10.5 days (GPU).

Detection method

final CNN architecture

- The number of trained parameters of CNN: 2 803 618

	Layers	kernel number	kernel size	Activation function
1	Input	...	matrix(size: 2×262144)	...
2	Convolution	32	matrix(size:1×34)	relu
3	Pooling	16	matrix(size:1×8)	relu
4	Convolution	16	matrix(size: 1×8)	relu
5	Pooling	16	matrix(size: 1×6)	relu
6	Convolution	16	matrix(size: 1×6)	relu
7	Pooling	16	matrix(size: 1×4)	relu
8	Flatten	...	...	...
9	Dense	...	vector(size: 128)	relu
10	Dense	...	vector(size: 32)	relu
11	Output	...	vector(size: 2)	softmax

Detection method

Testing phase

- **7 groups signal setups $h(\theta)$ used as testing data**

number	waveform model	physical parameters distribution		signal samples number
1	AK	$\rho \in$ uniform [50,120]		500
2	AK	$\rho > 50$, astrophysical model M12		500
3	AAK	$\rho \in$ uniform [50,120]		500
4	AK	ρ enumerates $10, 20, ..., 130$		1000×13
5	AK	M enumerates $10^4, 10^{4.5}, ..., 10^7 M_\odot$, $a = 0.98$	$z = 0.1$ $z = 0.2$ $z = 0.3$	$1000 \times 7 \times 3$
6	AK	$M = 10^6 M_\odot$, a enumerates $0.0, 0.2, 0.4, 0.6, 0.8, 0.98$	$z = 0.1$ $z = 0.2$ $z = 0.3$	$1000 \times 6 \times 3$
7	AK	$M = 10^{5.5} M_\odot, a = 0.98$, z enumerates $0.1, 0.2, 0.3$ $M = 10^6 M_\odot, a = 0.0$, z enumerates $0.1, 0.2, 0.3$ $M = 10^6 M_\odot, a = 0.98$, z enumerates $0.1, 0.2, 0.3$		$1000 \times 3 \times 3$

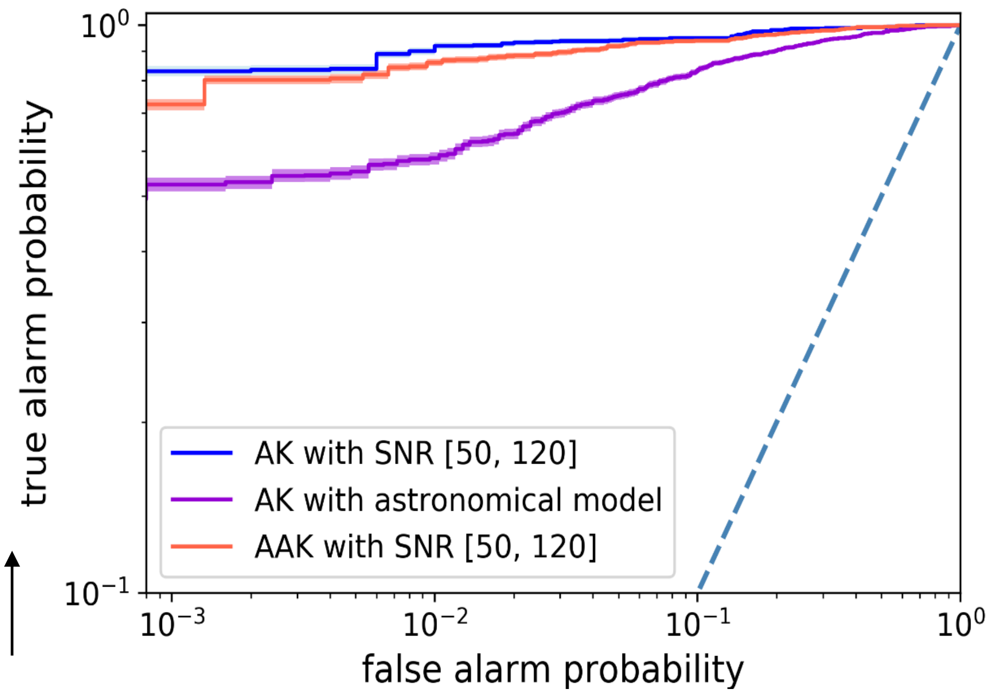
Results

☐ receiver operator characteristics (ROC) curve： group 1-3

- Blue: expected effectiveness, identical distribution to the training data
- Red: waveform model from AK to AAK waveform.
- Purple: parameters distribution is drawn from an astrophysical model.

$y > y^*$		prediction	
		Signal	noise
actual	signal	TP	FN
	noise	FP	TN

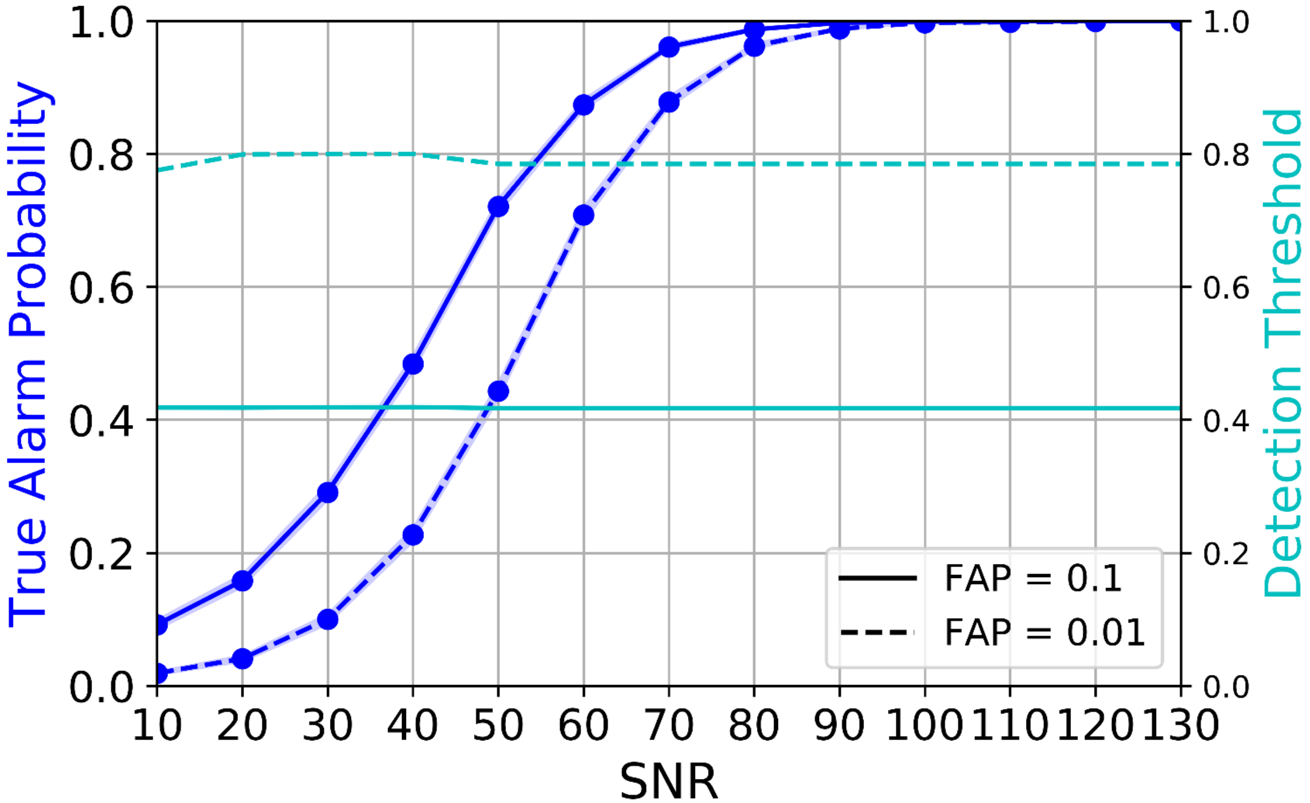
$$FAP = \frac{FP}{TN + FP}, \downarrow \qquad TAP = \frac{TP}{TP + FN} \uparrow$$



Results

□ Efficiency Curve: group 4

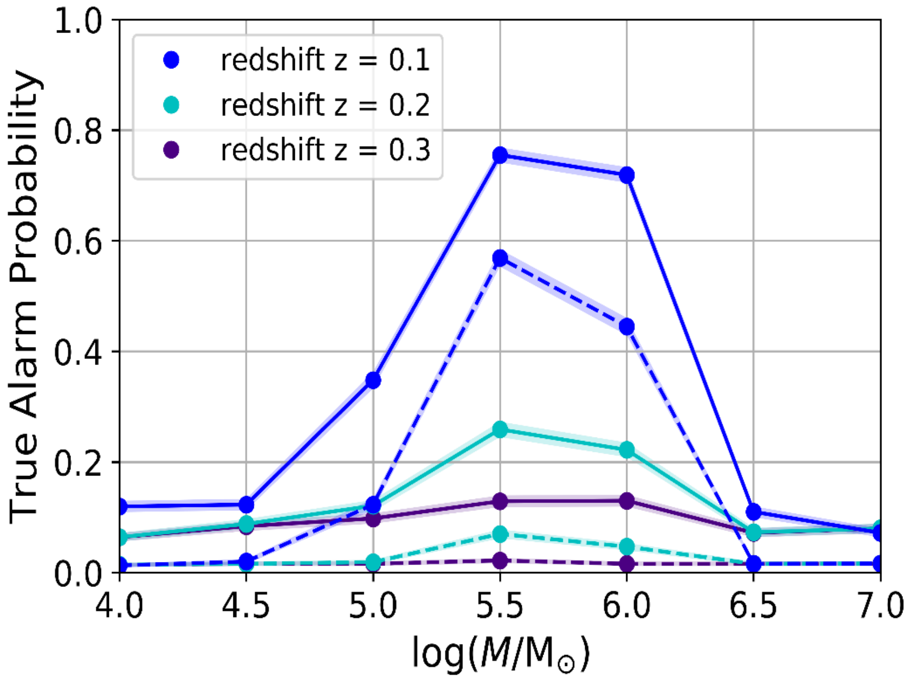
- consistent with the expectation that the CNN exhibits higher sensitivity toward stronger signal, and for SNR of higher than about 100

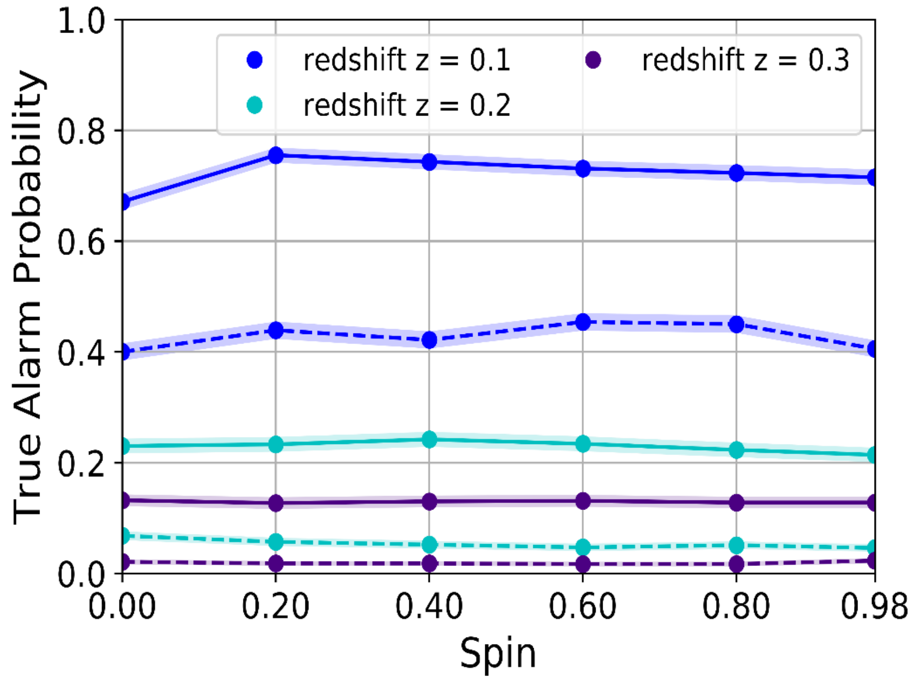


Results

□ Efficiency Curve：group 5-6

- Changes in other parameters can also lead to a different performance in TAP, but such differences can be mostly explained by the different SNRs



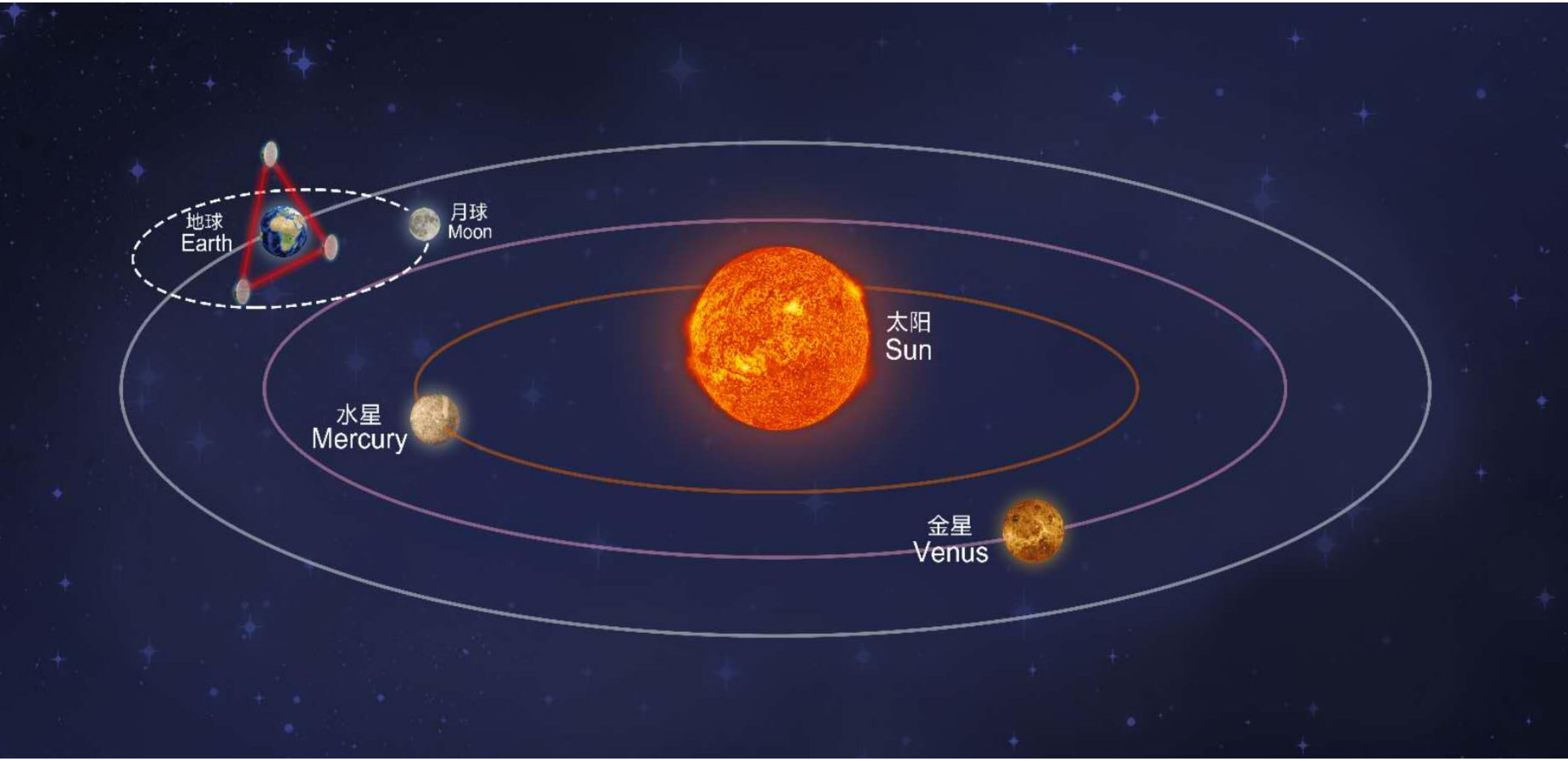


Conclusion

- We demonstrate a proof-of-principle application of a CNN on the EMRIs signals detections, covering a wide range of astrophysical parameters and giving FAP and TAP analysis.

- CNN shows a good generalization ability against a change of waveform models. [AK &AAK test]

- We recognize that there are still lots of challenges to implement a reliable CNN to detect EMRI signals. For example, one needs to push the SNR threshold to the values lower than 50.



谢谢！