An Investigation on the Effects of Non-Gaussian Noise Transients and Their Mitigations to Tests of General Relativity

Jack Y. L. Kwok*
Department of Physics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Mentors: Alan J. Weinstein, Rico K. L. Lo
LIGO, California Institute of Technology, Pasadena, California 91125, USA
(Dated: July 9, 2020)

I. INTRODUCTION

Over a century after its formulation in 1915, Einstein's General Relativity (GR) remains as the accepted theory of gravity, passing all precision tests to date [1]. In the weak-field, slow-motion regime, where the effects of metric theories of gravity can be approximated as higher-order *post-Newtonian* (PN) corrections to the Newtonian theory [2], GR lies within the stringent bounds set by solar-system tests and pulsar tests [3, 4].

Recent attention has turned to testing GR in the strong-field, highly-relativistic regime [3], which potentially suggests high-energy modifications which make GR compatible with standard quantum field theory [1]. One approach to probe the strong-field regime is through the detection of gravitational waves (GWs), a form of gravitational radiation predicted by GR which travels at the speed of light.

Far from the source, GWs are *small* perturbations about a background spacetime; assuming a flat Minkowski background spacetime, a propagating GW is determined by two time-dependent scalar quantities h_+ and $h_\times$, called the *plus* and *cross polarizations* respectively [2], which carries information about its astrophysical origin [5]. In the detector frame, GW exerts tidal forces and changes the separation between freely-falling objects [2]. Using laser interferometry, a beam-splitter splits incident laser beam into two equal parts and directs them to Fabry-Pérot cavities in two orthogonal multi-kilometer-long arms [6]; a GW passing through the detector would change the separation of "freely-falling" mirrors at the ends of each arm to different extents, resulting in a phase difference between the two beams that could be measured in terms of the light intensity of the recombined beam [6].

Since 2015, Advanced LIGO [7] and Advanced Virgo [8] have jointly announced 14 confident detections of GWs, all of which are generated by the *coalescence* of binary black holes or binary neutron starss [9–12]. The coalescence of binary compact objects begins as their orbital separation continuously decreases due to emission of GWs during the *inspiral* phase, until the point when the compact objects are so close to each other they plunge together close to the speed of light and *merge* into a single black hole, which quickly settles down to a Kerr black hole during the *ringdown* phase [13, 14].

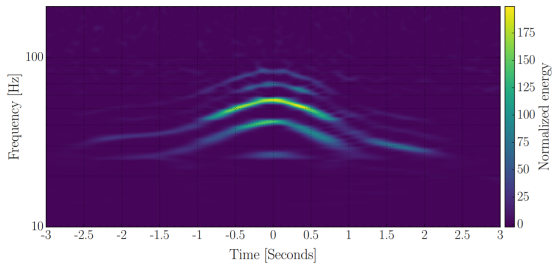
Of all strong-field astrophysical events that could be probed, the coalescence of stellar-mass binary black holes (BBHs) plays a crucial role in testing GR [1]: Being compact objects, the orbital separation of the BBH can reach far below the last stable orbit before merging, generating a gravitational field many order of magnitudes larger than other observed astrophysical events [14–16]; moreover, GWs emitted by coalescing BBHs offers one of the cleanest test of GR, as environmental effects such as accretion disks and electromagnetic fields are negligible for most sources [17], enabling precision tests of the strong-field dynamics of GR.

As no high-accuracy coalescing BBH waveform models for alternative theories are avaliable [16], several *generic* tests of GR are devised: Consistency tests analyze the discrepancy between observational data and the most probable waveform model derived from GR for the data [18], and compare source parameters inferred only from inspiral data to that inferred only from merger-ringdown data [18]; parameterized tests introduce deviation to waveform models derived from GR, where a deviation from GR can be inferred using methods of parameter estimation or hypothesis testing [16, 19]. To this date, no evidence for violations of GR has been identified from GWs emitted by coalescing BBHs [20].

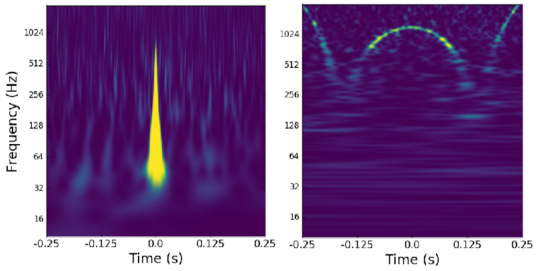
Aside from GWs, the detector output can be attributed to many independent sources of random noise [6]; assuming that noise characteristics remain stationary over time, and by appealing to the central limit theorem [21], noise in GW detectors are typically modeled to be Gaussian within data analysis routines used in tests of GR [22, 23]. However, these assumptions cannot account for transient, non-Gaussian noise features which enter gravitational-wave detectors, commonly-known as *glitches* [24–26]. Three classes of commonly-seen glitches are shown in Figure 1. If the presence of glitches were not accounted for, one may infer from the detected waveform that a deviation from GR has occurred. The extent to which glitches mimic the effects of a deviation of GR and the effectiveness of glitch mitigations to tests of GR certainly deserve a study.

This report is structured as follows: In Section II we introduce the waveform model we use to generate simulated signals and perform tests of GR. In Section III

* Email: jackkwok@link.cuhk.edu.hk



(a) A light scattering glitch [27]



(b) A blip glitch [25] (c) A whistle glitch [25]

FIG. 1. Three normalized spectrograms of commonly-seen glitches. The colour represents the 'loudness' of the signal at each time-frequency bin [25]. (a) A light scattering glitch has a characteristic arch shape; it is caused by slight misalignments of the laser beam and the mirrors [27]. (b) A blip glitch has a characteristic 'teardrop' shape; its noise source has not been identified [28]. (c) A whistle glitch has a characteristic 'W' or 'V' shape; it is caused by radio signals generated by the interferometer control system [24].

we discuss the parameterization of the waveform model and the two tests that we employ. In Section IV we introduce three commonly-used mitigation measures for glitches. In Section V, we describe our method of studying the effect of glitches and their mitigations on tests of GR.

II. WAVEFORM MODEL

In our study, we will use the frequency-domain precessing inspiral-merger-ringdown waveform model `IMRPhenomPv2` derived from GR [29–31]to generate simulated signals and perform test of GR in virtue of its good match with Numerical Relativity waveforms [31] and low computational costs.

`IMRPhenomPv2` is a "hybrid" waveform model [32] constructed by matching and combining PN waveforms [33] applicable to early inspiral, with Numerical Relativity waveforms [34] used in merger and ringdown. Its phase composes of terms with known frequency dependence; the coefficients of these terms, denoted as the *phase coefficients* p_i, are the subjects of parameterized tests of GR in Section III. The phase coefficients p_i can be categorized into three groups, depending on the stages of coalescence in which they predominantly assert their effect on [16, 31]: (i) *inspiral* PN coefficients $\{\varphi_0, ..., \varphi_5, \varphi_{5l}, \varphi_6, \varphi_{6l}, \varphi_7\}$ and phenomenological coefficients $\{\sigma_0, ..., \sigma_4\}$; (ii) the *intermediate* phenomenological coefficients $\{\beta_0, ..., \beta_3\}$; (iii) the *merger-ringdown* phenomenological and black hole perturbation theory coefficients $\{\alpha_0, ..., \alpha_5\}$.

Stage of coalescence	δp_i	f-dependence
Inspiral	$\delta\varphi_0$	$f^{-5/3}$
	$\delta\varphi_1$	$f^{-4/3}$
	$\delta\varphi_2$	f^{-1}
	$\delta\varphi_3$	$f^{-2/3}$
	$\delta\varphi_4$	$f^{-1/3}$
	$\delta\varphi_{5l}$	$\log(f)$
	$\delta\varphi_6$	$f^{1/3}$
	$\delta\varphi_{6l}$	$f^{1/3}\log(f)$
	$\delta\varphi_7$	$f^{2/3}$
Intermediate	$\delta\beta_2$	$\log f$
	$\delta\beta_3$	f^{-3}
Merger-Ringdown	$\delta\alpha_2$	f^{-1}
	$\delta\alpha_3$	$f^{3/4}$
	$\delta\alpha_4$	$\tan^{-1}(af+b)$

TABLE I. The frequency dependence of `IMRPhenomPv2` dephasing coefficients used in parameterized tests of GR. The table is reproduced from Table 1 of Ref. [18].

III. PARAMETERIZED TESTS OF GR

In this project, we will focus on a parameterized test of GR, which introduces *fractional* deviations δp_i of the `IMRPhenomPv2` phase coefficients p_i, also known as *dephasing coefficients*, from their prescribed value: [35]:

$$p_i \mapsto p_i[1 + \delta p_i] \,. \tag{1}$$

In practice, we do not allow some of the `IMRPhenomPv2` phase coefficients to deviate from their prescribed value as they have large uncertainties or are degenerate with with other coefficients or physical parameters [16]. We therefore perform tests with the remaining 13 dephasing coefficients, henceforth denoted as the testing dephasing coefficients [16]:

$$\{\delta p_i\} = \{\delta\varphi_0, ..., \delta\varphi_4, \delta\varphi_{5l}, \delta\varphi_6, \delta\varphi_{6l}, \delta\varphi_7, \delta\beta_2, \delta\beta_3, \delta\alpha_2, \delta\alpha_3, \delta\alpha_4\} \,, \tag{2}$$

the frequency dependence of the testing dephasing coefficients δp_i is shown in Table I [18, 30].

To quantify a deviation from GR, we can either compare GR against a modified theory of gravity through hypothesis testing, or infer the most probable values of δp_i through parameter estimation.

A. Hypothesis Testing

With in mind that the two-body self-gravitating problem cannot be solved analytically in GR, the orbital evolution of coalescing BBHs and tests of GR derived from it can only be computed *up to a precision*. We follow the TIGER framework [16, 19] and define $\mathcal{H}_{\rm GR}$ as the hypothesis that the gravitational-wave signal has the exact functional form predicted by a chosen approximation to GR, i.e. $\delta p_i = 0$ [35]; to test against this hypothesis, we denote $\mathcal{H}_{\rm modGR}$ as the hypothesis that $\boldsymbol{h}$ has the functional form predicted by the chosen approximation to GR but with at least one $\delta p_i \neq 0$.

It is evident that the two hypotheses $\mathcal{H}_{\rm GR}$ and $\mathcal{H}_{\rm modGR}$ are mutually exclusive. Given data $\boldsymbol{d}$ and prior information I, we prefer the hypothesis which is relatively more probable. To quantify this statement, we can define the *odds*

$$O^{\rm modGR}_{\rm GR} \equiv \frac{P(\mathcal{H}_{\rm modGR}|\boldsymbol{d}, I)}{P(\mathcal{H}_{\rm GR}|\boldsymbol{d}, I)} \,. \tag{3}$$

If the odds is much greater than one, we prefer the hypothesis $\mathcal{H}_{\rm modGR}$; if it is much less than one, we prefer $\mathcal{H}_{\rm GR}$. If the odds is close to 1, then the current data is inconclusive [36]. Invoking Bayes' Theorem, the odds can be rewritten as

$$O^{\rm modGR}_{\rm GR} = \frac{P(\mathcal{H}_{\rm modGR}|I)}{P(\mathcal{H}_{\rm GR}|I)} \times \frac{P(\boldsymbol{d}|\mathcal{H}_{\rm modGR}, I)}{P(\boldsymbol{d}|\mathcal{H}_{\rm GR}, I)} \,. \tag{4}$$

The second term on the R.H.S. of Eq. (4) is called the *Bayes factor*, which could be readily computed with the support of Bayesian inference libraries [22, 37].

Applying the treatment of $\mathcal{H}_{\rm modGR}$ introduced in TIGER [19], we can compare the odds of unmitigated and mitigated samples against each other, or against a predefined threshold set as the odds obtained for the collection of detected GW signals so far [35]. A deviation from GR due to the false assumption of Gaussian noise can then be visualized by plotting the obtained odds on top of a simulated *background distribution*, defined as the probability density of the odds $\mathcal{O}^{\rm modGR}_{\rm GR}$ [35], for simulated GW signals derived from GR injected in Gaussian noise, as shown in Figure 2.

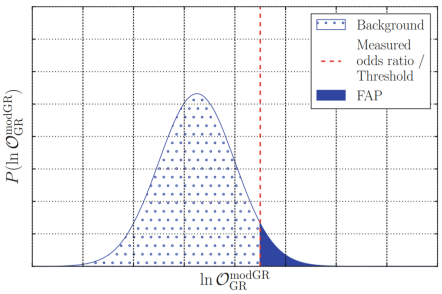


FIG. 2. The background distribution (region under blue curve) is plotted for a collection of simulated GR-consistent signals in Gaussian noise. Deviations of GR with odds higher than a defined *threshold* (red dotted line) are vetoed; the threshold can be set as the odds obtained for the collection of detected gravitational-wave signals. The *false alarm probability* (FAP) (shaded region), defined as the probability of obtaining a odds above the threshold, is shown to be an invariant quantity under choices of prior during hypothesis testing [35]. This figure is retrieved from Li [35].

IV. GLITCH MITIGATION

Many efforts are made to develop algorithms that identify glitches [38–41]; these algorithms play an important role in gravitational-wave searches. Once a glitch is identified, it could be eliminated either by hand or automatically by search pipelines [42, 43] through a process called *gating*, which zeroes out the time interval containing the glitch by multiplying the time series with an *inverse* window function [42, 43]. An example of gating is illustrated in Figure 3.

A similar procedure can be done in the frequency domain: if the glitch is localized in certain intervals of frequency, zeroing out the corresponding frequency bins through band-pass filtering would eliminate the glitch.

A more sophisticated approach would be to subtract off a glitch model from the original time series. This procedure, called *de-glitching*, was employed for the glitch-contaminated GW170817 data [44], as illustrated in Figure 3. The de-glitching procedure can be extended beyond well-modeled glitches using the `BayesWave` framework, which introduces a method to model glitches using sine-Gaussian wavelets and infer the most probable model using Bayesian statistics [45, 46].

In our study, we will separately apply the three standard mitigation measures of 1) gating, 2) band-pass filtering and 3) de-glitching to data samples.

V. METHODOLOGY

Our goal is to investigate the extent to which glitches mimic the effects of a deviation of GR, and evaluate the effectiveness of common glitch-related mitigation measures presented in Section IV on parameterized tests of GR. To this end, we first prepare a collection of data samples by injecting simulated `IMRPhenomPv2` signals *coherently* into LIGO Hanford (H1) and Livingston (L1) data segments where commonly-seen glitches are present.

On the one hand, the simulated signal is chosen to be the most probable (*maximum a posteriori*, or MAP) `IMRPhenomPv2` waveform for one or two GW events; a concrete set of criteria is not yet established for the selection of GW events. On the other hand, glitches with different signal-to-noise ratio (SNR) and morphologies are chosen for the study; each of them are made to overlap the simulated signal in the inspiral, intermediate and merger-ringdown regime. As a rough estimate, 40 injections will be performed and 160 tests of GR will be

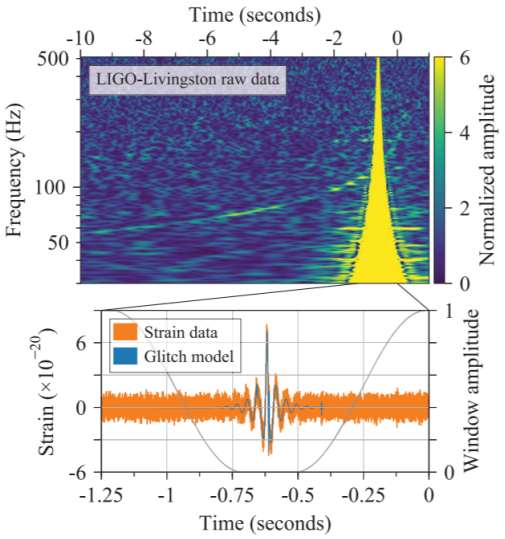


FIG. 3. The output data from the LIGO-Livingston detector during GW170817 is plotted over time in the bottom figure (orange curve). A glitch was identified around the time $t = -0.75$s to -0.5s in the figure. To determine the sky location of the event, data was gated in the time domain by multiplying the inverse window function (black curve) [44]. To infer the source properties during parameter estimation, data was de-glitched by subtracting off a glitch model (blue curve) reconstructed with `BayesWave` [44]. The upper figure shows a spectrogram of the raw LIGO-Livingston data. The figure is retrieved from Abbott et al. [44]

performed after applying the three mitigation measures after applying the three mitigation measures. Comparing the test results between different data samples, the effect of glitches under the false assumption of Gaussian noise and their mitigations can be determined for various GW source properties, glitch SNR and morphologies, and the degree of overlapping.

A. Injection Tool

To minimize human error and time spent on performing injections, a specialized injection tool in Python, `injhelper`, is under development. Under the hood, it is a wrapper which calls `LALSimulation` [22] to generate a time-domain waveform and converts it into a `GWpy` time series [47], it then retrieves (given LIGO credentials) LIGO data centered at a specified GPS time and returns a standard data-frame file (`.gwf` or `.hdf5`) containing the injection; the merger time of the injected signal is set as the specified GPS time. By looking up the LIGO glitch database, GlitchDB, one could specify the injection GPS time as the time corresponding to a glitch occurrence. A few *sanity checks* devised to catch obvious errors are discussed in Appendix B.

Upcoming features of `injhelper` include injection into multiple detectors coherently given the sky location of the BBH coalescence, which is necessary for de-glitching routines [45].

B. Preparing Data Samples

Using `injhelper`, the most probable `IMRPhenomPv2` waveform for the GW event S190828l are injected into H1 detector noise at a GPS time of 1238571602.625 where a light-scattering glitch with SNR of 10.856 is present, and at a GPS time of 1248851234.507 where a blip glitch with SNR of 10.862 is present. Figure 4 shows the alignment between the simulated signal and injected data at the specified injection times and the Q-scans of the injected data and that with a simulated signal subtracted.

The preparation of mitigated data samples is an easy extension from the injection process: the `.gwf` or `.hdf5` files returned by `injhelper` can be directly parsed into `GWpy`, which offers inbuilt functions for gating and band-pass filtering; the `BayesWave` de-glitching routine also supports the returned `GWpy` files.

Appendix A: Updated Timeline

16 / 06 - 28 / 06: Developed Injection Tool
29 / 06 - 12 / 07: Run TIGER on glitched data
First Interim Report
Finish injection tool
13 / 07 - 26 / 07: Visualize and summarize first batch of results
27 / 07 - 09 / 08: **Second Interim Report**
Visualize and summarize all results
10 / 08 - 21 / 08: **Presentation, Final Report**

Appendix B: Sanity Checks for Injection Tool

Results generated from different parts of the `injhelper` routine is used to check for obvious errors. Figure 5 and 6 plots the waveform generated by `LALSimulation` in time and frequency domain respectively and check for unexpected features; Figure 7 and 8 attempts to retrieve a simulated signal which was "blindly" injected into colored Gaussian noise and real H1 detector noise. We conclude that `injhelper` do not produce obvious errors. The signal used in the sanity checks is the most probable `IMRPhenomPv2` waveform for the event S190828l.

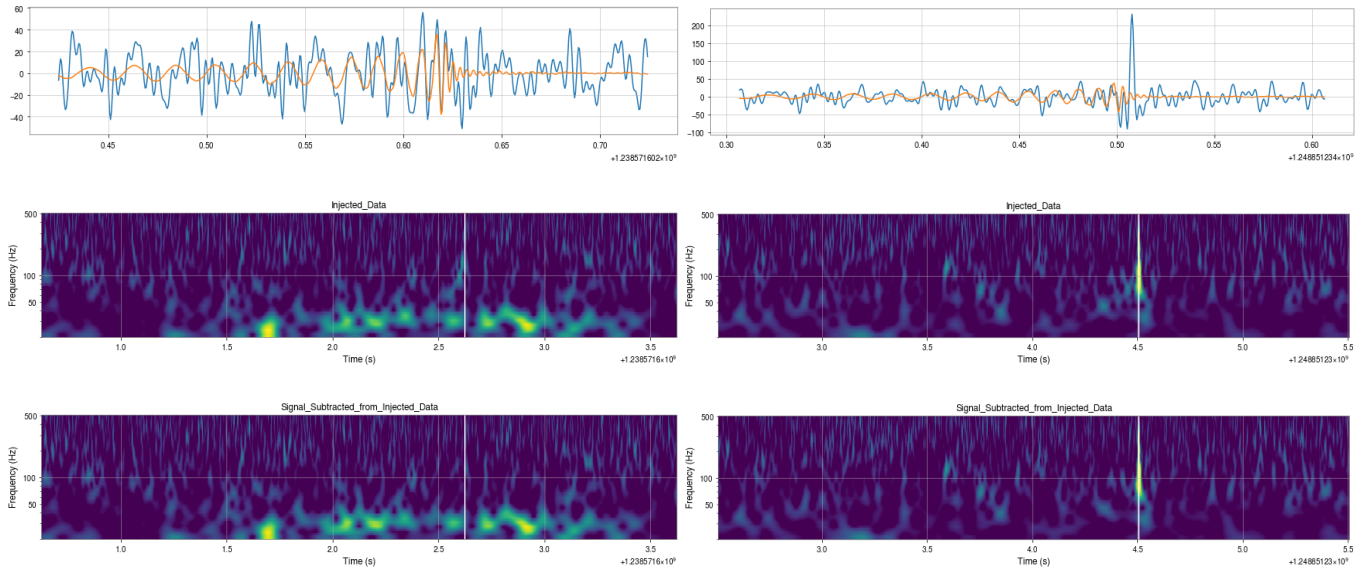


FIG. 4. A simulated signal is injected in real H1 noise where a light scattering glitch (left, GPS time = 1238571602.625) and a blip glitch (right, GPS time = 1248851234.507) is present. The signal (orange) and injected data (blue) are whitened using their respective PSD and plotted in the top figures; the good fit between the signal and data in both cases suggests successful injections. Q-scans of the whitened injected data is shown in the middle figures, where the color denotes the "loudness" of data in the corresponding time-frequency bin, a signal which merges at the specified injection time (white vertical line) can be observed; on the left, a light-scattering glitch can be observed in the low-frequency range spanning over two seconds, on the right, a "loud" blip glitch can be observed at the specified injection time. The characteristic tract of a GW signal due to the coalescence of BBHs cannot be observed from the Q-scans of the remnant data plotted in the bottom figures where a simulated signal merging at the specified injection time is subtracted from the data.

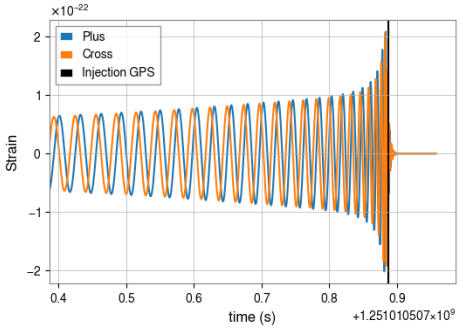


FIG. 5. The time-domain waveforms of a plus-polarized GW (blue) and a cross-polarized GW (orange) generated from the same source parameters are plotted. The cross-polarized waveform is $-\pi/2$ out of phase with respect to the plus-polarized waveform. The merger occurs roughly at specified merger time indicated by the black vertical line.

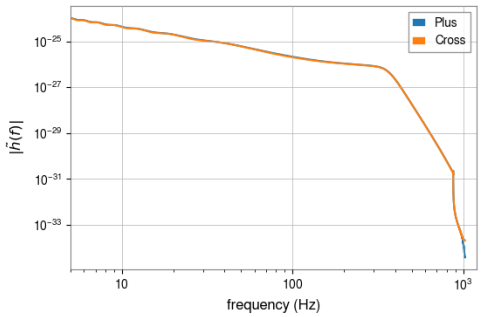


FIG. 6. The modulus of the frequency-domain waveforms for a plus-polarized GW (blue) and a cross-polarized GW (orange) are plotted. The "knee" of the plot, which corresponds to the merger, has a frequency of around 370Hz; this is a reasonable value for a binary system with a total mass of $M \sim 46M_{\odot}$: according to the no-hair theorem, the "knee" frequency $f_{\rm knee}$ scales with $1/M$ [*citation needed*], and for the BBH merger GW150914 where $M \sim 69M_{\odot}$, the "knee" frequency is about 250Hz.

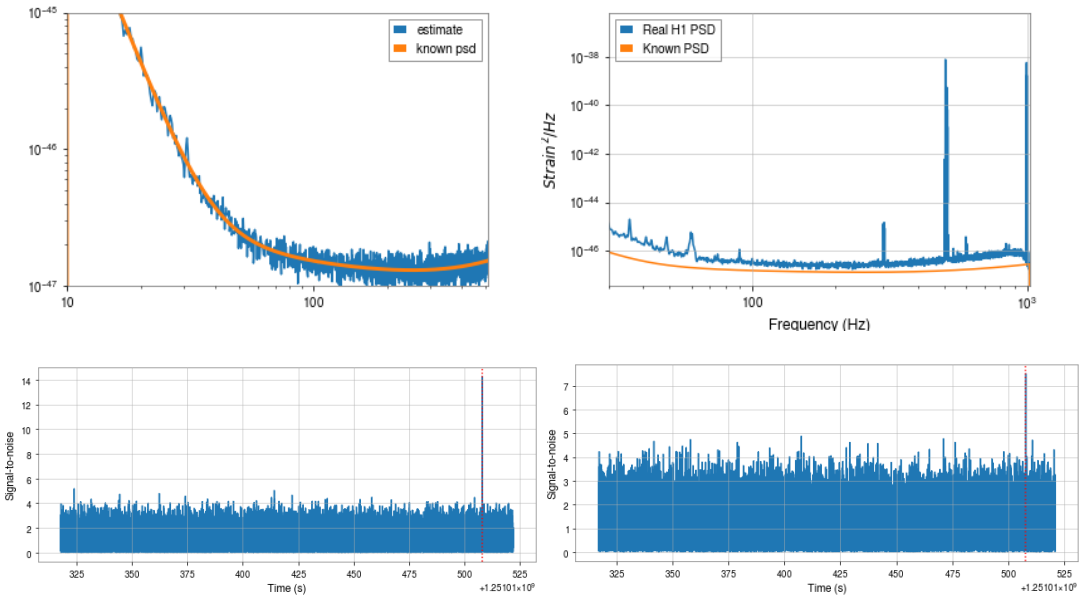


FIG. 7. "Blind" injections are performed on simulated Gaussian noise colored with the predicted zero-detuned high-power PSD (left) and real H1 detector noise (right). The obtained PSD (blue) and the zero-detuned high-power prediction (orange) for the two cases are plotted in the top figures; matched-filtering was performed and the SNR peak is identified at the specified injection time (red dotted line) in the bottom figures. This indicates that successful injections are performed by `injhelper`.

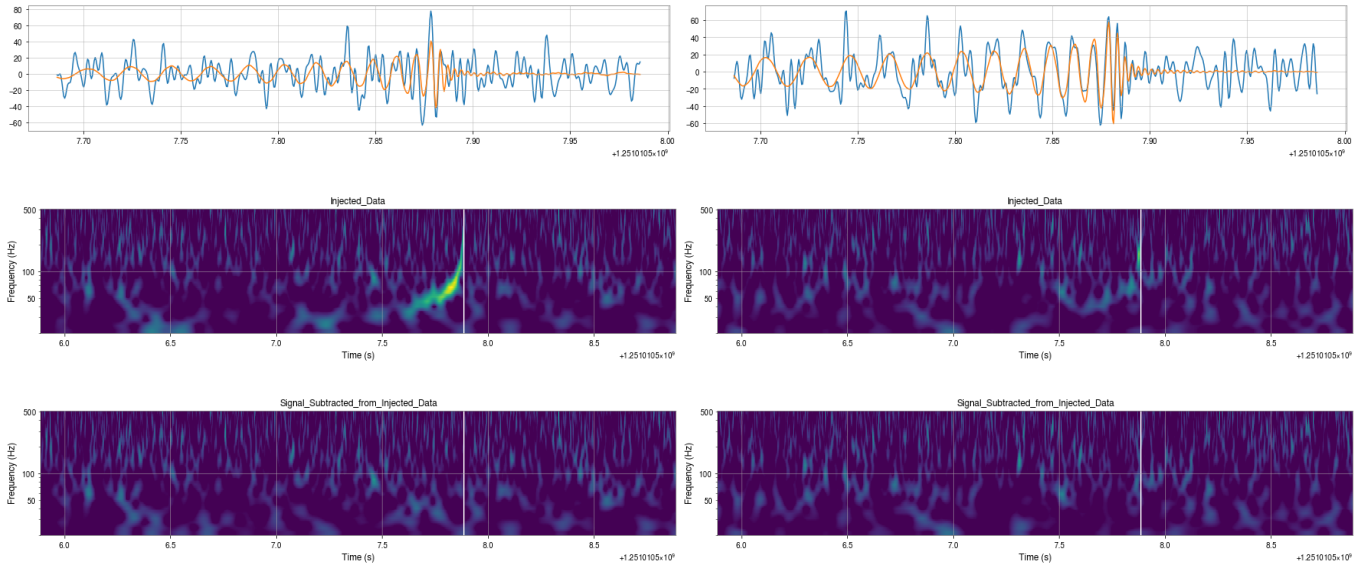


FIG. 8. The simulated signal is aligned to the SNR peaks for simulated Gaussian noise colored with the predicted zero-detuned high-power PSD (left) and real H1 detector noise (right) obtained in Figure 7; both the signal (orange) and data (blue) are whitened using the respective PSD and plotted in the top figures, a great fit between the signal and data can be observed. Q-scans of the whitened injected data is shown in the middle figures, where the color denotes the "loudness" of the data in the corresponding time-frequency bin; a signal which merges at the specified injection time (white vertical line) can be observed. Excess power cannot be observed from the Q-scans of the remnant data plotted in the bottom figures where a simulated signal aligned using the SNR peak is subtracted from the injected data in both cases.

[1] E. Berti, E. Barausse, V. Cardoso, L. Gualtieri, P. Pani, U. Sperhake, L. C. Stein, N. Wex, K. Yagi, T. Baker, *et al.*, Classical and Quantum Gravity **32**, 243001 (2015), arXiv:1501.07274.

[2] C. W. Misner, K. S. Thorne, J. A. Wheeler, *et al.*, *Gravitation* (Macmillan, 1973).

[3] C. M. Will, Living reviews in relativity **9**, 3 (2006), arXiv:0103036.

[4] C. M. Will, Living reviews in relativity **17**, 4 (2014).

[5] B. Sathyaprakash, Living Rev. Relativity **12**, 2 (2009).

[6] P. R. Saulson, *Fundamentals of interferometric gravitational wave detectors* (World Scientific, 1994).

[7] J. Aasi, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, *et al.*, Classical and quantum gravity **32**, 074001 (2015), arXiv:1411.4547 [gr-qc].

[8] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin, *et al.*, Classical and Quantum Gravity **32**, 024001 (2014), arXiv:1408.3978 [gr-qc].

[9] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, *et al.*, Physical Review X **9**, 031040 (2019), arXiv:1811.12907 [astro-ph.HE].

[10] R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, M. Agathos, *et al.*, arXiv preprint arXiv:2004.08342 (2020).

[11] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, *et al.*, The Astrophysical Journal Letters **892**, L3 (2020).

[12] R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, M. Agathos, *et al.*, The Astrophysical Journal Letters **896**, L44 (2020).

[13] F. Pretorius, in *Physics of relativistic objects in compact binaries: From birth to coalescence* (Springer, 2009) pp. 305–369.

[14] N. Yunes, K. Yagi, and F. Pretorius, Physical review D **94**, 084002 (2016).

[15] N. Yunes and X. Siemens, Living Reviews in Relativity **16**, 9 (2013).

[16] J. Meidam, K. W. Tsang, J. Goldstein, M. Agathos, A. Ghosh, C.-J. Haster, V. Raymond, A. Samajdar, P. Schmidt, R. Smith, *et al.*, Physical Review D **97**, 044033 (2018), arXiv:1712.08772 [gr-qc].

[17] E. Barausse, V. Cardoso, and P. Pani, Physical Review D **89**, 104059 (2014).

[18] L. Scientific, V. Collaborations, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, *et al.*, Physical review letters **116**, 221101 (2016).

[19] T. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos, J. Veitch, K. Grover, T. Sidery, R. Sturani, and A. Vecchio, Physical Review D **85**, 082003 (2012), arXiv:1111.5274 [gr-qc].

[20] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, *et al.*, Physical Review D **100**, 104036 (2019).

[21] W. B. Davenport, W. L. Root, *et al.*, *An introduction to the theory of random signals and noise*, Vol. 159 (McGraw-Hill New York, 1958).

[22] LIGO Scientific Collaboration, LIGO Algorithm Library - LALSuite, free software (GPL) (2018).

[23] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, *et al.*, Physical Review D **91**, 042003 (2015), arXiv:1409.7215 [gr-qc].

[24] L. K. Nuttall, T. Massinger, J. Areeda, J. Betzwieser, S. Dwyer, A. Effler, R. Fisher, P. Fritschel, J. Kissel, A. Lundgren, *et al.*, Classical and Quantum Gravity **32**, 245005 (2015), arXiv:1508.07316 [gr-qc].

[25] M. Zevin, S. Coughlin, S. Bahaadini, E. Besler, N. Rohani, S. Allen, M. Cabero, K. Crowston, A. K. Katsaggelos, S. L. Larson, *et al.*, Classical and Quantum Gravity **34**, 064003 (2017), arXiv:1611.04596 [gr-qc].

[26] B. P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, M. Adamo, C. Adams, T. Adams, P. Addesso, *et al.*, Classical and Quantum Gravity **33**, 134001 (2016), arXiv:1602.03844 [gr-qc].

[27] L. Nuttall, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **376**, 20170286 (2018), arXiv:1804.07592 [astro-ph.IM].

[28] M. Cabero, A. Lundgren, A. H. Nitz, T. Dent, D. Barker, E. Goetz, J. S. Kissel, L. K. Nuttall, P. Schale, R. Schofield, *et al.*, Classical and Quantum Gravity **36**, 155010 (2019), arXiv:1901.05093 [physics.ins-det].

[29] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Physical review letters **113**, 151101 (2014), arXiv:1308.3271 [gr-qc].

[30] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, Physical Review D **93**, 044006 (2016), arXiv:1508.07250 [gr-qc].

[31] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Physical Review D **93**, 044007 (2016), arXiv:1508.07253 [gr-qc].

[32] P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, J. Whelan, B. Bruegmann, P. Diener, J. Gonzalez, M. Hannam, *et al.*, Classical and Quantum Gravity **24**, S689 (2007), arXiv:0704.3764 [gr-qc].

[33] A. Buonanno, B. R. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, Physical Review D **80**, 084043 (2009), arXiv:0907.0700 [gr-qc].

[34] A. H. Mroue, M. A. Scheel, B. Szilagyi, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, S. Ossokine, N. W. Taylor, *et al.*, Physical Review Letters **111**, 241104 (2013), arXiv:1304.6077 [gr-qc].

[35] T. G. Li, *Extracting physics from gravitational waves: Testing the strong-field dynamics of general relativity and inferring the large-scale structure of the Universe* (Springer, 2015).

[36] D. Sivia and J. Skilling, *Data analysis: a Bayesian tutorial* (OUP Oxford, 2006).

[37] G. Ashton, M. Hübner, P. D. Lasky, C. Talbot, K. Ackley, S. Biscoveanu, Q. Chu, A. Divakarla, P. J. Easter, B. Goncharov, *et al.*, The Astrophysical Journal Supplement Series **241**, 27 (2019), arXiv:1811.02042 [gr-qc].

[38] J. R. Smith, T. Abbott, E. Hirose, N. Leroy, D. MacLeod, J. McIver, P. Saulson, and P. Shawhan, Classical and Quantum Gravity **28**, 235005 (2011), arXiv:1107.2948

[gr-qc].
[39] T. Isogai, L. S. Collaboration, V. Collaboration, *et al.*, in *Journal of Physics: Conference Series*, Vol. 243 (IOP Publishing, 2010) p. 012005.
[40] R. Essick, L. Blackburn, and E. Katsavounidis, Classical and Quantum Gravity **30**, 155010 (2013), arXiv:1303.7159 [astro-ph.IM].
[41] R. Biswas, L. Blackburn, J. Cao, R. Essick, K. A. Hodge, E. Katsavounidis, K. Kim, Y.-M. Kim, E.-O. Le Bigot, C.-H. Lee, *et al.*, Physical Review D **88**, 062003 (2013), arXiv:1303.6984 [astro-ph.IM].
[42] C. Messick, K. Blackburn, P. Brady, P. Brockill, K. Cannon, R. Cariou, S. Caudill, S. J. Chamberlin, J. D. Creighton, R. Everett, *et al.*, Physical Review D **95**, 042001 (2017), arXiv:1604.04324 [astro-ph.IM].
[43] 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.*, Classical and Quantum Gravity **33**, 215004 (2016), arXiv:1508.02357 [gr-qc].
[44] B. P. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya, *et al.*, Physical Review Letters **119**, 161101 (2017), arXiv:1710.05832 [gr-qc].
[45] N. J. Cornish and T. B. Littenberg, Classical and Quantum Gravity **32**, 135012 (2015), arXiv:1410.3835 [gr-qc].
[46] T. B. Littenberg and N. J. Cornish, Phys. Rev. D **91**, 084034 (2015), arXiv:1410.3852 [gr-qc].
[47] D. Macleod, A. L. Urban, S. Coughlin, T. Massinger, M. Pitkin, paulaltin, J. Areeda, E. Quintero, T. G. Badger, L. Singer, and K. Leinweber, gwpy.