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: December 2, 2020)

The detection of gravitational waves from compact binary coalescence by Advanced LIGO and Advanced Virgo provides an opportunity to study the strong-field, highly-relativistic regime of gravity. Gravitational-wave tests of General Relativity (GR) typically assume Gaussian and stationary detector noise, thus do not account for non-Gaussian, transient noise features (glitches). We present the results obtained by performing parameterized gravitational-wave tests on simulated signals from binary-black-hole coalescence overlapped with a scattered-light glitch. We then review and apply three glitch mitigation methods and evaluate their effect on reducing false deviations from GR. We show that a scattered-light glitch has negligible effect on parameterized test of GR for a three-detector observation, and a bias resulted from its mitigation can only be identified in one case which a significant portion of high-frequency contribution of the signal is removed.

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 corrections to the Einstein-Hilbert action [5], making GR compatible with standard quantum field theory [1]. One approach of probing the strong-field regime is through the detection of gravitational waves (GWs), which propagates at the speed of light and carries information about its astrophysical origin [6].

Of all strong-field astrophysical events that could be probed using GWs, the *coalescence* of stellar-mass binary black holes (BBHs), which can be schematically divided into *inspiral*, *merger* and *ringdown* (IMR) stages, plays a crucial role in testing GR [1]. Since the orbital separation of BBHs can reach far below the last stable orbit before merging, the generated gravitational field can be many order of magnitudes stronger than other astrophysical events observed so far [7–9]. Moreover, GWs emitted by coalescing BBHs offers one of the cleanest test of GR, as matter and electromagnetic fields are negligible for most sources [8, 10], and the emitted GWs essentially propagate through matter unimpeded [8], enabling precision tests of the strong-field dynamics of GR. Since 2015, Advanced LIGO [11] and Advanced Virgo [12] have jointly announced over 40 confident detections of GWs from coalescing BBHs [13, 14].

Several GW tests of GR using coalescing BBHs are developed to test for *generic* deviations from GR without the need for signal models from competing theories of gravity [8]. For example, consistency tests search for excess power in the residual noise after subtracting a best-fit GR waveform [15], or compare the source parameters inferred using only high-frequency data to that inferred using only low-frequency data [15]; parameterized tests introduce parameterized deformations to waveform approximations to GR and infer the extent of deviation using Bayesian parameter estimation [9]. To this date, no evidence for violations of GR has been identified using GWs emitted by coalescing BBHs [16, 17].

Aside from GWs, output from GW detectors is attributed to many independent sources of random noise [18]. Detector noise is typically modeled as stationary and Gaussian in GW data analysis in light of the central limit theorem, and by assuming that noise characteristics remain stationary over observation timescales [19, 20]. However, these assumptions cannot account for transient, non-Gaussian noise features, commonly referred to as *glitches* [21–23], which pose significant problems to GW searches [22] and may bias GW data analysis by violating the noise model. Three glitches from commonly-seen glitch classes during the O3 observing run are shown in Fig. 1.

Many efforts are made to identify and classify glitches [22, 24–28]. Once a glitch is identified, the data containing the glitch can be removed using various mitigation methods [29–33]. The effects of glitches and their mitigations to the inference of source parameters have been studied in the context of glitches similar to that affecting GW170817 [34]. It is of interest to extend the study to parameterized tests of GR, which share the same noise model and parameter estimation techniques but involving extra degree(s) of freedom as parameterized deviations from GR are introduced to the signal model, which

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

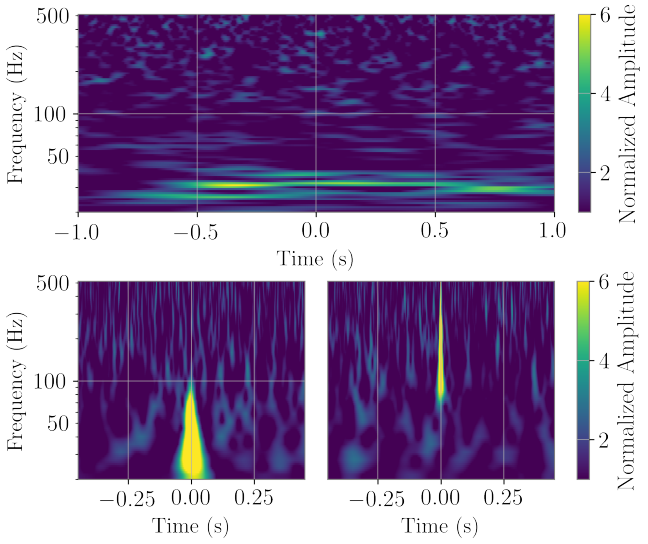


FIG. 1. Glitches with similar morphology can be algorithmically categorized into different classes [22]. A time-frequency representation, called a *Q-scan* (or *Omega scan*) [35], where the duration of each time-frequency bins varies inversely with frequency and linearly with a parameter Q, is commonly used to visualize glitches [22, 28]. Q-scans of three frequently-occurring glitches (top: scattered-light, bottom-left: tomte, bottom-right: blip) during the O3 observing run are shown. The value of Q used is 40, 8 and 8 respectively. The colour represents the normalized amplitude (square root of the normalized power) in each time-frequency bin.

may enhance such effects.

This report is structured as follows: Sec. II describes the typical data model used in GW data analyses [19, 20], which composes of a GW signal in additive stationary and Gaussian noise. Sec. III introduces a parameterized test of GR involving the parameterization of the phase of an IMR waveform model [36]. Sec. IV reviews three glitch mitigation methods, namely band-pass filtering, gating and inpainting, and discusses their potential impacts on tests of GR. Sec. V presents the results obtained by performing the parameterized test of GR to glitch-overlapped BBH-coalescence GW signals before and after glitch mitigations.

II. DATA MODEL

A GW detector is designed to respond linearly to the fractional change in arm length, or *strain* [18]. The time series of detector output data $\boldsymbol{d}$, sampled at time t_k at constant sampling interval Δt, can thus be expressed as a linear superposition of a time series of the GW strain signal $\boldsymbol{h}$ and a time series of detector noise $\boldsymbol{n}$:

$$\boldsymbol{d}(t_k) = \boldsymbol{h}(t_k) + \boldsymbol{n}(t_k) \ . \tag{1}$$

In Eq. (1) and in subsequent discussion, boldface denotes the matrix representation of specified quantities.

A. Stationary Gaussian Noise Model

Assuming that a *large* number of independent noise sources contribute linearly to the detector noise $\boldsymbol{n}$, the central limit theorem states that the probability density distribution of the noise $\boldsymbol{n}$ tends to follow a multivariate *Gaussian* distribution [37]:

$$P(\boldsymbol{n}) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\boldsymbol{n}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{n}-\boldsymbol{\mu})} \ , \tag{2}$$

which is uniquely defined by the *covariance matrix* $\Sigma_{ij} = E[(n(t_i) - \mu(t_i))(n(t_j) - \mu(t_j))]$ and the mean vector $\mu_i = E[n(t_i)]$, where $E[\cdot]$ and $|\cdot|$ denotes the expectation and determinant operation respectively. The diagonal (off-diagonal) terms of the covariance matrix are the variances at each instance of time (correlations between data from different instances of time).

If the number of samples N is large, it is undesirable to invert the $N \times N$ covariance matrix in Eq. (2). Instead, we consider the joint probability density in Fourier domain, which is also a multivariate Gaussian distribution [37]. With the assumption of stationarity, i.e. the joint probability density distribution is time-invariant, the covariance matrix in Fourier domain is diagonalized in the infinite-duration limit [38]. This relation can be approximated for the finite-duration discretely-sampled time series, giving the following approximation to the joint probability density in Fourier domain [38] (for even N), also known as the *Whittle likelihood* [39] in the context of statistical inference:

$$P(\boldsymbol{n}) \simeq \prod_{j=0}^{N/2-1} \frac{2\Delta f}{\pi S_n(f_j)} \exp\left(-\Delta f \frac{2|\tilde{n}_j|^2}{S_n(f_j)}\right) \ , \tag{3}$$

where $f_j \equiv j/N\Delta t$. The quantity $S_n(f_j) \equiv 2|\tilde{n}(f_j)|^2/T$ is scaled from the diagonal terms of the covariance matrix in Fourier domain, $\Delta f \equiv 1/T$ is the *frequency resolution* and the tilde denotes a discrete Fourier transformed (DFT) quantity:

$$\tilde{n}_j \equiv \Delta t \ \mathrm{DFT}[n(t_k)] = \Delta t \sum_{k=0}^{N-1} n(t_k) e^{-2\pi i j k/N} \ . \tag{4}$$

To motivate the quantity $S_n(f_j)$, called the *one-sided power spectral density* (PSD), we invoke Parseval's theorem [38]:

$$\sum_{j=0}^{N/2-1} S_n(f_j)\Delta f \equiv \frac{2}{T} \sum_{j=0}^{N/2-1} |\tilde{n}(f_j)|^2 \Delta f = \frac{1}{N} \sum_{k=0}^{N-1} |n(t_k)|^2 \ , \tag{5}$$

and note that the rightmost side of Eq. (5) returns the *power* of the time series. Since a time series is real, we have the property $\tilde{n}(f_j) = \tilde{n}^*(-f_j)$. Consequently, we can sample only the frequency bins from 0 Hz to up to the *Nyquist frequency* $1/2\Delta t$, and introduce the factor of 2 in Eq. (3) and Eq. (5).

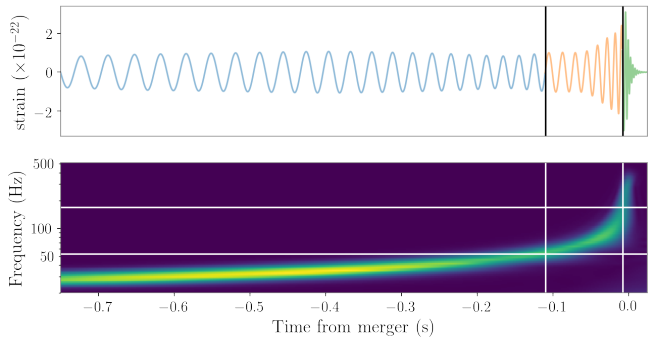


FIG. 2. An example `IMRPhenomPv2` time-domain GW waveform (upper figure) and the corresponding instantaneous frequency (lower figure) plotted against time. The two horizontal lines in the lower figure correspond to the frequencies $0.018/M$ (lower line) and $f_{\rm RD}/2$ (upper line), which defines the boundaries of the inspiral, intermediate and merger-ringdown stages in Fourier domain. The corresponding boundaries in time domain (vertical lines) are determined as the times when the instantaneous frequency of the signal intersects with the Fourier-domain boundaries.

B. Signal Model

Since the two-body self-gravitating problem cannot be solved analytically in GR, we generate simulated GW strain signals from coalescing BBHs using the frequency-domain precessing IMR waveform model `IMRPhenomPv2` [36] in virtue of its good match with Numerical Relativity (NR) waveforms [40] and low computational costs.

`IMRPhenomPv2` is a phenomenological waveform model constructed by combining PN-like inspiral waveforms with NR-calibrated merger-ringdown ansatz [41]. Its *inspiral* stage is modeled up to $f \sim 0.018/M$ (in natural units), where M is the total mass of the system. The region with $Mf \geq 0.018$ is subdivided into an *intermediate* stage with $0.018 \geq Mf \geq 0.5 f_{\rm RD}$, which bridges the inspiral stage to the *merger-ringdown* stage modeled above half the ringdown frequency $f_{\rm RD}$ [41]. Fig. 2 illustrates the stages of coalescence of an example `IMRPhenomPv2` GW strain and its frequency evolution over time.

The phase of `IMRPhenomPv2` 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 and the orbital evolution of the BBH depend only on the masses and spin angular momentum vectors of the component black holes [40], denoted as the *intrinsic* parameters. 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 [9, 41]: (i) the *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 $\{\alpha_0, ..., \alpha_5\}$.

TABLE I. The frequency dependence of `IMRPhenomPv2` testing parameters used in parameterized tests of GR. The table is reproduced from Table 1 of Ref. [15].

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}
	$\delta\alpha_3$	$f^{3/4}$
	$\delta\alpha_4$	$\tan^{-1}(af+b)$

Seven additional *extrinsic* parameters, including the sky location, luminosity distance, polarization angle of the source, and the spatial orientation and orbital phase of the system at a reference frequency and time respectively, are needed to determine the response of the GW detectors.

III. PARAMETERIZED TESTS OF GR

We will focus on a parameterized test of GR, which introduces *fractional* deviations δp_i, also known as *dephasing coefficients*, to `IMRPhenomPv2` phase coefficients p_i [9]:

$$p_i \mapsto p_i[1 + \delta p_i] \,. \tag{6}$$

For the exceptional case where $p_i = 0$, such as φ_1, an *absolute* deviation is instead introduced [9]. In practice, we do not allow some of the `IMRPhenomPv2` phase coefficients to deviate from their prescribed values as they have large uncertainties or are degenerate with with other coefficients or physical parameters [9]. We therefore perform tests with the remaining 14 dephasing coefficients, henceforth denoted as the *testing* dephasing coefficients [9]:

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

The frequency dependence of the testing parameters δp_i is shown in Table I [15, 42].

To quantify a deviation from GR, we can infer the most probable values of δp_i through Bayesian parameter estimation, as discussed in the following subsection.

A. Parameter Estimation

Recall our data model $\boldsymbol{d} = \boldsymbol{h} + \boldsymbol{n}$. Introducing parameterized phase deviations to the signal $\boldsymbol{h}$, we denote $\boldsymbol{\theta}(\theta, \delta p_i)$ as the parameter vector generating the signal, which consists of the intrinsic and extrinsic parameters generating the `IMRPhenomPv2` waveform, θ, and the testing parameters δp_i. In practice, the testing parameters are introduced *once at a time*, which is expected to capture a deviation from GR present in multiple phase coefficients, while returning narrower credible intervals [15].

Given the detector output $\boldsymbol{d}$ and prior information I, we wish to infer the conditional probability density of $\boldsymbol{\theta}$, referred to as the *posterior*, by invoking Bayes' theorem

$$P(\boldsymbol{\theta}|\boldsymbol{d}, I) = \frac{P(\boldsymbol{d}|\boldsymbol{\theta}, I) \times P(\boldsymbol{\theta}|I)}{P(\boldsymbol{d}|I)} \ , \tag{7}$$

which relates the posterior to three probability densities: the *likelihood* $P(\boldsymbol{d}|\boldsymbol{\theta}, I)$, the *prior* $P(\boldsymbol{\theta}|I)$ and the *evidence* $P(\boldsymbol{d}|I)$. During parameter estimation, the evidence, which do not depend explicitly on $\boldsymbol{\theta}$, can be seen as a proportionality constant since $\boldsymbol{d}$ and I are kept fixed. The likelihood and prior is separately discussed below.

Given $\boldsymbol{h}(\boldsymbol{\theta})$, the time series of the output data $\boldsymbol{d}$ uniquely defines a time series of the residual noise $\boldsymbol{d} - \boldsymbol{h}$, which is modeled as Gaussian and stationary. As such, the likelihood is approximated by the Whittle likelihood in Eq. (3):

$$P(\boldsymbol{d}|\boldsymbol{\theta}, I) \propto \exp\left[-\frac{1}{2}(\boldsymbol{d} - \boldsymbol{h}|\boldsymbol{d} - \boldsymbol{h})\right] \ , \tag{8}$$

where $(\cdot|\cdot)$ is the *noise-weighted inner product* [43]:

$$(\mathbf{a}|\mathbf{b}) \equiv \sum_{j=0}^{N/2-1} 4\Re\left(\frac{\tilde{a}_j^* \tilde{b}_j}{S_n(f_j)}\right)\Delta f \ . \tag{9}$$

Assuming that noise from multiple detectors, indexed l, are uncorrelated, the joint likelihood takes the form

$$P(\boldsymbol{d}_l|\boldsymbol{\theta}, I) \propto \exp\left[-\frac{1}{2}\sum_l (\boldsymbol{d}_l - \boldsymbol{h}_l|\boldsymbol{d}_l - \boldsymbol{h}_l)\right] \ . \tag{10}$$

The prior $P(\boldsymbol{\theta}|I)$ incorporates our beliefs about $\boldsymbol{\theta}$ prior to the observation. We follow the default choice of prior in `LALInference` [20], which include uniform priors for the component masses m_1 and m_2, with $m_2 \leq m_1$, a log-uniform prior for the luminosity distance, an isotropic prior for the sky location of the source and the spin angular momentum vectors of the component black holes, and uniform priors for the remaining parameters. We note that in `LALInference`, the uniform priors specified for component masses are transformed to non-uniform, correlated priors for the chirp mass $\mathcal{M} \equiv (m_1 m_2)^{3/5}(m_1 + m_2)^{-1/5}$ and the mass ratio $q \equiv m_2/m_1$ for more efficient sampling [20].

In parameterized tests of GR, parameters of primary interest are the testing parameters δp_i, while the posterior distribution spans the full 16-dimensional parameter space. We therefore compute the *marginalized* posterior distribution for introduced the testing parameter δp_i:

$$P(\delta p_i|\boldsymbol{d}, I) = \int P(\boldsymbol{\theta}|\boldsymbol{d}, I) d\theta \ , \tag{11}$$

where θ denotes the parameters generating the underlying `IMRPhenomPv2` waveform.

IV. GLITCH MITIGATION METHODS

In this section, we introduce three mitigation methods to disregard or remove data containing glitches. In the context of analyzing glitch-overlapped GW signals from coalescing BBHs, a preferred mitigation method should meet two additional requirements: 1) the mitigation itself should not bias the noise-weighted inner product $(\boldsymbol{d} - \boldsymbol{h}|\boldsymbol{d} - \boldsymbol{h})$ in the likelihood calculation, which consequently bias parameterized tests of GR, and 2) since GW signals from coalescing BBHs are typically observed at sub-second timescales, the mitigation should remove the minimum possible amount of data.

Preferred mitigation methods would differ from case to case depending on the morphology of glitches and their location within the analysis segment. For example, disregarding affected frequency bins is not preferred for mitigating a broadband glitch, whereas removing affected time bins is not preferred for mitigating a long-duration glitch. Owing to the unpredictable morphology and occurrence rates of glitches in future observing runs, we do not make the attempt to identify the most preferred mitigation method for specific classes of glitches in this study.

In our study, we will separately apply the three mitigation methods of band-pass filtering (configured in `LALInference`), gating with an inverse Tukey window and inpainting (both implemented in `PyCBC` [44]) to glitch-overlapped data samples.

A. Band-pass filtering

Assuming stationary and Gaussian noise, components of the noise-weighted inner product from different frequency bins of equal bandwidth and from different detectors contribute linearly to the log likelihood, as seen from Eq. (10). A direct way of removing the glitch in Fourier domain is by excluding the frequency bins containing the glitch from the likelihood calculation. In `LALInference`, this can be done by specifying the high-pass and low-pass cutoff frequency for the affected detector such that data containing the glitch is filtered out. Only the passed frequency bins are considered in the likelihood calculation. By default, data is high-passed at 20 Hz in `LALInference` [20].

Although the frequency bins that contribute to the noise-weighted inner product are essentially unaffected by the band-pass filter, bias may be resulted if a large portion of signal from the most sensitive bands, which contributes significantly to inference, is disregarded. If the glitch is localized in frequency, the method of band-pass filtering affects the minimal amount of data.

B. Gating in Time Domain

A similar procedure can be done in the time domain, commonly known as *gating*, in which data containing the glitch is zeroed out by multiplying an inverse window function. The inverse window function reduces the spectral leakage in Fourier domain due to discontinuity of data at the boundary of the region to be zeroed out [45].

The inverse *Tukey*, or *cosine-tapered*, window [46] is commonly-used in GW searches [29, 30]. It is adopted in the mitigation of glitch-overlapped GW170817 signal in LIGO-Livingston during the rapid localization of the source [47], illustrated in Fig. 3, which successfully led to follow-up electromagnetic observations [48]. A less common choice of window for gating, implemented in `GWpy` [49], is the inverse *Planck* window [50], which is also a considerable choice of windowing as the same reduction of spectral leakage can be achieved with a shorter window duration than an inverse Tukey window, if implemented correctly [50].

An inverse Tukey (Planck) window consists of a flat "bottom" with values of zero, which is smoothly tapered to unity at both ends using a cosine function (Planck-distribution-like function). The parameter α denotes the ratio of the duration of tapering to the total window duration, which are both adjustable for both windows.

As remarked in Ref. [34], gating can introduce errors to parameterized tests of GR, as it affects the signal power in frequency bins that count towards the noise-weighed inner product. Gating also produces long-duration noise artifacts at high frequencies (of around 500 Hz) outside the gated interval, which may overlap the ringdown stage in frequency for lower-mass BBHs. On the other hand, the duration of data affected by fully removing a glitch exceeds that by the glitch by at least a factor of $(1-\alpha)^{-1}$. The specification of α can only be subjectively chosen to "optimally" balance the reduction of spectral leakage, which increases with α, and the duration affected, which decreases with α. For short-duration glitches, the minimal duration of window is further limited by the spectral leakage in time and frequency domain by low-frequency lines in the PSD, which would be resulted if the window duration is shorter than the inverse width of the spectral line [33], producing high-amplitude glitch-like noise artifacts at the boundaries in time [33].

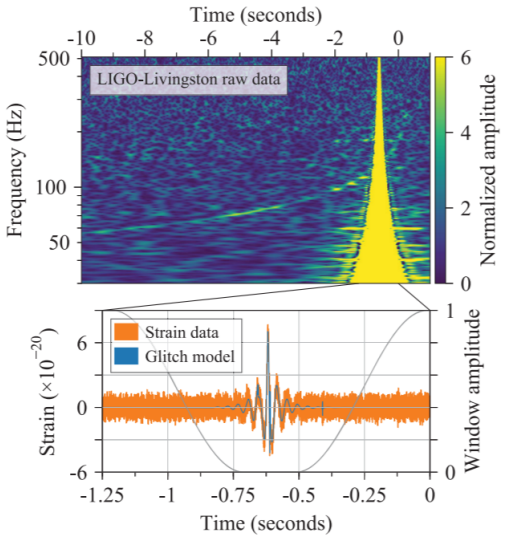


FIG. 3. The output data from LIGO-Livingston for GW170817 is plotted over time in the bottom figure (orange curve). A glitch was identified around the time $t = -0.75$ s to -0.5 s in the figure. To infer the sky location of the event during rapid localization, data was multiplied by an inverse Tukey window function (black curve) [47]. To infer the source properties during parameter estimation, a glitch model (blue curve) reconstructed with `BayesWave` [31, 32] is subtracted from the data [47], which is currently not included in our study. The upper figure shows a spectrogram of the raw LIGO-Livingston data. The figure is retrieved from Abbott et al. [47]

C. Inpainting in Time Domain

A new method, called *impainting* or *hole filling* [33], is developed to address the noise artifacts and statistical bias that may be resulted from gating. After specifying the time interval to be mitigated, new values are assigned for data within the interval, or *hole*, according to an *inpainting filter*, while data outside the hole are unaffected. The inpainting filter depends on the PSD of the stationary Gaussian noise. Inpainted data within the hole is identically zero upon twice-whitening by the same PSD, and the quantity $(\boldsymbol{d}|\boldsymbol{h})$ is independent of the template waveform $\boldsymbol{h}$ inside the interval [33]. Since the hole can be made arbitrarily narrow, inpainting affect the minimal amount of data if the glitch is localized in time.

If the PSD used in the inpainting filter equals the PSD estimation used in the likelihood calculation, inpainting is not expected to bias parameterized tests of GR: re-expressing the noise-weighed inner product in the likelihood calculation:

$$P(\boldsymbol{d}|\boldsymbol{h}) \propto \exp\left[-\frac{1}{2}(\boldsymbol{d}-\boldsymbol{h}|\boldsymbol{d}-\boldsymbol{h})\right] = \exp\left[-\frac{1}{2}(\boldsymbol{d}|\boldsymbol{d}) + (\boldsymbol{d}|\boldsymbol{h}) - \frac{1}{2}(\boldsymbol{h}|\boldsymbol{h})\right] . \quad (12)$$

Given inpainted data $\boldsymbol{d}$, only the terms $(\boldsymbol{d}|\boldsymbol{h})$ and $(\boldsymbol{h}|\boldsymbol{h})$

differ across waveform templates $\boldsymbol{h}$; between these two terms, only $(\boldsymbol{d}|\boldsymbol{h})$ explicitly depends on the inpainted data. As $(\boldsymbol{d}|\boldsymbol{h})$ is independent of the template waveform inside the hole, inpainted data inside the hole is not expected to affect parameterized tests. Again, bias may be introduced if a significant portion of signal is removed by inpainting. The desirable behaviors of the inapinting filter may not hold if different estimates of the PSD are used in the inpainting filter and the likelihood calculation. The sensitivity of the inpainting filter towards the PSD deserves a study [51].

V. RESULTS

A. Scattered-light Glitch overlapped with a GW190828_065509-like Signal

We are motivated to consider a signal similar to that of the high-mass-ratio BBH-merger event GW190828_065509 [14], in which the mitigation of potential glitches overlapping the event in L1 through band-pass filtering resulted in pathological features in parameterized tests of GR [52]. Values of some selected generating parameters of the GW190828_065509-like signal is tabulated in Table II.

TABLE II. Injected values of some selected generating parameters of a GW190828_065509-like signal using the `IMRPhenomPv2` waveform model. The GW190828_065509-like signal is taken to be the *maximum likelihood* waveform inferred for real GW190828_065509 data using the `IMRPhenomPv2` template waveform model. Due to the high mass ratio and strong spins of the GW190828_065509-like signal, precession effects are significant.

Waveform Parameter	Value
Chirp mass $\mathcal{M}$ ($M_\odot$)	16.86
Mass ratio q	0.14
Dimensionless primary spin magnitude a_1	0.92
Dimensionless secondary spin magnitude a_2	0.75
Right ascension α (rad)	2.54
Declination δ (rad)	-0.84

The simulated GW190828_065509-like signal, generated with a `IMRPhenomPv2` waveform model, is then injected coherently into real Hanford (H1), Livingston (L1) and Virgo (V1) detector data [53] at the time where a scattered-light glitch is present in H1. The glitch lasts for a long duration of ~ 2 seconds, and affect frequency bins lower than 40 Hz. The duration and peak frequency are representative of scattered-light glitches in the O3 observing run [22]. In addition, H1, L1 and V1 are all operating in the science mode when the glitch occurred. The top Q-scan of Fig. 1 shows a representation of the scattered-light glitch selected for our study. The merger time of the simulated signal is slightly adjusted in between injections so that the glitch overlaps the inspiral, inspiral-intermediate and the intermediate-merger-ringdown stages in time domain in between data samples. We then performed parameterized test of GR on the three data samples. The posteriors of the testing parameters for the three cases are plotted in on the left of each violin plot in Fig. 4. In all three cases, we found no evidence in support of false violations of GR due to the effect of the scattered-light glitch. GR ($\delta p_i = 0$) is not excluded from the posterior of all testing parameters except $\delta\alpha_4$ for the case of scattered-light-glitch-overlapped intermediate-merger-ringdown in time.

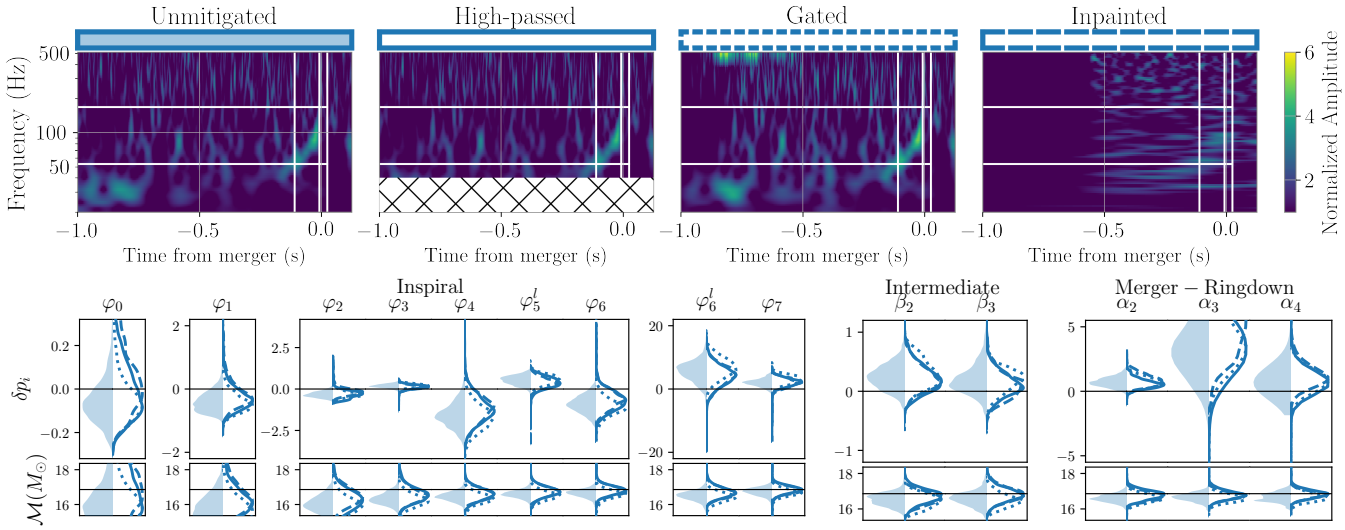
Next, we applied band-pass filtering, gating and inpainting discussed in Sec. IV on the three glitch-overlapped data samples and performed parameterized test of GR on the mitigated samples. The posterior distributions of the testing parameters for the unmitigated case (left) and the mitigated cases (right) are presented as violin plots in Fig. 4. In most cases, mitigating the scattered-light glitch with different independent methods do not lead to statistically significant changes in the posterior of most testing parameters. This suggests that this particular glitch has negligible effect on parameterized tests of GR, irrespective of the stage of coalescence it overlaps in time domain. We get an exception when the glitch-overlapped intermediate-merger-ringdown stage in time domain is inpainted, which result in biased posteriors for the merger-ringdown testing parameters. A closer inspection is warranted in order to identify the cause of the bias. This can be done by repeating the study using data from two detectors (H1, V1) instead of three detectors (H1, L1, V1) and observe whether a more significant bias would be resulted. Another bias can be observed in $\delta\beta_2$ when the same data sample is high-passed or gated. The cause(s) of this bias is unidentified.

Aside from posterior distributions of the testing parameters, the recovered posteriors of chirp mass by parameterized tests of GR are also presented in Fig. 4. The injected chirp mass can be accurately recovered in most cases, and less accurately recovered when parameterized deviations are introduced to PN testing parameters when the glitch overlaps the intermediate-merger-ringdown stage in time. The match between posteriors of chirp mass for the unmitigated and mitigated cases again suggests that the accuracy of chirp mass recovery do not depend on the scattered-light glitch itself.

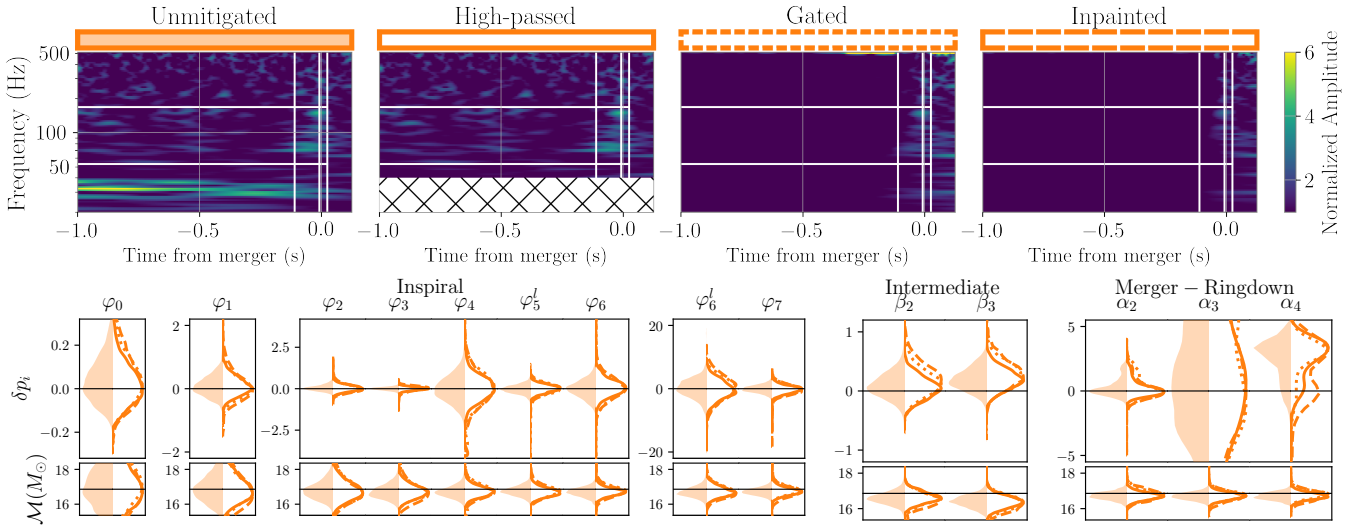
VI. CONCLUSION AND OUTLOOK

We injected a high-mass-ratio coalescing BBH signal coherently into the LIGO-Hanford, LIGO-Livingston and Virgo detector output at a time when a scattered-light glitch was present in LIGO-Hanford. The merger time of the injected signal is adjusted slightly so that the glitch overlaps with different stages of coalescence of the signal.

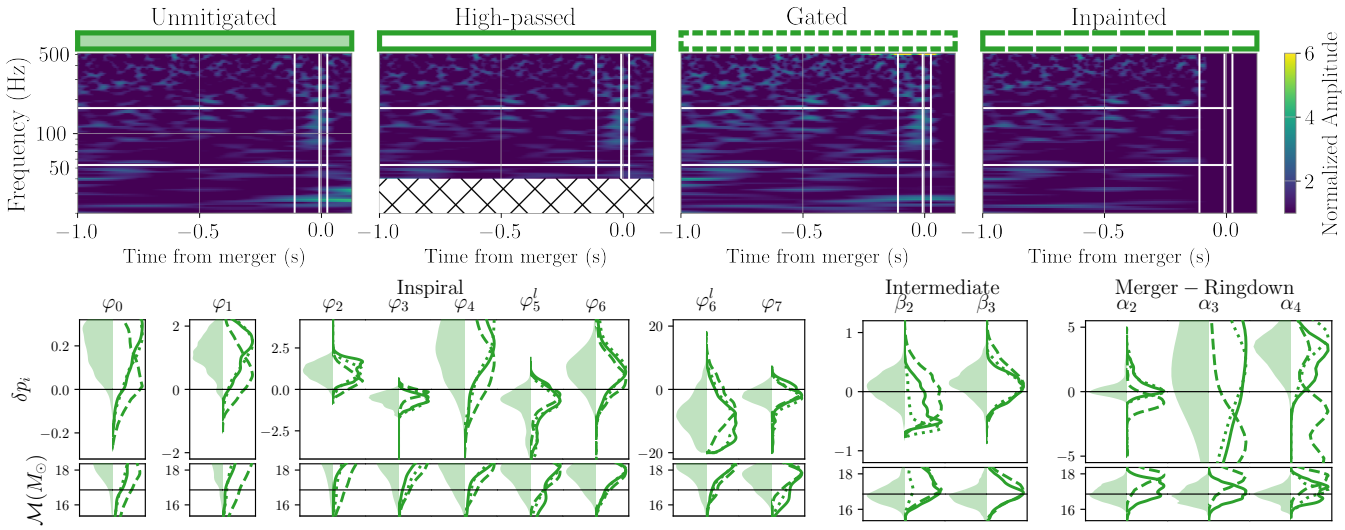
We then investigated the effect of the scattered-light



(a) Scattered-light-glitch-overlapped GW190828_065509-like signal at inspiral stage in time domain.



(b) Scattered-light-glitch-overlapped GW190828_065509-like signal at inspiral-intermediate stage in time domain.



(c) Scattered-light-glitch-overlapped GW190828_065509-like signal at intermediate-merger-ringdown stage in time domain.

FIG. 4. Posterior distributions of testing parameters and the recovered chirp mass obtained by performing parameterized tests on unmitigated glitch-overlapped cases during a three-detector observation (left of violin plot) and the corresponding mitigated cases (right of violin plot) where band-pass filtering (solid line), gating (dotted line) and inpainting (dashed line) are performed. The GR-value of the testing parameters and the injected value of chirp mass are indicated by vertical black lines.

glitch and three mitigation methods, namely band-pass filtering, gating and inpainting, on a parameterized test of GR. We conclude that the scattered-light glitch do not result in a false violation of GR. Moreover, the glitch and the three mitigations have negligible effect on parameterized test of GR. It is speculated that if a significant proportion of signal from the network of detectors is removed, either in frequency or time domain, bias may be resulted in parameterized test of GR. Work is in progress to verify this speculation. Repeating this study on GW190701, which was similarly overlapped with a scattered-light glitch in LIGO-Livingston [14], will provide us with a reality check on the implications drawn above.

The investigation described in this report will be reproduced for broadband glitches and their mitigations, which can lead to distinctly different results for parameterized tests of GR. Their effect on a two-detector and a three-detector observation will likewise be studied and compared. Two short-duration, broadband glitches from the *tomte* and *blip* class, which overlap the inspiral-intermediate and intermediate-merger-ringdown frequency bands respectively, are chosen for further investigation. The bottom two Q-scans in Fig. 1 shows the representations of the tomte and blip glitches respectively.

A major improvement on the LIGO detectors are expected to be completed in a few years, doubling the detector sensitivities [54]. The increased sensitivity in turn suggests more frequent occurrence of glitches overlapping signals. As mitigating glitch-overlapped signals may become a regularity in the future, a systematic study on the effects of glitches and their mitigation on parameter estimation and tests of GR will be crucial to the next generation of GW astronomy; this will be left for future work.

VII. ACKNOWLEDGMENTS

The author thanks Alan J. Weinstein and Rico K. L. Lo for providing instructive comments and reviewing the manuscript, Tjonnie G. F. Li for preparations before the LIGO SURF program, Derek Davis and Jonah Kanner for insightful discussions, Tyson Littenberg for providing the `BayesWave` glitch subtraction specifications. The author also thanks the National Science Foundation (NSF) and NSF REU for supporting the LIGO SURF program. Computing resources for this study was supported by the LIGO Laboratory and supported by the NSF Grants PHY-0757058 and PHY-0823459. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the NSF and operates under cooperative agreement PHY-0757058. The LIGO SURF Program is supported by NSF award PHY-1852081. This report carries LIGO Document Number LIGO-T2000374.

[1] E. Berti, E. Barausse, V. Cardoso, L. Gualtieri, P. Pani, U. Sperhake, L. C. Stein, N. Wex, K. Yagi, T. Baker, *et al.*, *Testing general relativity with present and future astrophysical observations*, 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, *The Confrontation between General Relativity and Experiment*, Living Reviews in Relativity **9**, 3 (2006), arXiv:gr-qc/0510072 [gr-qc].

[4] C. M. Will, *The confrontation between general relativity and experiment*, Living reviews in relativity **17**, 4 (2014), arXiv:1403.7377 [gr-qc].

[5] K. Stelle, *Renormalization of higher-derivative quantum gravity*, Physical Review D **16**, 953 (1977).

[6] B. S. Sathyaprakash and B. F. Schutz, *Physics, Astrophysics and Cosmology with Gravitational Waves*, Living Reviews in Relativity **12**, 2 (2009), arXiv:0903.0338 [gr-qc].

[7] N. Yunes and X. Siemens, *Gravitational-wave tests of general relativity with ground-based detectors and pulsar-timing arrays*, Living Reviews in Relativity **16**, 9 (2013), arXiv:1304.3474 [gr-qc].

[8] N. Yunes, K. Yagi, and F. Pretorius, *Theoretical physics implications of the binary black-hole mergers gw150914 and gw151226*, Physical review D **94**, 084002 (2016), arXiv:1603.08955 [gr-qc].

[9] J. Meidam, K. W. Tsang, J. Goldstein, M. Agathos, A. Ghosh, C.-J. Haster, V. Raymond, A. Samajdar, P. Schmidt, R. Smith, *et al.*, *Parametrized tests of the strong-field dynamics of general relativity using gravitational wave signals from coalescing binary black holes: Fast likelihood calculations and sensitivity of the method*, Physical Review D **97**, 044033 (2018), arXiv:1712.08772 [gr-qc].

[10] E. Barausse, V. Cardoso, and P. Pani, *Can environmental effects spoil precision gravitational-wave astrophysics?*, Physical Review D **89**, 104059 (2014), arXiv:1404.7149 [gr-qc].

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

[12] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin, *et al.*, *Advanced virgo: a second-generation interferometric gravitational wave detector*, Classical and Quantum Gravity **32**, 024001 (2014), arXiv:1408.3978 [gr-qc].

[13] 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*, Physical Review X **9**, 031040 (2019), arXiv:1811.12907 [astro-

ph.HE].
[14] R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. Adhikari, V. Adya, C. Affeldt, *et al.*, *Gwtc-2: Compact binary coalescences observed by ligo and virgo during the first half of the third observing run*, arXiv preprint arXiv:2010.14527 (2020), arXiv:2010.14527 [gr-qc].
[15] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, *et al.*, *Tests of General Relativity with GW150914*, Phys. Rev. Lett. **116**, 221101 (2016), arXiv:1602.03841 [gr-qc].
[16] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, *et al.*, *Tests of general relativity with the binary black hole signals from the ligo-virgo catalog gwtc-1*, Physical Review D **100**, 104036 (2019), arXiv:1903.04467 [gr-qc].
[17] R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. Adhikari, V. Adya, C. Affeldt, *et al.*, *Tests of general relativity with binary black holes from the second ligo-virgo gravitational-wave transient catalog*, arXiv preprint arXiv:2010.14529 (2020), arXiv:2010.14529 [gr-qc].
[18] P. R. Saulson, *Fundamentals of Interferometric Gravitational Wave Detectors* (2017).
[19] LIGO Scientific Collaboration, LIGO Algorithm Library - LALSuite, free software (GPL) (2018).
[20] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, *et al.*, *Parameter estimation for compact binaries with ground-based gravitational-wave observations using the lalinference software library*, Physical Review D **91**, 042003 (2015), arXiv:1409.7215 [gr-qc].
[21] L. K. Nuttall, T. Massinger, J. Areeda, J. Betzwieser, S. Dwyer, A. Effler, R. Fisher, P. Fritschel, J. Kissel, A. Lundgren, *et al.*, *Improving the data quality of advanced ligo based on early engineering run results*, Classical and Quantum Gravity **32**, 245005 (2015), arXiv:1508.07316 [gr-qc].
[22] M. Zevin, S. Coughlin, S. Bahaadini, E. Besler, N. Rohani, S. Allen, M. Cabero, K. Crowston, A. K. Katsaggelos, S. L. Larson, *et al.*, *Gravity spy: integrating advanced ligo detector characterization, machine learning, and citizen science*, Classical and Quantum Gravity **34**, 064003 (2017), arXiv:1611.04596 [gr-qc].
[23] B. P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, M. Adamo, C. Adams, T. Adams, P. Addesso, *et al.*, *Characterization of transient noise in advanced ligo relevant to gravitational wave signal gw150914*, Classical and Quantum Gravity **33**, 134001 (2016), arXiv:1602.03844 [gr-qc].
[24] J. R. Smith, T. Abbott, E. Hirose, N. Leroy, D. MacLeod, J. McIver, P. Saulson, and P. Shawhan, *A hierarchical method for vetoing noise transients in gravitational-wave detectors*, Classical and Quantum Gravity **28**, 235005 (2011), arXiv:1107.2948 [gr-qc].
[25] T. Isogai, L. S. Collaboration, V. Collaboration, *et al.*, in *Journal of Physics: Conference Series*, Vol. 243 (IOP Publishing, 2010) p. 012005.
[26] R. Essick, L. Blackburn, and E. Katsavounidis, *Optimizing vetoes for gravitational-wave transient searches*, Classical and Quantum Gravity **30**, 155010 (2013), arXiv:1303.7159 [astro-ph.IM].
[27] 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.*, *Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data*, Physical Review D **88**, 062003 (2013), arXiv:1303.6984 [astro-ph.IM].
[28] F. Robinet, N. Arnaud, N. Leroy, A. Lundgren, D. Macleod, and J. McIver, *Omicron: a tool to characterize transient noise in gravitational-wave detectors*, arXiv preprint arXiv:2007.11374 (2020), arXiv:2007.11374 [astro-ph.IM].
[29] C. Messick, K. Blackburn, P. Brady, P. Brockill, K. Cannon, R. Cariou, S. Caudill, S. J. Chamberlin, J. D. Creighton, R. Everett, *et al.*, *Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data*, Physical Review D **95**, 042001 (2017), arXiv:1604.04324 [astro-ph.IM].
[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*, Classical and Quantum Gravity **33**, 215004 (2016), arXiv:1508.02357 [gr-qc].
[31] N. J. Cornish and T. B. Littenberg, *Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches*, Classical and Quantum Gravity **32**, 135012 (2015), arXiv:1410.3835 [gr-qc].
[32] T. B. Littenberg and N. J. Cornish, *Bayesian inference for spectral estimation of gravitational wave detector noise*, Phys. Rev. D **91**, 084034 (2015), arXiv:1410.3852 [gr-qc].
[33] B. Zackay, T. Venumadhav, J. Roulet, L. Dai, and M. Zaldarriaga, *Detecting gravitational waves in data with non-gaussian noise*, arXiv preprint (2019), arXiv:1908.05644 [astro-ph.IM].
[34] C. Pankow, K. Chatziioannou, E. A. Chase, T. B. Littenberg, M. Evans, J. McIver, N. J. Cornish, C.-J. Haster, J. Kanner, V. Raymond, *et al.*, *Mitigation of the instrumental noise transient in gravitational-wave data surrounding gw170817*, Physical Review D **98**, 084016 (2018), arXiv:1808.03619 [gr-qc].
[35] S. Chatterji, L. Blackburn, G. Martin, and E. Katsavounidis, *Multiresolution techniques for the detection of gravitational-wave bursts*, Classical and Quantum Gravity **21**, S1809 (2004), arXiv:gr-qc/0412119 [gr-qc].
[36] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, *Simple model of complete precessing black-hole-binary gravitational waveforms*, Physical review letters **113**, 151101 (2014), arXiv:1308.3271 [gr-qc].
[37] 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).
[38] J. D. Romano and N. J. Cornish, *Detection methods for stochastic gravitational-wave backgrounds: a unified treatment*, Living reviews in relativity **20**, 2 (2017), arXiv:1608.06889 [gr-qc].
[39] P. Whittle, *Hypothesis testing in time series analysis*, Vol. 4 (Almqvist & Wiksells boktr., 1951).
[40] S. Khan, K. Chatziioannou, M. Hannam, and F. Ohme, *Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects*, Physical Review D **100**, 024059 (2019), arXiv:1809.10113 [gr-qc].

[41] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, *Frequency-domain gravitational waves from nonprecessing black-hole binaries. ii. a phenomenological model for the advanced detector era*, Physical Review D **93**, 044007 (2016), arXiv:1508.07253 [gr-qc].

[42] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, *Frequency-domain gravitational waves from nonprecessing black-hole binaries. i. new numerical waveforms and anatomy of the signal*, Physical Review D **93**, 044006 (2016), arXiv:1508.07250 [gr-qc].

[43] C. Cutler and E. E. Flanagan, *Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform?*, Physical Review D **49**, 2658 (1994), arXiv:gr-qc/9402014 [gr-qc].

[44] A. Nitz, I. Harry, D. Brown, C. M. Biwer, J. Willis, T. D. Canton, C. Capano, L. Pekowsky, T. Dent, A. R. Williamson, *et al.*, gwastro/pycbc: Pycbc release v1.16.9 (2020).

[45] F. J. Harris, *On the use of windows for harmonic analysis with the discrete fourier transform*, Proceedings of the IEEE **66**, 51 (1978).

[46] J. W. Tukey, *An intoroduction to the calculation of numerical spectrum analysis*, Spectra Analysis of Time Series , 25 (1967).

[47] B. P. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya, *et al.*, *Gw170817: observation of gravitational waves from a binary neutron star inspiral*, Physical Review Letters **119**, 161101 (2017), arXiv:1710.05832 [gr-qc].

[48] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, *et al.*, *Multi-messenger observations of a binary neutron star merger*, The Astrophysical Journal **848**, L12 (2017), arXiv:1710.05833 [astro-ph.HE].

[49] 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.

[50] D. McKechan, C. Robinson, and B. S. Sathyaprakash, *A tapering window for time-domain templates and simulated signals in the detection of gravitational waves from coalescing compact binaries*, Classical and Quantum Gravity **27**, 084020 (2010), arXiv:1003.2939 [gr-qc].

[51] Private communication with D. Davis.

[52] Private communication with Rico K. L. Lo.

[53] Calibrated, cleaned data from H1 and L1 are taken from the strain channel `DCS-CALIB_STRAIN_CLEAN_C01`. Reproduced data from V1 is taken from the strain channel `Hrec_hoft_V1O3Repro1A_16384Hz`.

[54] J. Miller, L. Barsotti, S. Vitale, P. Fritschel, M. Evans, and D. Sigg, *Prospects for doubling the range of advanced ligo*, Physical Review D **91**, 062005 (2015), arXiv:1410.5882 [gr-qc].