Phys. Lett. B 858 (2024) 139016

Contents lists available at ScienceDirect

Physics Letters B

journal homepage: www.elsevier.com/locate/physletb

Letter

Gravitational wave signal extraction against non-stationary instrumental noises with deep neural network

Yuxiang Xu [a,b,c,d], Minghui Du [b], Peng Xu [b,a,e,d,*], Bo Liang [a,b,c,d], He Wang [f,g,d]

[a] Hangzhou Institute for Advanced Study, UCAS, Hangzhou 310024, China
[b] Center for Gravitational Wave Experiment, National Microgravity Laboratory, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China
[c] Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai 201800, China
[d] Taiji Laboratory for Gravitational Wave Universe (Beijing/Hangzhou), University of Chinese Academy of Sciences (UCAS), Beijing 100049, China
[e] Lanzhou Center of Theoretical Physics, Lanzhou University, Lanzhou 730000, China
[f] CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China
[g] International Centre for Theoretical Physics Asia-Pacific, University of Chinese Academy of Sciences, 100190 Beijing, China

ARTICLE INFO

Editor: M. Doser

ABSTRACT

Sapce-borne gravitational wave antennas, such as LISA and LISA-like mission (Taiji and Tianqin), will offer novel perspectives for exploring our Universe while introduce new challenges, especially in data analysis. Aside from the known challenges like high parameter space dimension, superposition of large number of signals etc., gravitational wave detections in space would be more seriously affected by anomalies or non-stationarities in the science measurements. Considering the three types of foreseeable non-stationarities including data gaps, transients (glitches), and time-varying noise auto-correlations, which may come from routine maintenance or unexpected disturbances during science operations, we developed a deep learning model for accurate signal extractions confronted with such anomalous scenarios. Our model exhibits the same performance as the current state-of-the-art models do for the ideal and anomaly free scenario, while shows remarkable adaptability in extractions of coalescing massive black hole binary signal against all three types of non-stationarities and even their mixtures. This also provide new explorations into the robustness studies of deep learning models for data processing in space-borne gravitational wave missions.

1. Introduction

Since the first landmark event GW150914 captured by Adv-LIGO [1], today nearly a hundred Gravitational Wave (GW) signals from mergers of stellar mass compact binaries had been observed by the LIGO-Virgo collaboration, which lead to numerous accomplishments in astrophysics and fundamental physics [2–14]. However, limited by seismic noises, ground-based GW detectors are typically sensitive to signals at frequencies higher than 10 Hz [15–17]. To enclose the exciting sources of larger and heavier astrophysical systems, one needs to explore the low frequency band of the GW spectrum with detectors of much longer baselines. The first generation space-borne antennas, including LISA (Laser Interferometer Space Antenna) [18,19] and the LISA-like missions Taiji [20,21] and Tianqin [22], would be launched in the 2030s and cover the mHz band. Offering novel perspectives for exploring the Universe, such space-borne antennas will also introduce new challenges [19], especially in data analysis [23–25].

Space antennas are supposed to response to a large amount of superimposed GW signals, including coalescing (super) Massive Black Hole Binaries (MBHB), extreme mass ratio inspirals, galactic compact binaries, evolving topological defects, primordial GW background, and also un-modeled sources. To achieve the expected scientific objectives, throughout analysis of the observational data and estimations of the relevant parameters from the largely superimposed GW signals against complicated noises are needed. Many works have been done to address these issues [26–28], and in present days the most accurate results can be obtained by matched filtering method for signals buried in Gaussian noises [29,30]. However, the construction of a full-fledged pipeline of scientific data analysis for space-borne antennas remains still an unfinished task, especially with the boost, in recent years, from machine

* Corresponding author at: Center for Gravitational Wave Experiment, National Microgravity Laboratory, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China.
E-mail address: xupeng@imech.ac.cn (P. Xu).

https://doi.org/10.1016/j.physletb.2024.139016
Received 21 February 2024; Received in revised form 7 August 2024; Accepted 9 September 2024
Available online 12 September 2024
0370-2693/© 2024 The Author(s). Published by Elsevier B.V. Funded by SCOAP³. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

learning methods been included. Following the studies in [31,32], accurate signal waveform extractions can be viewed as an intermediate step of the subsequent high-precision parameter estimations with machine learning methods, and the improvements in its performances and robustness are worth further explorations and investigations.

As lessons learned from the LIGO-Virgo observatories [33] and precedent missions like GRACE/GFO [34], LISA PathFinder (LPF) [35] and also Taiji-1 [36], new issues in data analysis for space-borne GW detections begin to draw more attentions [25,37–40]. As mentioned, to precisely measure the related parameters and infer the physical properties of the sources, continues measurements without disruptions in the data are normally demanded, which however imposes hard challenges on the long-term stability and robustness of both the payloads and satellite platforms. GW detection in space would be more seriously affected by anomalies or non-stationarities in the science measurements, including instrument transients or noise bursts, data gaps, slowly varying noise auto-correlations, trends etc., as the detailed knowledge of instrumental noises is crucial to GW signals extraction and parameter estimations. According to the designs of the LISA and Taiji missions, the foreseeable anomalies or non-stationarities in science measurements may come from the routine maintenance of the satellites or the onboard payloads and unexpected environment disturbances or instrumental anomalies, which will be the main subjects of this work. Cyclostationary properties also exist mainly due to the orbital modulations of the stochastic GW foreground from compact binaries, which are not within the scope of this work and will be left for future studies.

Routine maintenance, such as antenna re-orientations, payloads calibrations etc., could cause important and even continuous disturbances of satellite platforms and affect heavily the performances of the key payloads, therefore result into varying noise auto-correlations and even data loss or gaps. For example, the scheduled re-orientations of the transmission antennas would lead to data gaps about 3.5 hours per week or 7 hours every two-weeks [23]. In addition, unforeseen anomalies or failures due to hardware problems would take place and produce random unscheduled data gaps. Based on the experiences from LPF [24,38,40,41], GRACE/GFO [42,43] and also Taiji-1 [44], instrumental transients may also contaminate the science data, such as the acceleration glitches in gravitational reference sensor (GRS) systems, phase jumps in interferometers etc. Early studies on data processing algorithms against such possible anomalies can be found in [37,39,40]. Recently, Dey et al. [23] have used Bayesian analysis techniques to assess the impact of data gaps and concluded that the effects of unscheduled gaps were often more significant, and Spadaro et al. [45] used a joint parameter estimation approach to evaluate the influences of GRS glitches demonstrating that it is possible to accurately identify glitches in the absence of overlapping GW signals. These studies treated the above non-stationarities independently, and, for instrumental transients, prior knowledge of transient models are required. In this work, given the potentials and capacities of deep neural networks, we seek to develop an anomaly-model-independent signal extraction method for space GW antennas that can thoroughly treat and overcome the impacts of the above three types of data non-stationarities or anomalies.

Deep learning methods have already achieved considerable success in data processing of GW detections [31,32,46–55]. Compared to traditional algorithms, such as matched filtering [29,30,56], which generally require a vast template bank and thus consume substantial computational resources, deep learning methods are well known for their fast processing speed and excellent generalization ability, making it possible to achieve rapid and anomaly-model-independent signal extraction. For GW detection in space, Conv-TasNet [31] has demonstrated its performance in various scenarios, completing signal extraction in $\lesssim 10^{-2}$ seconds, however, robustness tests under anomalous conditions are overlooked. In this work, focusing on the three types of possible non-stationarities for LISA or LISA-like missions, including gaps, transients (glitches), and time varying noise auto-correlations, we developed a deep learning model for accurate extractions of coalescing MBHB signals confronted with such anomalous scenarios. This also provide new explorations into the robustness studies of deep learning models for data processing in space-borne GW missions. Our model exhibits the same performance as the Conv-TasNet model [31] does for the ideal and anomaly free scenario, while shows remarkable adaptability in signal extractions against all three types of non-stationarities and even their mixtures.

2. Dense-LSTM model

Autoencoders, commonly employed as denoising architectures, are extensively applied in fields like image denoising [57–60], signal extractions [61–64], and speech enhancements, et al. [65–68]. Autoencoder models include an encoder and a decoder. The encoder is designed to extract feature mappings from the input data, while the decoder is to generate reconstructed data from these mappings. For GW signal extractions in LIGO-Virgo data analysis, the CNN-LSTM denoising autoencoder model has been used [69] and shown superior performance.

As discussed in Sec. 1, the long periods and consequently the much more complicated data anomalies for science measurements of LISA and LISA-like missions have driven us to develop models capable of uncovering deeper mappings within the signals and thereby reducing the loss and alteration of the crucial information under significant anomalous conditions. Simple CNN architectures fall short in meeting these demands. In this work, we choose to use a deeper CNN network different from the CNN-LSTM denoising autoencoder suggested by Chatterjee et al. [69]. The DenseNet is chosen as the encoder for its advantage in reducing computational overhead while maintaining comparable accuracy compared with other deep CNN networks [70–73] and at the same time ensuring faster feature extraction. More importantly, DenseNet thoroughly implements the concept of feature reuse, with each layer directly connected to all preceding layers. This enables the comprehensive utilization of low-complexity shallow features, facilitating the derivation of a smoother and more robust decision function. Such a design ensures considerable extraction capabilities even when confronted with data strongly contaminated by the possible non-stationarities or anomalies for space-borne GW detections. For the decoder, we retain the bidirectional LSTM network due to its superiority in capturing long-term dependencies within sequences.

The structure of our model is depicted in Fig. 1. The input GW data for the model have a sampling frequency of 0.1 Hz and a duration of 160,000 seconds, amounting to 16,000 data points. The detailed data preparations can be found in Sec 3. All data are normalized between -1 and 1, followed by segmentation into 16,000 overlapping subsequences of length 4. Each subsequence is then fed into the decoder network for feature extraction. The feature vector of each subsequence is processed through three bidirectional LSTM layers for predicting the output of the subsequent time step. Finally, a complete waveform output is obtained through a dense layer.

3. Data preparations

For LISA and LISA-like missions, laser frequency instability noise is dominant in the science measurements, and a pre-processing technology called Time Delay Interferometry (TDI) [74–76] is employed to efficiently suppress such noise. Following this, our data is generated through the first generation noise orthogonal TDI channels A and E, see [74] for a detailed introduction. The total TDI output data stream of channel I (dubbed $s^I(t)$) is the combination of signal $h^I(t)$ and noise $n^I(t)$:

$$s^I(t) = h^I(t) + n^I(t), \quad I \in \{A, E\}, \tag{1}$$

where the detector response in terms of channel I to some incident GW with polarizations h_α reads

$$h^I(t) = \sum_\alpha T^I_\alpha(t)\, h_\alpha(t), \quad \alpha \in \{+, \times\}, \tag{2}$$

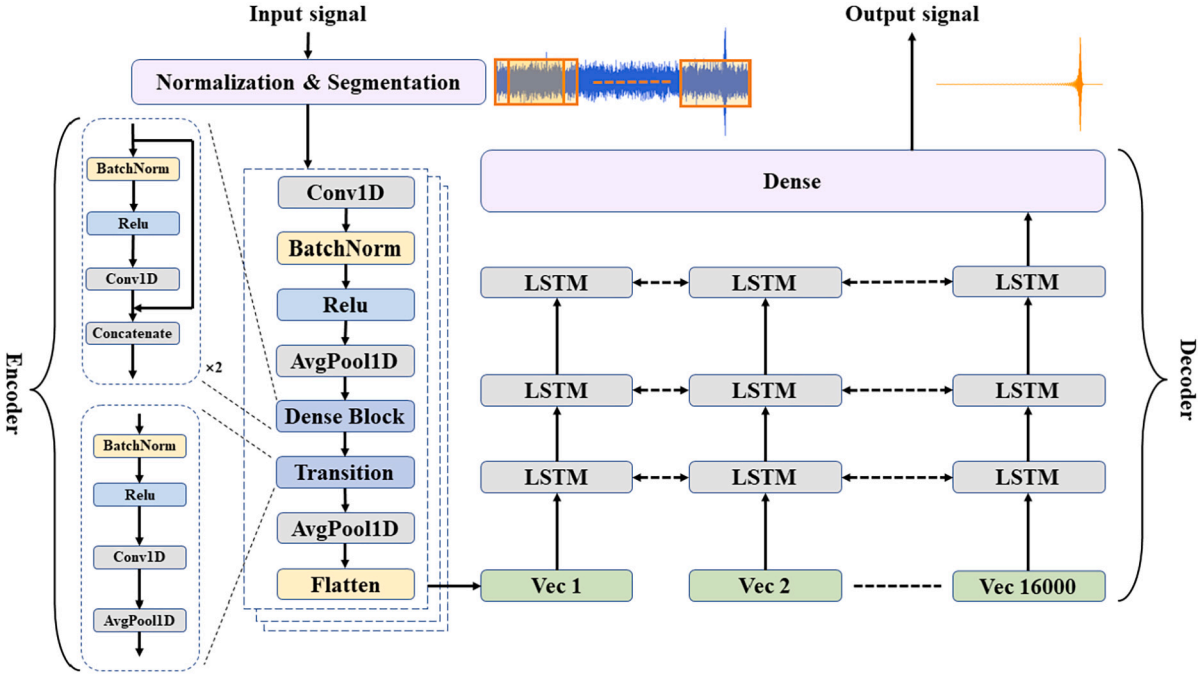


Fig. 1. The architecture of our denoising autoencoder is delineated as follows. The input data is first subjected to normalization and segmentation (top left), resulting in the formation of overlapping subsequences. Each subsequence is then processed through the encoder to extract its characteristic feature vector. This is followed by passage through three bidirectional LSTM layers to yield predictive values. The final reconstructed waveform is then attained in the Dense layer (top right).

where $T_\alpha^I(t)$ is the total transfer function of signal including antenna responses and TDI combinations, and $h_\alpha(t)$ represents the polarization components of the coalescing MBHB GW signal. Specific details of this transfer function can be found in [74,76]. In this paper, the TDI transfer functions of signals are derived according to Taiji's mission concept and orbit configurations [77], which has a nominal arm length of $\sim 3\times10^9$ m. The PyCBC package [78] and the waveform model SEOBNRv4 [79] are employed to generate the coalescing MBHB GW templates $h_\alpha(t)$. On the other hand, the noise floor of the output data from a certain TDI channel is in principle determined by the corresponding TDI combinations of GRS residual acceleration noises $n_{\rm ACC}(t)$ and optical metrology system noises $n_{\rm OMS}(t)$. The resulting total instrumental noise of TDI channel I can be written as

$$n^I(t) = T^I_{\rm OMS}(t) n_{\rm OMS}(t) + T^I_{\rm ACC}(t) n_{\rm ACC}(t), \tag{3}$$

with $T^I_{\rm ACC}(t)$ and $T^I_{\rm OMS}(t)$ being the transfer functions of GRS residual acceleration noise and optical metrology system noise, respectively [74,76], which are common to LISA and Taiji missions. The time series of basic instrumental noises $n_{\rm ACC}(t)$ and $n_{\rm OMS}(t)$ are generated according to the Power Spectral Densities (PSDs) of these two noise components, denoted as $S^{1/2}_{\rm OMS}(f)$ and $S^{1/2}_{\rm ACC}(f)$, and according to the designs of Taiji [77] we assume the nominal values

$$S^{1/2}_{\rm OMS}(f) = 8\times10^{-12}\sqrt{1+\left(\frac{2\ \rm mHz}{f}\right)^4}\ \frac{\rm m}{\sqrt{\rm Hz}}, \tag{4}$$

$$S^{1/2}_{\rm ACC}(f) = 3\times10^{-15}\sqrt{1+\left(\frac{0.4\ \rm mHz}{f}\right)^2} \times\sqrt{1+\left(\frac{f}{8\ \rm mHz}\right)^4}\ \frac{\rm m/s^2}{\sqrt{\rm Hz}}. \tag{5}$$

Specifically, to generate the noise data in time domain according to the PSDs, we convert Gaussian white noises to the frequency domain through Fourier transform, then multiply them by the amplitude spectral density value at each frequency to adjust the spectra of the noises, and finally obtain time series of noises through inverse Fourier transform.

Table 1
Summary of parameter setups in coalescing MBHB GW signals generations.

Parameter	Lower bound	Upper bound
$M_{\rm tot}$ [a]	$10^6 M_\odot$	$10^8 M_\odot$
q [b]	0.01	1
s_1^z [c]	-0.99	0.99
s_2^z [d]	-0.99	0.99

[a] Total mass of the system.
[b] Mass ratio of the objects in the binary system.
[c] Dimensionless spin parameter for the primary object.
[d] Dimensionless spin parameter for the secondary object.

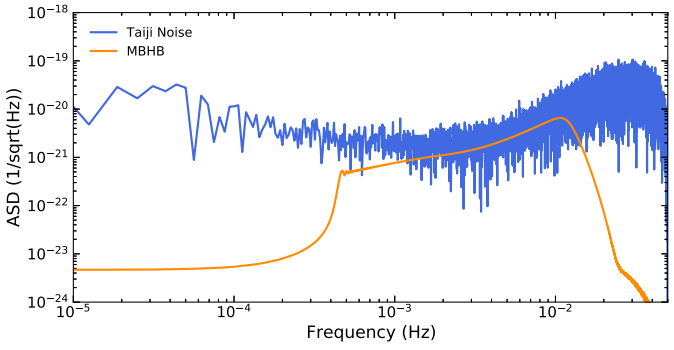


Fig. 2. The noise curve and the typical coalescing MBHB signal (SNR$\sim$ 35.0) in the frequency domain.

Theoretically, the total noise PSDs of the A, E channels $S_n^I(f)$, $I\in\{A,E\}$ are the TDI combinations of $S_{\rm OMS}(f)$ and $S_{\rm ACC}(f)$ calculated according to Eq. (3), see Refs. [74,76,80,81] for their specific forms.

We established the parameter space for coalescing MBHBs as in Table 1, and employed a uniform grid to generate the diverse waveforms. Parameters such as times and phases at coalescence, sky locations, polarization angles, as well as inclinations are drawn from uniform distributions. We inject the signal with specific optimal Signal-to-Noise Ratios (SNR)

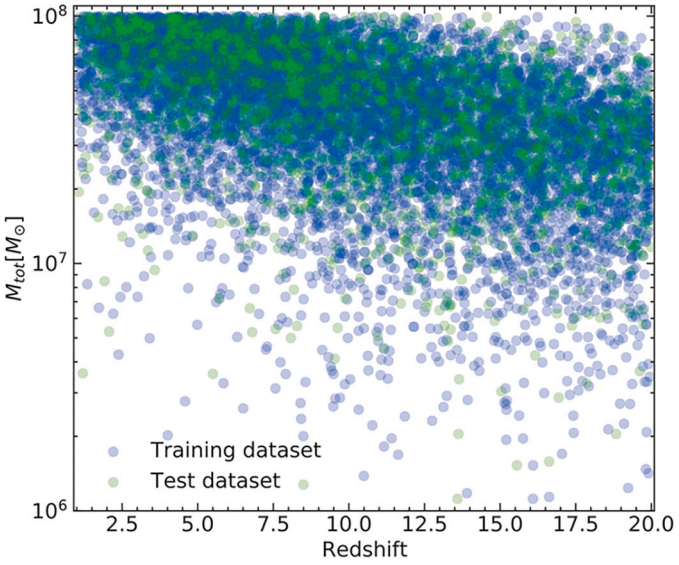


Fig. 3. The figure illustrates the distribution of our training and test datasets across redshifts ranging from 1 to 20.

$$\mathrm{SNR} = (h^I|h^I)^{1/2}, \tag{6}$$

where h^I represents the signal template through the TDI channel I, the inner product $(s^I|h^I)$ is defined as,

$$(s^I|h^I) = 2\int_{f_{\min}}^{f_{\max}} \left(\tilde{s}^I(f)\tilde{h}^{I*}(f) + \tilde{s}^{I*}(f)\tilde{h}^I(f)\right) df, \tag{7}$$

here, $f_{\min} = 3\times10^{-5}$ Hz and $f_{\max} = 0.05$ Hz. The tilde symbol represents the Fourier transform and $*$ the complex conjugation. This inner product can also be used to calculate the overlap between the signal waveform o extracted by the model and the whitened template waveform h through channel A or E can be evaluated in terms of the overlap function

$$\mathcal{O}(o,h) = \max_{t_c,\phi_c}(\hat{o}|\hat{h}), \tag{8}$$

where $\hat{o}$ and $\hat{h}$ are the normalized ones, and t_c and ϕ_c are the instantaneous time and phase corresponding to the maximum overlap between $\hat{o}$ and $\hat{h}$.

We set the SNR level for the training dataset as 50, and generate 10,000 samples. The distances of the sources are adjusted according to the specific optimal SNR. All samples are subjected to whitening processing, and a Tukey window with $\alpha = \frac{1}{8}$ is adopted.

Following the same routine, we produce 2,000 test samples, see Fig. 2 for illustration. In order to assess the model's performance for signals extractions with relative low SNRs and test the robustness of our model against various data anomalies, these samples were generated with SNRs ranging from 30 to 70. Our training and test datasets were sampled from redshifts ranging from 1 to 20. The resulting distribution is shown in Fig. 3.

4. Training strategy

For denoising autoencoders, the Mean Squared Error (MSE) is a prevalent loss function employed to quantify the discrepancy between the network's predictions and the true values. However, given that the GW signals fed into the network have undergone normalization, the contributions to the overall MSE from the earlier stages of the coalescing signals are rather weak compared to those from the merger phases. This can hinder the model from achieving optimal refinements. To address this, the loss function defined in [69] is adopted which is the difference of the MSE and the fractal Tanimoto similarity coefficient [82] terms

$$L_{o,h} = \frac{\sum_i^n (h_i - o_i)^2}{n} - r^d_{w,o,h}, \quad (i = 1,\ldots,n) \tag{9}$$

the first term on the right hand side is the MSE term and the fractal Tanimoto similarity coefficient is defined as

$$r^d_{w,o,h} = \frac{\sum_i^n w_i \cdot o_i \cdot h_i}{2^d \sum_i^n w_i \cdot (o_i^2 + h_i^2) - (2^{d+1}-1)\sum_i^n w_i \cdot o_i \cdot h_i}. \tag{10}$$

The MSE is used for the optimizations of individual data points, while the fractal Tanimoto similarity coefficient is used for the optimizations of the overall data. Tanimoto coefficient is commonly used to measure the similarity between two vectors, and the fractal coefficient, as an enhanced version, introduces a parameter d to facilitate deeper levels of optimizations. This approach ensures that the extracted data achieves good amplitude and phase agreements with the template data on a global level.

Another worth-noting modification is that, unlike in the original work [69] where the weight w_i is set to $1/o_i$, we set w_i equal to 1 in our approach. This is because, the long observation time in space borne GW detection will give rise to large number of data points that having amplitudes close to zero in the generated templates, the high weights, such as $\omega_i = 1/o_i$, of these data points with $o_i \sim 0$ will reduce the weights of the data from the merger phases, and therefore the optimizations will be highly affected by the data segments away from the important merge phases and consequently prevent the network from converging. For the parameter d, which is used to adjust the depth of optimization in similarity, is initially set to 0. As the training converges, d grows gradually to facilitate deeper levels of optimization.

To summarize, at the beginning of the training, we set the learning rate to 10^{-3} and d to 0. As the model approaches convergence, we set $d = 5$ and reduce the learning rate by a factor of 10. We have chosen the learning rate as 10^{-3}, 10^{-4}, and 10^{-5} with corresponding d taking values as 0, 5, and 10 for case studies, and trained for 200 epochs for each case.

5. Results

We first tested the extraction performance of our model for the ideal case without data non-stationarities or anomalies, see illustrations in Fig. 4. More than 99% of the signals extracted by the model achieve an overlap ≥ 0.9 with the corresponding template, and each signal could be extracted in less than 10^{-2} seconds. This demonstrates a similar performance of our model compared to the current state-of-the-art models [31]. In the followings, we conduct various tests of our model against the possible non-stationarities or anomalies known to the literature, including data gaps, time-varying noise auto-correlations, glitches, and even their mixtures. We expect that at least 80% of the signals extracted by the model should attain an overlap ≥ 0.9 with the template signals,

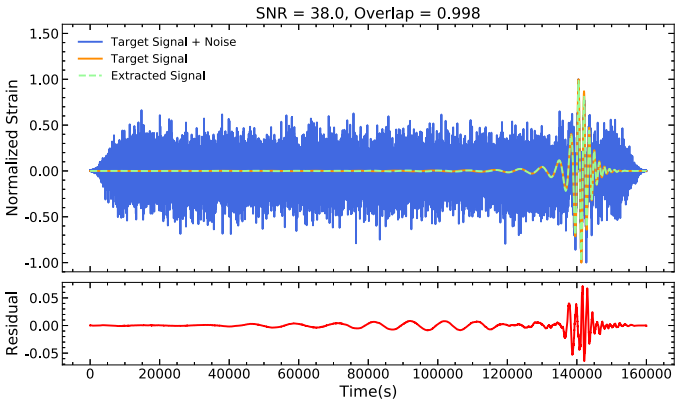


Fig. 4. Example of signal extraction with our model when there is no anomaly or non-stationarity present in the data.

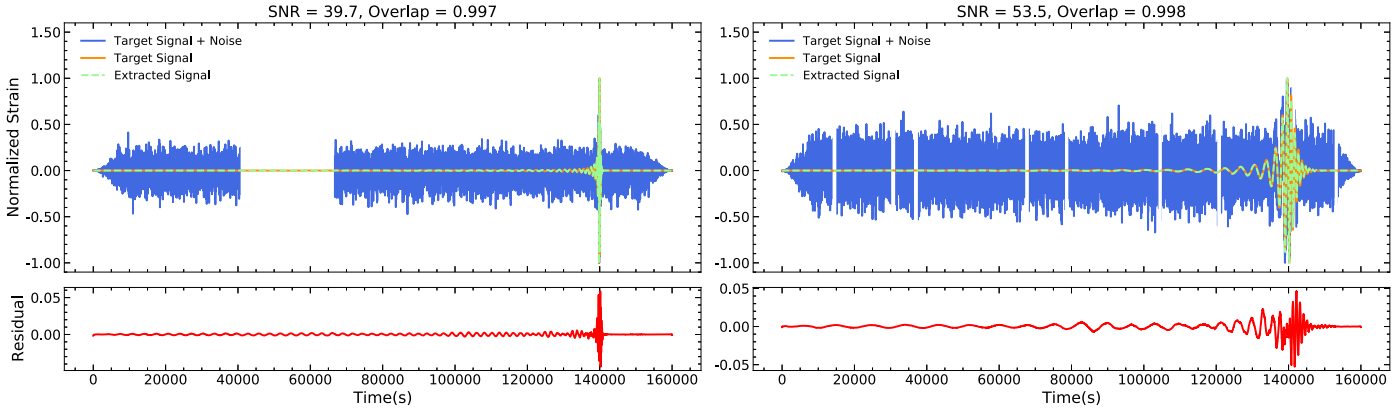


Fig. 5. The left image illustrates the model's extraction result in the presence of scheduled long gap, while the right image depicts the outcomes when facing unscheduled gaps, with a occurrence frequency about once every four hours.

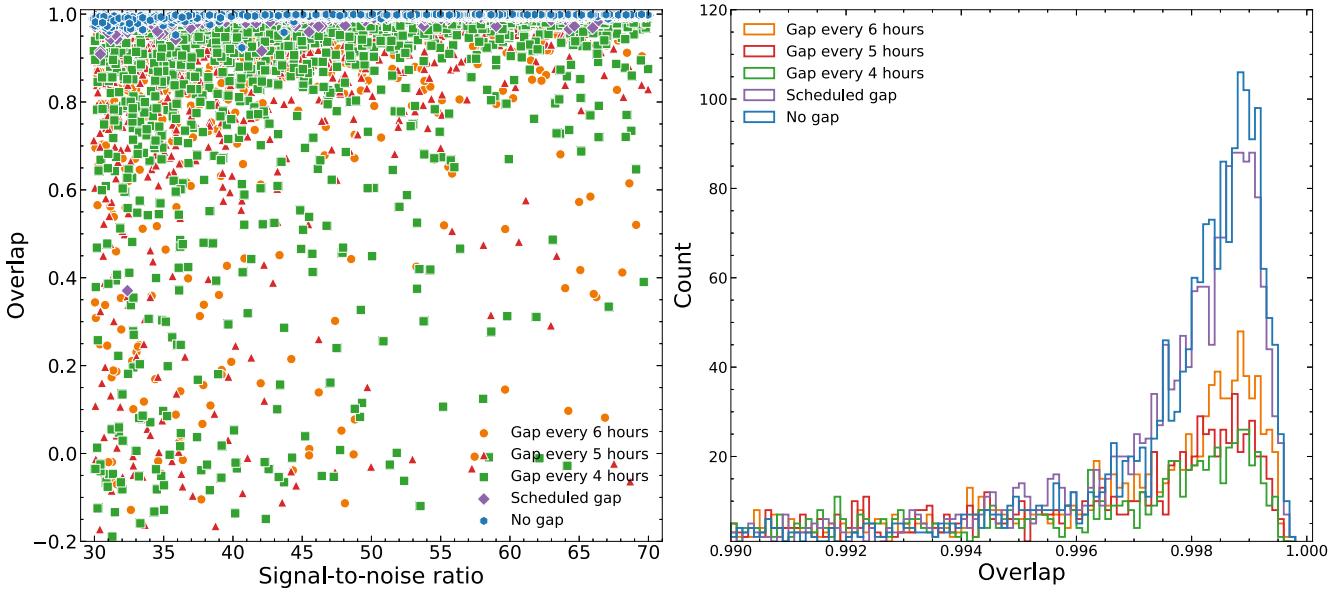


Fig. 6. The left image displays the distributions of the overlaps for our test set confronting with data gaps of different occurrence frequencies. The image on the right presents the histogram for instances where the model achieves an overlap ≥ 0.99, corresponding to our test set confronting with data gaps of different occurrence frequencies.

which we will use as the criterion for assessing the limits of the model's performance. The physical meanings of a more realistic criterion or threshold should come from more detailed studies on false positive, uncertainties in parameters estimations and so on, which will be left for future studies.

5.1. Data gaps

Following the discussions in [23], we use window functions to generate data gaps (see Ref [23] for the specific forms), which, as discussed in Sec. 1 can be classified into two types, that the scheduled prolonged gaps and unscheduled random but short gaps. For the scheduled gaps, we emulate an extreme situation, resulting into a data loss of 7 hours during the inspiral phases. For unscheduled gaps, it can last up to days or cause data loss of a few minutes per day due to glitch masking according to [37]. Considering the length of our data, here we choose the second type of unscheduled gap. We set each gap to cause 5 to 8 minutes of data loss, and progressively increase the frequency of occurrences of such randomly happened gaps to a threshold at which our model can hardly tolerate. Moreover, the unscheduled gaps are allowed to break out at any place of our data, while a prolonged data gap during the merger phase is ignored for obvious reason. Our extraction results are presented in Fig. 5. Our model can proficiently extract the signals despite facing these two types of data gaps. We delved further into the threshold performance of our model, as detailed in Fig. 6. The model exhibits minimal accuracy loss in scenarios with scheduled gaps before mergers. When confronted with the random unscheduled gaps, the model's results gradually worsen as the frequency of occurrences increases. As expected, signals with lower SNRs are more susceptible to data gaps.

Table 2
For different types of data gaps, the ratio of overlaps ≥ 0.9 between the model-extracted signals and the templates across the entire dataset. The last two columns represent the means and standard deviations of the overlap.

Type of gap	Overlaps ≥ 0.9	Mean	Std. Dev.
No gap	100.0%	0.997	0.005
Scheduled gap	99.9%	0.996	0.016
Gap every 6 hours	86.7%	0.935	0.169
Gap every 5 hours	83.1%	0.923	0.181
Gap every 4 hours	77.3%	0.893	0.220

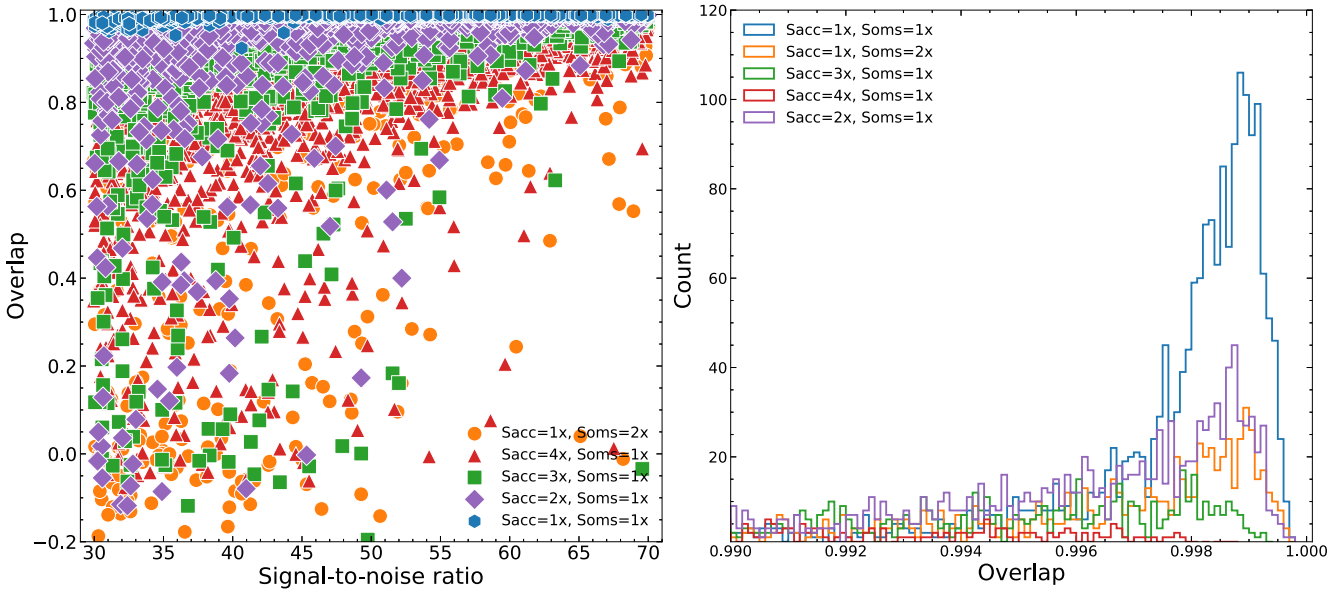


Fig. 7. The left image displays the distributions of the overlaps for our test set confronting with time-varying noise auto-correlations of varying intensities. The image on the right presents the histogram for instances where the model achieves an overlap ≥ 0.99, corresponding to our test set confronting with time-varying noise auto-correlations of varying intensities.

In Table 2, we present more detailed results under different conditions. According to the predetermined criterion, our model can maintain robustness even when the frequency of unscheduled gaps reaches once per 5-hours.

5.2. *Time-varying noise auto-correlations*

We simulated non-stationary noises with time-varying auto-correlations or PSDs by altering the magnitudes of $S_{\mathrm{OMS}}(f)$ and $S_{\mathrm{ACC}}(f)$ of the two noise components, with the modified noise segment lasting for 80,000 seconds. In Fig. 8, we present one of our extraction samples. The limits of our model's performance against such non-stationary noise is investigated and shown in Fig. 7. As expected, data with lower SNRs will be more significantly affected by time-varying noise PSDs, and the detailed results are presented in Table 3. Our model can successfully extract the signal without being affected by such variations in total noise PSDs.

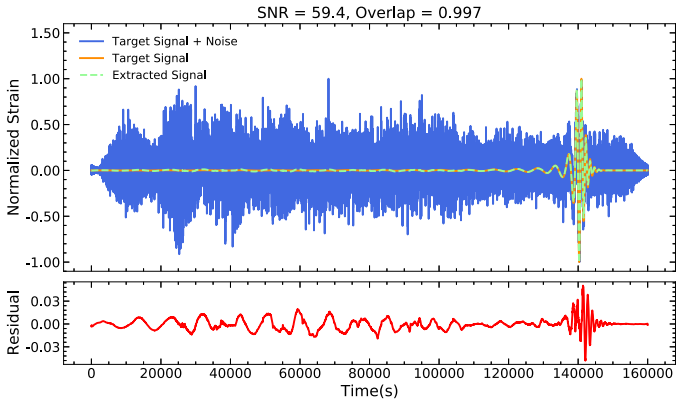


Fig. 8. Example of signal extraction of the model when contending with time-varying noise auto-correlations. Here, a segment of the data spanning 80,000 seconds has the residual acceleration noise set to threefold its nominal magnitude. The result demonstrate that our model effectively extracted the signal, achieving a high overlap.

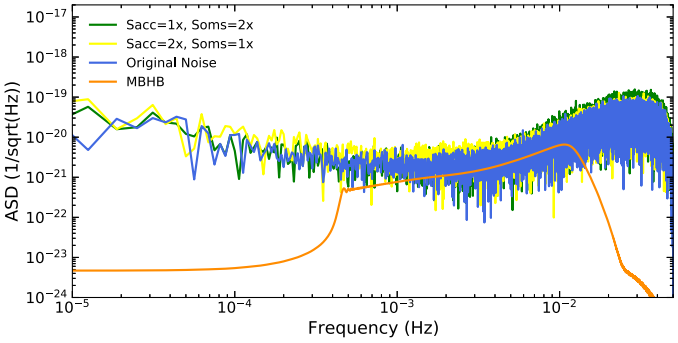


Fig. 9. The changes in the overall noise spectrum when the magnitudes of the optical metrology system noise and acceleration noise are altered. For our test data set, the optical noises affect the total SNR (SNR $\sim$ 35 in this example) more significantly.

While, given the predefined criterion, one finds that noises originating from the optical metrology system has a more significant impact on the final outcomes. Further investigation shows that this is because that, increasing the optical metrology system noise (concentrated in the relatively high frequency band) will significantly reduce the SNR of coalescing MBHB signals due to the specific forms of the PSDs of the total noise and expected signal, see Fig. 9 for illustrations. This, but not the

Table 3
For time-varying noise auto-correlations, the ratio of overlaps ≥ 0.9 between the model-extracted signals and the template signals across the entire dataset. The last two columns represent the means and standard deviations of the overlap.

$S_{\mathrm{OMS}}(f)$	$S_{\mathrm{ACC}}(f)$	Overlaps ≥ 0.9	Mean	Std. Dev.
1x	1x	100.0%	0.997	0.005
1x	2x	92.2%	0.962	0.117
1x	3x	81.5%	0.923	0.158
1x	4x	56.0%	0.842	0.205
2x	1x	78.3%	0.894	0.224

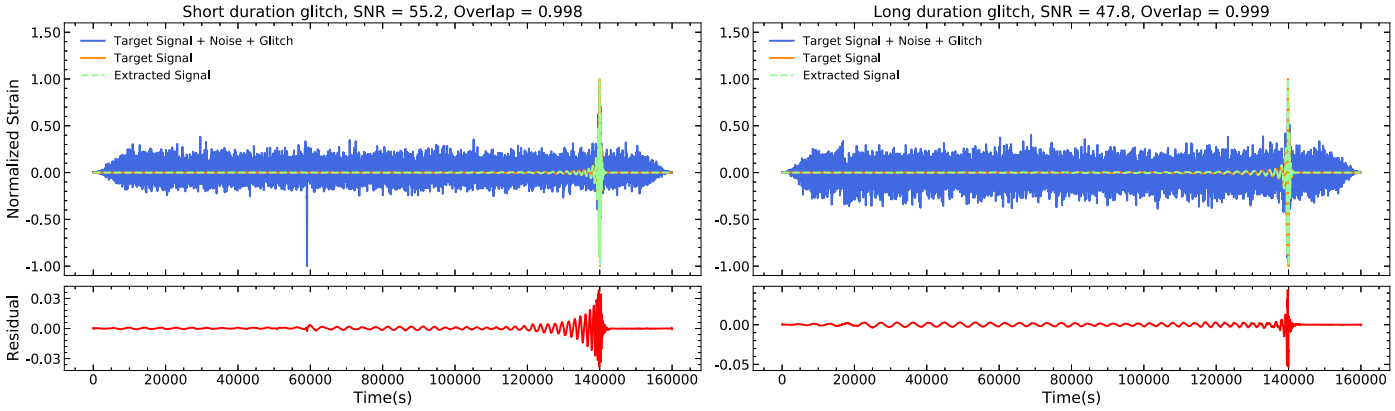


Fig. 10. The left image displays the extraction result of our model when encountering a short duration glitch with a peak value of $1 \times 10^{-13}\,\mathrm{ms}^{-2}$. The right image shows the extraction result of our model when encountering a long duration glitch with a peak value of $1 \times 10^{-14}\,\mathrm{ms}^{-2}$.

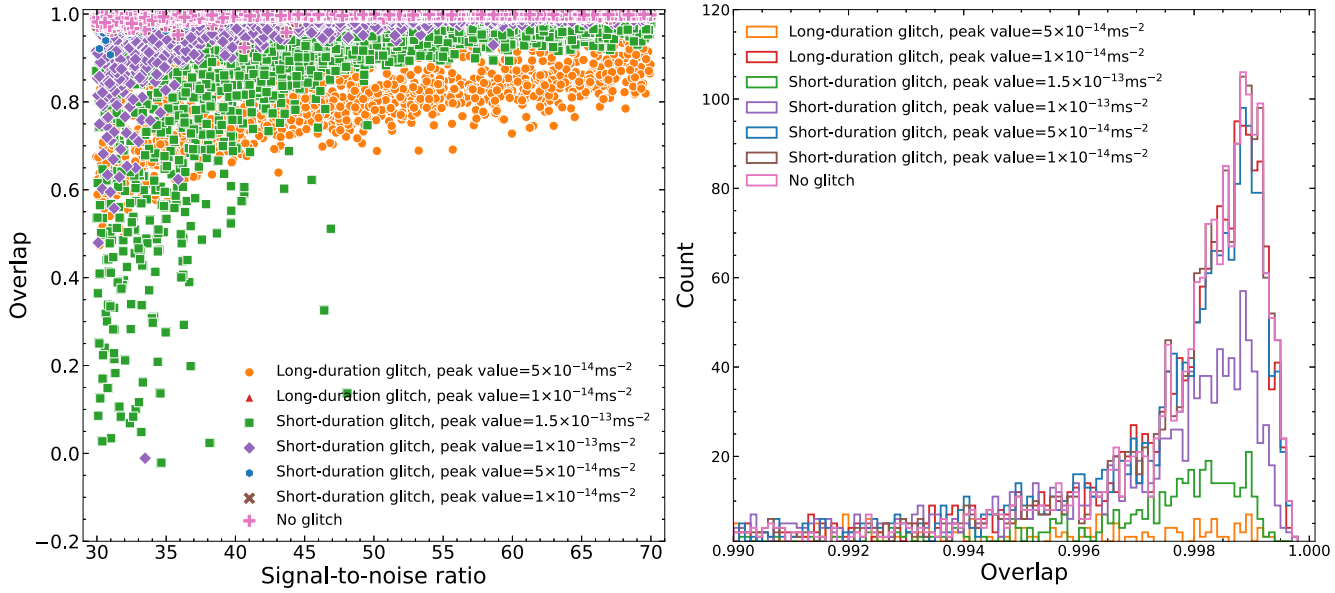


Fig. 11. The left image displays the distributions of the overlaps for our test set confronting with glitches of varying intensities and duration. The image on the right presents the histogram for instances where the model achieves an overlap ≥ 0.99, corresponding to our test set confronting with glitches of varying intensities and duration.

variations in noise PSDs, will make it more difficult for the model to extract signals and produce the results shown Fig. 7. and Table 3.

5.3. Glitches

For instrumental transients, we consider the legacy model from LISA Pathfinder for GRS glitches as representatives [40]

$$g(t) = \frac{\Delta v}{\tau_1 - \tau_2}\left(e^{\frac{-(t-t_0)}{\tau_1}} - e^{\frac{-(t-t_0)}{\tau_2}}\right)\Theta(t-t_0), \tag{11}$$

where t_0 is the time of glitch injection, Δv is the impulse transferred by the glitch (i.e., the gain of test mass velocity caused by the glitch), τ_1, τ_2 are time scales, and $\Theta(t-t_0)$ denotes the Heaviside step function. Glitches are simulated and injected to the GRS residual acceleration noise $n_{\mathrm{ACC}}(t)$ of TM_{12} (*i.e.* the test-mass onboard spacecraft 1 and facing spacecraft 2). They are then passed through $T^{I}_{\mathrm{ACC}}(t)$ (in Eq. (3)), together with the "normal" acceleration noises, to get the total TDI combinations with glitches. We considered two types of glitches according

Table 4
The parameters corresponding to the two types of glitches, under this parameter setting, their amplitudes are both $1 \times 10^{-14}\,\mathrm{ms}^{-2}$.

Duration	Δv (m/s)	τ_1 (s)	τ_2 (s)
Short	2.86×10^{-13}	10	11
Long	5.44×10^{-11}	2000	2001

Table 5
For glitches with different intensities and duration, the ratio of overlaps ≥ 0.9 between the model-extracted signals and the template signals across the entire dataset. The last two columns represent the means and standard deviations of the overlap.

Duration	Peak value(ms^{-2})	Overlaps ≥ 0.9	Mean	Std. Dev.
None	None	100.0%	0.997	0.005
Short	1×10^{-14}	100.0%	0.997	0.005
Short	5×10^{-14}	99.9%	0.995	0.010
Short	1×10^{-13}	94.9%	0.977	0.057
Short	1.5×10^{-13}	72.0%	0.893	0.161
Long	1×10^{-14}	100.0%	0.996	0.006
Long	5×10^{-14}	31%	0.830	0.112

Table 6
After mixing all anomalous or non-stationary conditions, the ratio of overlaps ≥ 0.9 between the model-extracted signals and the template signals across the entire dataset. The results also show the means and standard deviations of the obtained overlaps.

Data gaps		Non-stationary Gaussian noise		Glitches		Results		
Scheduled gap	Unscheduled gap	$S_{\mathrm{OMS}}(f)$	$S_{\mathrm{ACC}}(f)$	Duration	Peak value (ms^{-2})	Overlaps ≥ 0.9	Mean	Std. Dev.
✓	×	1x	3x	Short	5×10^{-14}	81.6%	0.919	0.167
✓	×	1x	3x	Long	1×10^{-14}	80.5%	0.921	0.160
×	Gap every 10 hours	1x	2x	Short	5×10^{-14}	82.1%	0.914	0.194
×	Gap every 10 hours	1x	2x	Long	1×10^{-14}	81.8%	0.905	0.216

to Ref. [40], that the short-duration glitches lasting approximately 70 seconds and long-duration glitches for about 3 hours. The initial amplitudes of both types of glitches were set to $1\times10^{-14}\,\mathrm{ms}^{-2}$, see Table 4 for detailed parameters, and we tested the robustness of our model by increasing the amplitudes by 0.5 times for each new experiment until the performance fell below the threshold we set. All glitches were injected with an initial time $t_0 = 0$, and the time difference between the injection points of the glitches and the mergers rang from 22 to 38 hours. Our model proficiently extracts the coalescing MBHB signals given such instrumental transients. Fig. 10 shows an example of the model's extraction result.

The limits of our model's performance against glitches is shown in Fig. 11. The results indicate that the performance deteriorates as the duration of the glitch increases, and detailed results are presented in Table 5. Specifically, for short-duration glitches, the maximum tolerable magnitude of the glitch is about $1\times10^{-13}\,\mathrm{ms}^{-2}$, while for long-duration glitches, the maximum tolerable magnitude is about $1\times10^{-14}\,\mathrm{ms}^{-2}$. We can see that the long-duration glitch has a greater impact on the extraction results. This is because we adjust its duration to be consistent with the duration of the merge phase of the MBHB signal, which greatly affects the model's judgment, resulting in erroneous extraction.

5.4. Mixing three types of non-stationarities

At last but not least, we conducted tests by mixing three types of non-stationarities and have listed in Table 6 the results achieved by our model under different conditions. When all these combine, our model still exhibits considerable extraction capability. When long-term scheduled gap exists, our model can tolerate with time-varying noise PSDs together with short-duration glitches of maximum peak value $\sim 5\times10^{-14}\,\mathrm{ms}^{-2}$ or long-duration glitches $\sim 1\times10^{-14}\,\mathrm{ms}^{-2}$. When the occurrence frequency of unscheduled gaps reaches about once every 10 hours, our model can tolerate with relative weaker changes in noise (residual acceleration noise) PSDs combined with short-duration glitches of maximum peak value $\sim 5\times10^{-14}\,\mathrm{ms}^{-2}$ or long-duration glitches $\sim 1\times10^{-14}\,\mathrm{ms}^{-2}$. These tests give typically the possible extreme cases in real science operations and measurements of space-borne GW antennas, and our model have shown high robustness and feasibility in typical signal extractions. It can be observed that high-frequency unscheduled gaps may have important impacts, which is predominantly due to substantial loss of spectral information and further compounded by additional anomalies that significantly obfuscate the model's signal extraction capabilities. In future works, we will try to improve the model's ability to tolerate these mixed situations in order to improve performance in more extreme situations, and we will focus on unscheduled gaps, improving the model's tolerance for this situation and mitigating its impact on other anomalies.

6. Concluding remarks

In this work, we developed a deep learning model for accurate GW extractions against possible data anomalies or non-stationarities for space-borne GW antennas, such as LISA and LISA-like missions, which also provide robustness studies of deep learning models in data processing of space-borne GW missions. Our research focuses on the three types of non-stationarities, including gaps, glitches, and time varying noise auto-correlations. Compared with the current state-of-the-art models, our model exhibits the same performance for the ideal and anomaly free cases in signal extractions. While, confronted with the three types of non-stationarities and even their mixtures in some extreme cases, our model still shows considerable adaptability. It is worth noting that when all anomalies occur at the same time, the model's tolerance for these anomalies would degrade compared with each stand-alone situation, and among these the unscheduled gap turns out to be a more important factor.

Deep learning has provided a method for the rapid detection of space-borne GW missions, and robustness research is key to ensuring that deep learning models are applicable in real detection scenarios. In the future, we will use this signal extraction model as a pre-processing model to reduce the difficulty of subsequent scientific analysis such as GW signal detection and parameter estimation through robust extraction of GW signals, becoming an important part of the Taiji detection pipeline. It is planned to process longer datasets in future works and to incorporate a greater variety of realistic simulation scenarios, including various levels of SNR and types of anomalies. Additionally, we will focus on unscheduled gaps, seeking new improvements to mitigate the substantial impact of such anomalies on model performance.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work is supported by the National Key Research and Development Program of China No. 2021YFC2201901, No. 2021YFC2201903, No. 2020YFC2200601 and No. 2020YFC2200901.

References

[1] B.P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al., Gw150914: the advanced ligo detectors in the era of first discoveries, Phys. Rev. Lett. 116 (2016) 131103.
[2] B.P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al., Observation of gravitational waves from a binary black hole merger, Phys. Rev. Lett. 116 (2016) 061102.
[3] B.P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al., Gw151226: observation of gravitational waves from a 22-solar-mass binary black hole coalescence, Phys. Rev. Lett. 116 (2016) 241103.
[4] B.P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al., Binary black hole mergers in the first advanced ligo observing run, Phys. Rev. X 6 (2016) 041015.
[5] L. Scientific, B.P. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al., Gw170104: observation of a 50-solar-mass binary black hole coalescence at redshift 0.2, Phys. Rev. Lett. 118 (2017) 221101.

[6] B.P. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya, et al., Gw170608: observation of a 19 solar-mass binary black hole coalescence, Astrophys. J. Lett. 851 (2017) L35.

[7] B.P. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R.X. Adhikari, V.B. Adya, et al., Gw170814: a three-detector observation of gravitational waves from a binary black hole coalescence, Phys. Rev. Lett. 119 (2017) 141101.

[8] B.P. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V.B. Adya, et al., Gw170817: observation of gravitational waves from a binary neutron star inspiral, Phys. Rev. Lett. 119 (2017) 161101.

[9] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, et al., Gw190425: observation of a compact binary coalescence with total mass 3.4 m, Astrophys. J. 892 (2020) L3.

[10] R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R.X. Adhikari, V. Adya, C. Affeldt, M. Agathos, et al., Gw190412: observation of a binary-black-hole coalescence with asymmetric masses, Phys. Rev. D 102 (2020) 043015.

[11] R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, M. Agathos, et al., Gw190521: a binary black hole merger with a total mass of 150 m, Phys. Rev. Lett. 125 (2020) 101102.

[12] R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R.X. Adhikari, V. Adya, C. Affeldt, et al., Tests of general relativity with binary black holes from the second ligo-virgo gravitational-wave transient catalog, Phys. Rev. D 103 (2021) 122002.

[13] J.M. Ezquiaga, Hearing gravity from the cosmos: gwtc-2 probes general relativity at cosmological scales, Phys. Lett. B 822 (2021) 136665.

[14] G. Bozzola, V. Paschalidis, General relativistic simulations of the quasicircular inspiral and merger of charged black holes: Gw150914 and fundamental physics implications, Phys. Rev. Lett. 126 (2021) 041103.

[15] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, et al., Gwtc-1: a gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs, Phys. Rev. X 9 (2019) 031040.

[16] B.P. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. Adya, C. Affeldt, M. Agathos, et al., Prospects for observing and localizing gravitational-wave transients with advanced ligo, advanced virgo and kagra, Living Rev. Relativ. 23 (2020) 1.

[17] A. Freise, K. Strain, Interferometer techniques for gravitational-wave detection, Living Rev. Relativ. 13 (2010) 1.

[18] P. Amaro-Seoane, H. Audley, S. Babak, J. Baker, E. Barausse, P. Bender, E. Berti, P. Binetruy, M. Born, D. Bortoluzzi, et al., Laser interferometer space antenna, arXiv preprint arXiv:1702.00786, 2017.

[19] J. Baker, J. Bellovary, P.L. Bender, E. Berti, R. Caldwell, J. Camp, J.W. Conklin, N. Cornish, C. Cutler, R. DeRosa, et al., The laser interferometer space antenna: unveiling the millihertz gravitational wave sky, arXiv preprint arXiv:1907.06482, 2019.

[20] X. Gong, S. Xu, S. Bai, Z. Cao, G. Chen, Y. Chen, X. He, G. Heinzel, Y.-K. Lau, C. Liu, et al., A scientific case study of an advanced lisa mission, Class. Quantum Gravity 28 (2011) 094012.

[21] W.-R. Hu, Y.-L. Wu, The Taiji Program in Space for Gravitational Wave Physics and the Nature of Gravity, 2017.

[22] J. Luo, L.-S. Chen, H.-Z. Duan, Y.-G. Gong, S. Hu, J. Ji, Q. Liu, J. Mei, V. Milyukov, M. Sazhin, et al., Tianqin: a space-borne gravitational wave detector, Class. Quantum Gravity 33 (2016) 035010.

[23] K. Dey, N. Karnesis, A. Toubiana, E. Barausse, N. Korsakova, Q. Baghi, S. Basak, Effect of data gaps on the detectability and parameter estimation of massive black hole binaries with lisa, Phys. Rev. D 104 (2021) 044035.

[24] Q. Baghi, N. Korsakova, J. Slutsky, E. Castelli, N. Karnesis, J.-B. Bayle, Detection and characterization of instrumental transients in lisa pathfinder and their projection to lisa, Phys. Rev. D 105 (2022) 042002.

[25] M.C. Edwards, P. Maturana-Russel, R. Meyer, J. Gair, N. Korsakova, N. Christensen, Identifying and addressing nonstationary lisa noise, Phys. Rev. D 102 (2020) 084062.

[26] M.L.D.C.T. Force, K.A. Arnaud, S. Babak, J.G. Baker, M.J. Benacquista, N.J. Cornish, C. Cutler, S.L. Larson, B. Sathyaprakash, M. Vallisneri, et al., An Overview of the Mock Lisa Data Challenges, In AIP Conference Proceedings, vol. 873, American Institute of Physics, 2006, pp. 619–624.

[27] M.L. Katz, Fully automated end-to-end pipeline for massive black hole binary signal extraction from lisa data, Phys. Rev. D 105 (2022) 044055.

[28] M. Vallisneri, Synthetic lisa: simulating time delay interferometry in a model lisa, Phys. Rev. D 71 (2005) 022001.

[29] L.S. Finn, Detection, measurement, and gravitational radiation, Phys. Rev. D 46 (1992) 5236.

[30] S.A. Usman, A.H. Nitz, I.W. Harry, C.M. Biwer, D.A. Brown, M. Cabero, C.D. Capano, T. Dal Canton, T. Dent, S. Fairhurst, et al., The pycbc search for gravitational waves from compact binary coalescence, Class. Quantum Gravity 33 (2016) 215004.

[31] T. Zhao, R. Lyu, H. Wang, Z. Cao, Z. Ren, Space-based gravitational wave signal detection and extraction with deep neural network, Commun. Phys. 6 (2023) 212.

[32] H. Wang, S. Wu, Z. Cao, X. Liu, J.-Y. Zhu, Gravitational-wave signal recognition of ligo data by deep learning, Phys. Rev. D 101 (2020) 104003.

[33] V. De Luca, G. Franciolini, P. Pani, A. Riotto, Primordial black holes confront ligo/virgo data: current situation, J. Cosmol. Astropart. Phys. 06 (2020) 044.

[34] J. Chen, A. Cazenave, C. Dahle, W. Llovel, I. Panet, J. Pfeffer, L. Moreira, Applications and challenges of grace and grace follow-on satellite gravimetry, Surv. Geophys. 1 (2022).

[35] M. Armano, M. Benedetti, J. Bogenstahl, D. Bortoluzzi, P. Bosetti, N. Brandt, A. Cavalleri, G. Ciani, I. Cristofolini, A. Cruise, et al., Lisa pathfinder: the experiment and the route to lisa, Class. Quantum Gravity 26 (2009) 094001.

[36] T.S. Collaboration, Y.-L. Wu, Z.-R. Luo, J.-Y. Wang, M. Bai, W. Bian, H.-W. Cai, R.-G. Cai, Z.-M. Cai, J. Cao, et al., Taiji Program in Space for Gravitational Universe with the First Run Key Technologies Test in Taiji-1, 2021.

[37] Q. Baghi, J.I. Thorpe, J. Slutsky, J. Baker, T. Dal Canton, N. Korsakova, N. Karnesis, Gravitational-wave parameter estimation with gaps in lisa: a Bayesian data augmentation method, Phys. Rev. D 100 (2019) 022003.

[38] M. Armano, H. Audley, J. Baird, P. Binetruy, M. Born, D. Bortoluzzi, E. Castelli, A. Cavalleri, A. Cesarini, A. Cruise, et al., Beyond the required lisa free-fall performance: new lisa pathfinder results down to 20 μ hz, Phys. Rev. Lett. 120 (2018) 061101.

[39] T. Robson, N.J. Cornish, Detecting gravitational wave bursts with lisa in the presence of instrumental glitches, Phys. Rev. D 99 (2019) 024019.

[40] M. Armano, H. Audley, J. Baird, P. Binetruy, M. Born, D. Bortoluzzi, E. Castelli, A. Cavalleri, A. Cesarini, V. Chiavegato, et al., Transient acceleration events in lisa pathfinder data: properties and possible physical origin, Phys. Rev. D 106 (2022) 062001.

[41] M. Armano, H. Audley, G. Auger, J.T. Baird, M. Bassan, P. Binetruy, M. Born, D. Bortoluzzi, N. Brandt, M. Caleno, et al., Sub-femto-g free fall for space-based gravitational wave observatories: lisa pathfinder results, Phys. Rev. Lett. 116 (2016) 231101.

[42] B. Frommknecht, Integrated sensor analysis of the GRACE mission, Ph.D. thesis, Technische Universität München, 2007.

[43] B. Sheard, G. Heinzel, K. Danzmann, D. Shaddock, W. Klipstein, W. Folkner, Intersatellite laser ranging instrument for the grace follow-on mission, J. Geod. 86 (2012) 1083.

[44] China's First Step Towards Probing the Expanding Universe and the Nature of Gravity Using a Space Borne Gravitational Wave Antenna, Commun. Phys. 4 (2021) 34.

[45] A. Spadaro, R. Buscicchio, D. Vetrugno, A. Klein, D. Gerosa, S. Vitale, R. Dolesi, W.J. Weber, M. Colpi, Glitch systematics on the observation of massive black-hole binaries with lisa, arXiv preprint arXiv:2306.03923, 2023.

[46] D. George, E. Huerta, Deep neural networks to enable real-time multimessenger astrophysics, Phys. Rev. D 97 (2018) 044039.

[47] H. Gabbard, M. Williams, F. Hayes, C. Messenger, Matching matched filtering with deep networks for gravitational-wave astronomy, Phys. Rev. Lett. 120 (2018) 141103.

[48] H. Gabbard, C. Messenger, I.S. Heng, F. Tonolini, R. Murray-Smith, Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy, Nat. Phys. 18 (2022) 112.

[49] M. Dax, S.R. Green, J. Gair, J.H. Macke, A. Buonanno, B. Schölkopf, Real-time gravitational wave science with neural posterior estimation, Phys. Rev. Lett. 127 (2021) 241103.

[50] R.E. Colgan, K.R. Corley, Y. Lau, I. Bartos, J.N. Wright, Z. Márka, S. Márka, Efficient gravitational-wave glitch identification from environmental data through machine learning, Phys. Rev. D 101 (2020) 102003.

[51] M. Cavaglia, K. Staats, T. Gill, Finding the origin of noise transients in ligo data with machine learning, Commun. Comput. Phys. 25 (2019).

[52] M. Razzano, E. Cuoco, Image-based deep learning for classification of noise transients in gravitational wave detectors, Class. Quantum Gravity 35 (2018) 095016.

[53] R. Ormiston, T. Nguyen, M. Coughlin, R.X. Adhikari, E. Katsavounidis, Noise reduction in gravitational-wave data via deep learning, Phys. Rev. Res. 2 (2020) 033066.

[54] A. Torres-Forne, A. Marquina, J.A. Font, J.M. Ibanez, Denoising of gravitational wave signals via dictionary learning algorithms, Phys. Rev. D 94 (2016) 124040.

[55] W. Wei, E. Huerta, Gravitational wave denoising of binary black hole mergers with deep learning, Phys. Lett. B 800 (2020) 135081.

[56] K. Cannon, S. Caudill, C. Chan, B. Cousins, J.D. Creighton, B. Ewing, H. Fong, P. Godwin, C. Hanna, S. Hooper, et al., Gstlal: a software framework for gravitational wave discovery, SoftwareX 14 (2021) 100680.

[57] L. Gondara, Medical image denoising using convolutional denoising autoencoders, in: 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), IEEE, 2016, pp. 241–246.

[58] J. Xie, L. Xu, E. Chen, Image denoising and inpainting with deep neural networks, Adv. Neural Inf. Process. Syst. 25 (2012).

[59] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, L. Bottou, Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion, J. Mach. Learn. Res. 11 (2010).

[60] A. Creswell, A.A. Bharath, Denoising adversarial autoencoders, IEEE Trans. Neural Netw. Learn. Syst. 30 (2018) 968.

[61] H.-T. Chiang, Y.-Y. Hsieh, S.-W. Fu, K.-H. Hung, Y. Tsao, S.-Y. Chien, Noise reduction in ecg signals using fully convolutional denoising autoencoders, IEEE Access 7 (2019) 60806.

[62] O.M. Saad, Y. Chen, Deep denoising autoencoder for seismic random noise attenuation, Geophysics 85 (2020) V367.

[63] M. Xia, T. Li, L. Liu, L. Xu, C.W. de Silva, Intelligent fault diagnosis approach with unsupervised feature learning by stacked denoising autoencoder, IET Sci. Meas. Technol. 11 (2017) 687.

[64] E. Dasan, I. Panneerselvam, A novel dimensionality reduction approach for ecg signal via convolutional denoising autoencoder with lstm, Biomed. Signal Process. Control 63 (2021) 102225.
[65] S. Araki, T. Hayashi, M. Delcroix, M. Fujimoto, K. Takeda, T. Nakatani, Exploring multi-channel features for denoising-autoencoder-based speech enhancement, in: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2015, pp. 116–120.
[66] Y.-H. Lai, F. Chen, S.-S. Wang, X. Lu, Y. Tsao, C.-H. Lee, A deep denoising autoencoder approach to improving the intelligibility of vocoded speech in cochlear implant simulation, IEEE Trans. Biomed. Eng. 64 (2016) 1568.
[67] X. Feng, Y. Zhang, J. Glass, Speech feature denoising and dereverberation via deep autoencoders for noisy reverberant speech recognition, in: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2014, pp. 1759–1763.
[68] X. Lu, Y. Tsao, S. Matsuda, C. Hori, Speech enhancement based on deep denoising autoencoder, in: Interspeech, vol. 2013, 2013, pp. 436–440.
[69] C. Chatterjee, L. Wen, F. Diakogiannis, K. Vinsen, Extraction of binary black hole gravitational wave signals from detector data using deep learning, Phys. Rev. D 104 (2021) 064046.
[70] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. 25 (2012).
[71] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556, 2014.
[72] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[73] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436.
[74] M. Tinto, S.V. Dhurandhar, Time-delay interferometry, Living Rev. Relativ. 24 (2021) 1.
[75] J. Armstrong, F. Estabrook, M. Tinto, Time-delay interferometry for space-based gravitational wave searches, Astrophys. J. 527 (1999) 814.
[76] S. Babak, M. Hewitson, A. Petiteau, L.I.S.A. Sensitivity, S.N.R. Calculations, arXiv: 2108.01167 [astro-ph,physics:gr-qc], 2021.
[77] Z. Luo, Z. Guo, G. Jin, Y. Wu, W. Hu, A brief analysis to Taiji: science and technology, Results Phys. 16 (2020) 102918.
[78] A. Nitz, I. Harry, D. Brown, C.M. Biwer, J. Willis, T. Dal Canton, C. Capano, T. Dent, L. Pekowsky, A.R. Williamson, et al., Gwastro/Pycbc: v2. 0.2 Release of Pycbc, Zenodo, 2022.
[79] A. Bohé, L. Shao, A. Taracchini, A. Buonanno, S. Babak, I.W. Harry, I. Hinder, S. Ossokine, M. Pürrer, V. Raymond, et al., Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors, Phys. Rev. D 95 (2017) 044028.
[80] G. Wang, W.-T. Ni, W.-B. Han, S.-C. Yang, X.-Y. Zhong, Numerical simulation of sky localization for lisa-taiji joint observation, Phys. Rev. D 102 (2020) 024089.
[81] G. Wang, W.-T. Ni, Revisiting time delay interferometry for unequal-arm lisa and taiji, Phys. Scr. 98 (2023) 075005.
[82] F.I. Diakogiannis, F. Waldner, P. Caccetta, Looking for change? Roll the dice and demand attention, arXiv preprint arXiv:2009.02062, 2020.