

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Not quite black holes at LIGO

Bob Holdom

Phys. Rev. D **101**, 064063 — Published 26 March 2020

DOI: [10.1103/PhysRevD.101.064063](https://doi.org/10.1103/PhysRevD.101.064063)

# Not quite black holes at LIGO

Bob Holdom\*

*Department of Physics, University of Toronto, Toronto, Ontario, Canada M5S 1A7*

We provide more evidence of not quite black holes at LIGO. We update and streamline our previous search strategy and apply it to the 10 BH-BH merger events and the one NS-NS event. The strategy is aimed at the evenly-spaced resonance spectrum expected from not quite black holes, given that at low frequencies the radial wave equation describes the modes of a stretched 1D cavity. We describe various indications of the self-consistency of the apparent signals across all events in the context of a simple theoretical model. The merger with the largest final mass, spin and red-shift, GW170729, provides additional interesting support.

## I. INTRODUCTION

A “not quite black hole” is an horizonless object where strong gravity extends out to a radius just slightly beyond, by a distance  $\delta r$ , the would-be horizon. Such an object is very close to being a black hole since the distance  $\delta r$  is characterized by the Planck length. The possibility of having something other than a BH as the endpoint of gravitational collapse is related to having a UV complete theory, since the interior of these objects probe the UV completion. If this UV completion involves an action with terms quadratic in the curvature then these not quite BH solutions exist and are referred to as “2-2-holes” [1, 2]. At the same time quadratic gravity is an old candidate for a UV complete quantum field theory of gravity, since as such it shares with QCD the properties of renormalizability and asymptotic freedom [3]. Other properties of the quantum theory are still being elucidated.

Recently we found 2-2-hole solutions that are sourced by an ordinary relativistic gas [4] (for further analysis see [5]). The total entropy of the gas turns out to satisfy an area law, and this entropy is somewhat larger than the entropy of a black hole of the same mass. In this way 2-2-holes may be preferred as the endpoint of gravitational collapse. Like BHs, they have no upper limit on their size. Unlike BHs, they have a minimum size where they are cold,

---

\* [bob.holdom@utoronto.ca](mailto:bob.holdom@utoronto.ca)

and these Planck-size objects have an internal structure quite different than the large variety. This situation is not unlike QCD. The two types of solutions are analogous to the hadrons and the quark matter states of QCD. The latter states show that the effects of strong interactions can extend over macroscopically large regions even though the fundamental scale of strong interactions is microscopic. In the classically scale invariant version of quadratic gravity, the Planck scale can arise as the scale of strong interactions, just as the QCD scale arises in QCD [6].

We thus consider the possibility that all BHs in nature are in fact not quite BHs. This is a conjecture that can be tested, since a gravitational wave signal can distinguish not quite BHs from BHs. The effective radial description of low frequency waves around the not quite BHs is that of a 1D cavity. One end of the cavity is at the origin ( $r = 0$ ) and the other end is at the angular momentum potential barrier ( $r = 3M$ , in units with  $G = 1$ ). Since the potential barrier slightly leaks, a pulse that moves back and forth in the cavity can produce periodic pulses (echoes) observed on the outside [7, 8]. This observation motivated searches for echoes in LIGO data, beginning with [9]. But the perturbed state of a not quite BH, when newly formed after a merger, may be more complicated. If it produces something other than a single pulse moving back and forth, then the simple echo waveform will be replaced by something else.

A 1D cavity has a more general feature: an evenly-spaced resonance spectrum. This is a more robust and general search target than actual echoes [10, 11]. We can express the size of the cavity as a tortoise coordinate distance  $\Delta x$ . Then the spacing between resonances is  $\Delta f = 1/(2\Delta x) = 1/\Delta t$  where  $\Delta t$  is the round trip travel time (the time delay between the possible echoes). The cavity size is stretched compared to the physical size of the not quite BH since the dimensionless ratio  $\Delta t/M$  is large, something like 800 or so. Since an explicit solution for a rotating not quite BH does not exist yet, we instead use a truncated Kerr spacetime as an approximation. The truncation means that a boundary or wall is introduced slightly outside the horizon.  $\Delta x$  is now the distance between the wall and the potential barrier. This construction yields no angular momentum barrier at the wall, just as a 2-2-hole has no angular momentum barrier at its origin. We discuss the known formula for  $\Delta t/M$  in this model in Section III.

We focus on the dominant  $\ell = m = 2$  gravitational wave mode and obtain the corresponding solution of the Sasaki-Nakamura (SN) equation, and thus obtain the function  $\psi_\omega$

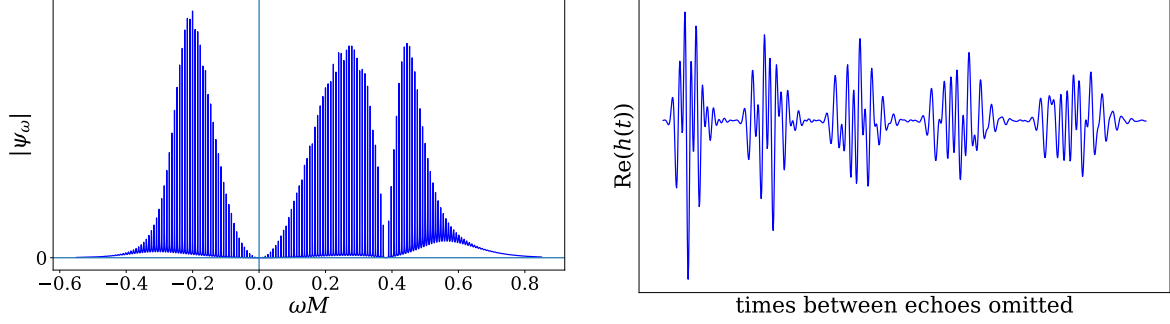


FIG. 1. a)  $|\psi_\omega|$  as obtained from the SN equation. b) The first five echoes.

that encodes the spectrum at asymptotic infinity [11]. This is a product of a transfer function  $K(\omega)$  and a source integral  $D(\omega)$ . The transfer function contains the resonance spectrum that depends on the mass  $M$  and dimensionless spin  $\chi$ , along with  $\Delta x$  and a quantity  $R_{\text{wall}}$  that describes the boundary condition.  $R_{\text{wall}} = -1$  is a purely reflecting boundary condition (and is the analog of the Dirichlet boundary condition found for the spinless 2-2-hole), but we choose a slightly less negative value,  $R_{\text{wall}} = -1 + \epsilon$ , to account for a small amount of damping or dissipation of the wave as it traverses the material inside the not quite BH.

The source integral  $D(\omega)$  modulates the resonance spectrum and its dependence on the initial condition was explored in [11]. The simplest initial condition is a single travelling pulse originating inside the cavity. The frequency content of the pulse can be characterized by a function  $f(\omega)$  that in turn determines the spectral flux density

$$\frac{dE}{d\omega} = 8\omega(\omega - \omega_0)|f(\omega)|^2. \quad (1)$$

This is negative for  $0 < \omega < \omega_0$  and we choose  $f(\omega) \propto \exp(-\kappa M^2 \omega(\omega - \omega_0))$  according to the expectation that the lowest energy modes are the most likely to be excited. (With  $m = 2$ ,  $\omega_0 = \chi/r_+$ , where  $r_+$  is the horizon radius.) An example of  $|\psi_\omega|$  is given in Fig. 1a for the values  $\chi = 2/3$ ,  $\Delta t/M = 800$  and  $\kappa = 13$ . Less trivial initial conditions can be modelled with more travelling pulses and will result in a less smooth envelope for the spikes. This envelope always vanishes at  $\omega_0$ .

From  $\psi_\omega$  the strain waveform in the time domain  $h(t)$  can be obtained by an inverse Laplace transform. As shown in [11],  $\psi_\omega$  contains a factor of a function  $c_0(\omega)$  that appears in the transformation of Teukolsky to SN amplitudes. But to obtain the strain  $h(t)$  from the output

of the SN equation  $\psi_\omega$ , this output must first be divided by  $c_0(\omega)$  before the inverse Laplace transform is performed [12], and thus the  $c_0(\omega)$  factor is cancelled. The resulting  $\text{Re}(h(t))$  is shown in Fig. 1b where we see irregular echoes for even the simplest of initial conditions.

In [10] a search strategy was developed to target the resonance spectrum. Some evidence for such spectra was presented for four BH merger events reported early by LIGO. Here we update this search and apply it to all ten confident detections of BH mergers in the first and second observing runs [13], using publicly available data [14]. Our goal is to simplify the strategy and to apply it uniformly to all events. This reduces look elsewhere effects and it makes our analysis and results more easily reproducible. Another feature of our analysis is that an optimized bandpass, a purely data-driven quantity, can be compared to the predicted spectrum for each event.

## II. THE SEARCH

The search strategy involves taking a range of whitened data starting at the merger time and having duration  $T$ . Then after a FFT, the absolute value is taken. A complication is that the gravitational wave arriving at earth has two polarizations which form the real and imaginary parts of a complex waveform. A LIGO detector, depending on its orientation, projects this complex waveform into a set of real numbers.

Thus to find the observable version of the signal spectrum, we first must model this projection, and we do this by taking the real part of the time-domain signal strain waveform  $h(t)$ . (Different projections give qualitatively similar results.) Carrying out the steps of the search strategy on  $\text{Re}(h(t))$  gives a “reconstructed” spectrum that we will label as  $h(f)$ . This quantity has an additional dependence on  $T = N_E \Delta t$ .  $N_E$  is the number of echoes, loosely speaking, since typically any echoes will have already merged by the time  $T$  is reached.

We show the resulting  $|h(f)|$  in Fig. 2a with  $N_E = 180$ ,  $M = 50M_\odot$ . The range of frequencies over which the resonance spectrum extends is proportional to  $1/M$ , and will thus be different for every event.  $N_E$  along with the choice of  $R_{\text{wall}}$ , for this example chosen to be  $-0.995$ , controls the overall height of the spikes. Relative to a fixed amount of noise, this overall height will increase with  $N_E$  up to roughly  $N_E \sim 200$ , after which the highest set of peaks, the ones that contribute most to a signal, begin to shrink. This behaviour is controlled by the value of  $R_{\text{wall}}$  and it occurs even though the energy radiated per time delay, for instance, is a

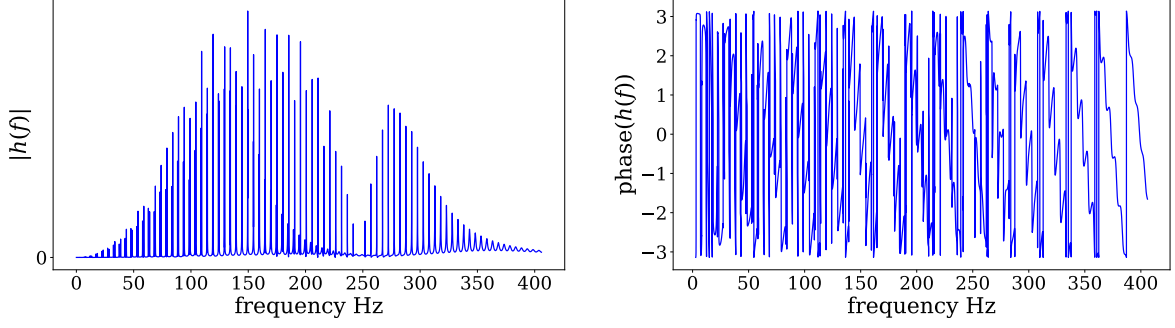


FIG. 2. a)  $|h(f)|$  as reconstructed from  $\text{Re}(h(t))$  b) The phase of  $h(f)$ .

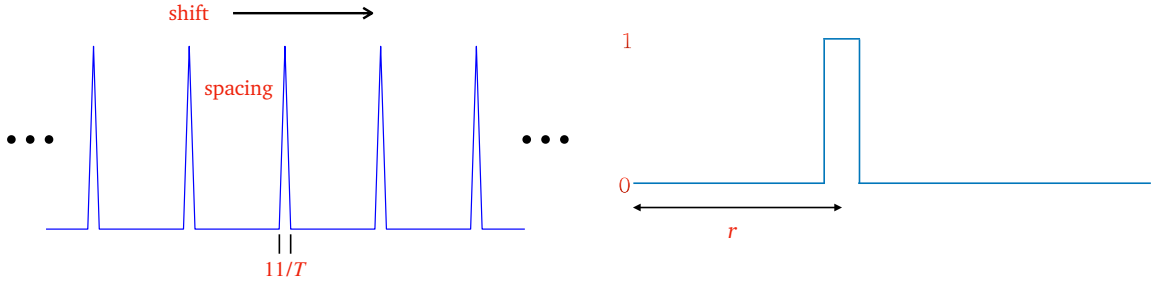


FIG. 3. a) Five teeth of a uniform comb in frequency space. b) The vector  $V(r)$  of length  $N_E$  is composed of 23 1's with location shown and the rest 0's.

monotonically decreasing function of time.

In Fig. 2a we see the two component structure of  $|h(f)|$ , which arises due to the real projection. The lower component originates from what was the negative frequency part of the original  $\psi_\omega$ . The relative size of the two components is controlled by the function  $f(\omega)$ , which as introduced above emphasizes positive frequencies. On the other hand, the source integral  $|D(\omega)|$  contains the factor  $|\omega - \omega_0|$  that enhances the original negative frequency part [11]. Thus in our example the negative frequency component in  $|h(f)|$  is still comparable in size to the positive frequency component.

Other echo searches have tended to try to model the full waveform. In frequency space the full information is in  $|h(f)|$  and the phase of  $h(f)$ . The latter is displayed in Fig. 2b. In a standard matched-filter search, the templates must be constructed to match this information as well. This makes clear why such searches are liable to fail. For a resonance search, only  $|h(f)|$  is needed.

The construction  $|\text{FFT}(\text{whitened data of duration } T \text{ after merger})|$  is a data series in the frequency domain. Onto this we apply a frequency bandpass  $f_{\min} < f < f_{\max}$  to model the

finite extent of the resonance spectrum. Then we multiply by a comb structure characterized by a spacing between teeth and an overall shift. The comb is uniform, meaning all teeth are the same with constant spacing, as shown in Fig. 3a.<sup>1</sup> We thus first define a vector that is the result of multiplying the comb onto the data for a fixed comb spacing,

$$U = \text{mean} [|\text{FFT}(\text{data})| \times \text{comb}(\text{shift})], \quad (2)$$

where the mean is with respect to the nonzero values.  $U$  is a vector in shift space with length  $N_E$ . A signal and the correct comb spacing will result in larger values around some shift. Let  $V(r)$  be another vector that represents an idealized bump centered at a shift of  $r$ . This is illustrated in Fig. 3b. Now we can take the two vectors  $U_H$  and  $U_L$ , from the Hanford and Livingston detectors respectively, and form the following quantity,

$$C = \max_r [\text{Correlation}(U_H, V(r)) \times \text{Correlation}(U_L, V(r))]. \quad (3)$$

$C$  is our correlation [10] between detectors and it is still a function of the comb spacing  $\Delta f = 1/\Delta t$ . A signal shows up as a peak in a plot of  $C(\Delta f)$ . This will be a sharp peak because the spacing needs to be tuned to achieve an overlap with the large number of sharp spikes.

The whitening of the original strain data is an important first step in the analysis, and to obtain the PSD required for whitening, an averaging over 300 segments is performed.<sup>2</sup> The range of data used for obtaining the PSD is larger than the search region, and it is centred on the midpoint of the search region. The duration of the segments (in seconds) used in each event is given in Table I below. Allowing a binary choice of .5 or 1 s seems to better treat the different noise characteristics in different events.<sup>3</sup> A trade-off is involved since although longer segment times will reduce noise on finer-grained frequency scales, it increases the risk of altering the signal that also occurs on fine-grained scales. One further noise cut is applied to the frequency series  $|\text{FFT}(\text{whitened data of duration } T \text{ after merger})|$ . Any value in this series that is more than three times the average value is reduced to the three times value. Such fluctuations should only occur about 0.1% of the time for Gaussian noise, but in the earlier events especially, there is substantial non-Gaussian noise. It is fair to assume that the signal

<sup>1</sup> This is a simplification of our previous analysis [10].

<sup>2</sup> We use the Welch method with the Hanning window and with some overlap and padding turned on. The same settings are used for all events.

<sup>3</sup> Incomplete data for GW151226 prevented the choice of 1 s.

spikes are typically not high enough to be much affected by this cut.

### III. RESULTS

For a truncated Kerr black hole, the time delay  $\Delta t$  has known dependence on mass  $M$ , spin  $\chi$  and redshift  $z$  [9, 15],

$$\frac{\Delta t}{M} = 4 \eta \log\left(\frac{M}{\ell_{\text{Pl}}}\right) \left(\frac{1 + (1 - \chi^2)^{-\frac{1}{2}}}{2}\right) (1 + z). \quad (4)$$

The redshift factor is due to  $\Delta t$  and  $M$  being measured in the detector and source frames respectively.  $\Delta t/M$  is driven large by the large log, due to the presence of the Planck scale. We have introduced the parameter  $\eta$  to characterize the small distance  $\delta r$  from the would-be horizon out to where strong gravity extends, as follows,

$$\delta r \approx \left(\frac{\ell_{\text{Pl}}}{M}\right)^{(\eta-1)} \ell_{\text{Pl}} \approx \left(\frac{M}{\ell_{\text{Pl}}}\right)^{(2-\eta)} \bar{\ell}_{\text{Pl}}, \quad (5)$$

where  $\bar{\ell}_{\text{Pl}}$  is the proper Planck length. Thus if  $\eta$  is between 1 and 2, then  $\delta r$  is smaller than the coordinate Planck length  $\ell_{\text{Pl}}$  and is larger than the proper Planck length. Because of the many evenly-spaced spikes, a resonance signal will determine  $\Delta t$  with negligible error. The LIGO measurements of  $M$ ,  $\chi$  and  $z$  for the final BH have more significant errors, and by ignoring correlations in these errors for simplicity, we can thus determine an  $\eta$  and its error from a measurement of  $\Delta t$ .

We find resonance signals for all events. Before elaborating on this we first provide a summary plot in Fig. 4, where we show our determination of  $\eta$  from all ten BH merger events. The consistency for a common value of  $\eta$  describing all events is excellent. The value  $\eta = 1.72 \pm 0.06$  means that  $\delta r \approx 10^{-28} \ell_{\text{Pl}} \approx 10^{12} \bar{\ell}_{\text{Pl}}$ . We stress that the size of the error bars has nothing to do with the strength of the resonance signal for each event. We shall address the question of signal strengths below.

These results provide a good test of the spin and redshift dependence in (4). This is shown more clearly in Fig. 5, where the left plot is the same as Fig. 4 but with an enlarged range of  $\eta$ , while the right plot shows the result when removing the spin and red-shift dependent factors from (4). The indication of a common value for  $\eta$  is lost, and the individual values of  $\eta$



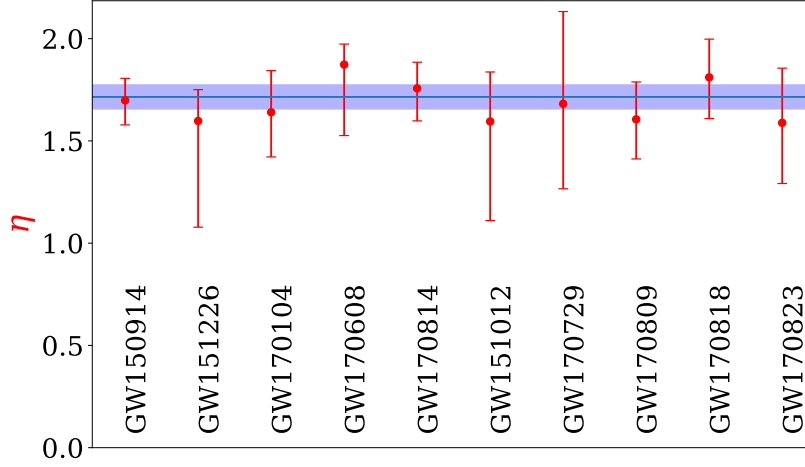


FIG. 4. Determination of  $\eta$  using (4).

rise above 2. Values this large would correspond to  $\delta r$  values that are much smaller than the proper Planck length, which does not seem reasonable. Thus the data is already supporting the truncated Kerr BH model in (4). Fig. 5 also draws attention to event GW170729, for which the three quantities,  $M$ ,  $\chi$  and  $z$ , are all significantly larger than for the other events.

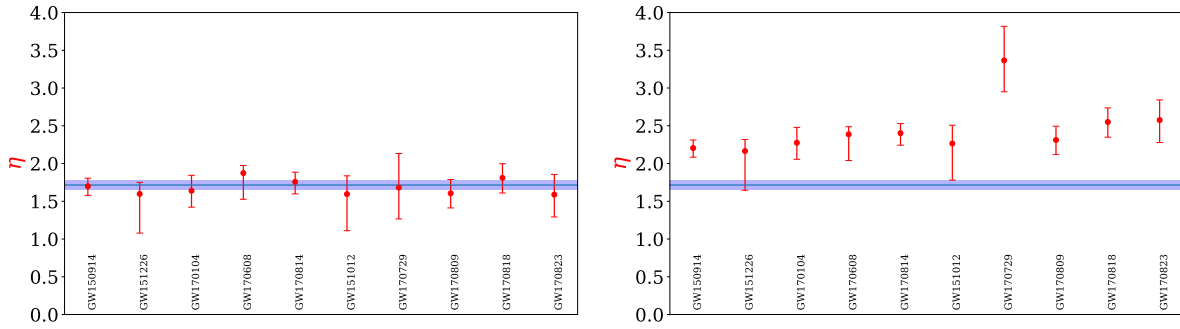


FIG. 5. Left: same as Fig. 4. Right: removal of spin and red-shift factors in (4).

We now take a closer look at the analysis and the results for each event. For each event we will show the predicted spectrum  $|h(f)|$ . To simplify the generation of these spectra we assign one of three spin values to each event. These values are well within the spin uncertainties, and so we group the ten events as in Table I. There we also give our chosen values of parameters  $R_{\text{wall}}$  and  $\kappa$  for these three groups of events, which along with  $M$ ,  $\chi$ ,  $\Delta t$  and  $N_E$  are used to obtain  $|h(f)|$ .

Except for the binary choice of a whitening parameter, as indicated in Table I, the analysis

spin ( $\chi$ )	2/3	0.72	0.81
	GW150914 (.5)	GW151226 (.8)	GW170729 (.5)
	GW170104 (1)	GW170814 (1)	
	GW170608 (1)	GW170809 (1)	
	GW151012 (.5)	GW170823 (.5)	
	GW170818 (.5)		
$R_{\text{wall}}$	-0.995	-0.994	-0.992
$\kappa$	13	10	2

TABLE I. The ten events are grouped according to spin, along with other chosen parameters. The numbers beside event names are segment durations in seconds used in whitening.

carried out for each event is the exactly the same. There are no other procedures or parameters that vary between events. We will express the comb spacing  $\Delta f$  as an integer,  $n = \Delta f / (1/T) = T/\Delta t$ . It is respect to  $n$  that we determine the correlation between detectors as described earlier. As we have said, a signal peak should be very sensitive to choice of  $\Delta t$ . In the process of determining the true  $\Delta t$ , we can choose  $T$  such that the peak location is fixed at a particular integer value of  $n = N_E$ , that is that  $T$  is  $N_E$  times the true  $\Delta t$ .

We then further optimize the signal (the peak height) by varying three quantities:  $N_E$  and the upper and lower limits of the bandpass  $f_{\min}$  and  $f_{\max}$ . For each event shall display the optimal bandpass as a band on a plot. We also give the optimal  $N_E$  (as a multiple of ten) and the measured values of  $\Delta f$  and  $\Delta t/M$  (where the uncertainty inherent in  $M$  is not displayed). We have already seen that it is the spin and redshift factors that account for the different values of  $\Delta t/M$  from different events. The plot will compare the optimal bandpass to the resonance pattern in  $|h(f)|$ . We stress that the optimal bandpass is determined solely from the data, while the resonance pattern is a theoretical quantity. It is interesting to note the consistency between the bandpass and the resonance pattern across all events.

We also wish to give a sense of the size of the signal relative to noise, with the latter being the square root of the power spectral density, or ASD, for each detector (as was used in the whitening). To make a comparison of the signal  $|h(f)|$  with the noise possible, we shall effectively calculate  $|h(f)|$  also as an ASD, but without employing the averaging and windowing (ie. the Welch method) that was used to calculate the noise ASD. This is equivalent to producing  $|h(f)|$  through a FFT, but in a way that gives the proper normalization to allow comparison of signal and noise.

To help pin down the strength of the signal we inject our time-domain waveform  $\text{Re}(h(t))$  into some stretch of strain data from each detector, away from the signal region. By running this signal plus noise through our analysis we find that the resulting peak on the injected signal plot can be similar to that on the actual signal plot. This then provides a rough estimate of the required strength of the signal, and we incorporate this into the normalization of  $|h(f)|$  on the plot. We also provide the corresponding peak value of  $|h(t)|$ , where the peak value occurs in the first echo.<sup>4</sup> We do not include the first pulse, that is the pulse before the first echo, in our injected waveform as this pulse is to be associated with the actual merger and ringdown. The starting time of our signal region is always taken to be after the ringdown has occurred. Our plots show separately the small ASD of the first pulse as a black line.

We also show a second plot for each event that shows the distribution of background values of the correlation along with the size of the correlation at the signal peak. We will discuss this more in the next section on p-values.

These plots and results are presented below for the five spin 2/3 events first (Figs. 6-10) and then the four spin 0.72 events (Figs. 11-14). Each figure shows the spectral plot and the correlation plot. The remaining event GW170729 (Fig. 15) has been flagged above, and we find that its results are both interesting and different. Here we find two resonance signals, for two non-overlapping choices of the bandpass. And both of these signals are strong, as seen from the two correlation plots. Each signal produces its own value of  $\Delta f$  that happen to be essentially the same. From the displayed  $|h(f)|$  we see that the high spin causes the positive and negative frequency components to become more separated, and the two signals are consistent with observing these two components separately. When attempting to use a broader bandpass that would encompass both components, the signal effectively disappears. Given that each signal has the same resonance spacing, the implication is that the two resonance patterns are displaced relative to each other, in such a way that the broad uniform comb misses one or the other. Such a relative shift is natural and it depends sensitively on  $\omega_0$ , which pins the overall shift of  $\psi_\omega$ . Since the spin and thus  $\omega_0$  is not measured accurately enough, this shift cannot be predicted.

Finally we present results for the NS-NS merger event, GW170817 in Fig. 16. We reported a resonance signal for this event in [10], but now we pass it through the current analysis. The

---

<sup>4</sup> Our estimate of the peak value assumes the simplest initial perturbation, that of a single pulse moving back and forth in the cavity. For a more general perturbation, the required peak value of  $|h(t)|$  may be substantially less.

results remain similar to before. The time scales are shorter for this event and the high sampling rate data (16384 Hz) is used. We use 200 segments of 1/4 s duration for the whitening. LIGO does not provide the final mass and spin, so here we make choices ( $M = 2.2M_{\odot}$  and  $\chi = 2/3$ ) that give a sensible looking  $|h(f)|$  when compared to the optimal bandpass. The measured  $\Delta t$  then corresponds to  $\eta = 1.6$ , and increasing  $M$  would decrease  $\eta$  further. We do not use this event in the analysis of the next section.

### P-VALUES

In this section we shall obtain probabilities useful for estimating p-values. By using many stretches of data we can quite accurately determine the background distribution of our correlation quantity, separately for each event. Then we can find the probability of finding a value at least as high as our signal peak. A p-value of each event should also account for the fact that the measured  $\Delta t$  gives an  $\eta$  that is consistent with the nine other events, as per Fig. 4. So for example if we require that  $\eta$  be within  $2\sigma$  of its central value, then this gives the range of  $\Delta t$  (or a range of  $n$  on a signal plot) over which we can look for a background correlation as high as our signal correlation.

We use data throughout the 4096s segment that contains the merger event, and within this we use segments of the same length as in the signal analysis with locations chosen randomly for each detector. The whitening of the data is repeated and the already determined values of  $\Delta t$ ,  $f_{\min}$ ,  $f_{\max}$  and  $N_E$  are used to produce many background versions of the signal plot. Each such plot has  $n$  ranging from .6 to 1.4 times  $N_E$ . This procedure has an important benefit, namely that we can verify that the correlation at the central value, where the peak occurs in the signal plot, does not have any excess strength in the background plots.<sup>5</sup> A correlation that exists outside the signal region would indicate that it is being generated by noise.

By collecting together all the correlation values from all the background plots, we can then determine their distribution. We find that a good fit is obtained to a function of the form  $x^{\alpha}e^{-\beta x}$ , which when normalized is a generalized gamma probability density function  $P_{gg}(x)$ . The fit (different for each event) has four parameters,  $\alpha$ ,  $\beta$ , and a scale and a location parameter. These parameters are determined by a maximum likelihood estimate directly from

---

<sup>5</sup> The one exception is GW150914 where there is some excess correlation at or close to the central value. This may be related to how close  $\Delta f = 3.991$  Hz is to 4 Hz, thus suggesting that there is some 4 Hz correlation generated by noise. This excess correlation is weaker than the signal.

the correlation values ( $\sim 20000$  values). Typically  $\beta$  is between 0.5 and 1 and the location parameter is close to vanishing.

A plot for each event shows  $P_{\text{gg}}(x)$  (red curve) for the correlations from background. The histogram is added to indicate how well  $P_{\text{gg}}(x)$  fits the data. Also shown are the signal peak values of the correlation. The signal peak value  $x_{\text{sig}}$  determines the probability  $\int_{x_{\text{sig}}}^{\infty} dx P_{\text{gg}}(x)$  that any single background measurement can produce a value of  $x$  at least as high. We then multiply by the number values of  $n$  that lie within  $2\sigma$  of  $n = N_E$  (explicitly, this number of values is  $0.14N_E$ ). The resulting numbers are simple estimates of p-values and they are shown in Table II.

GW150914	0.008	GW151226	0.014
GW170104	0.33	GW170814	0.098
GW170608	0.038	GW170809	0.081
GW151012	0.0016	GW170823	0.026
GW170818	0.0094	GW170729	0.0010 & 0.0006

TABLE II. Simple p-values for the 10 BBH merger events. These are the probabilities that background noise can generate the signal peak height or higher for any value of  $n$  within 7% ( $2\sigma$ ) of  $N_E$ .

These simple p-values have the advantage of being well-defined and they usefully show the relative signal strength of the different events. They highlight again the significance of the GW170729 results. From the individual p-values we can contemplate a global p-value. Combining p-values is not a completely well-defined procedure and various methods are used. Combining the 11 p-values in Table II according to the Fisher and Stouffer methods<sup>6</sup> (which assume independence of the p-values to be combined) gives  $5 \times 10^{-11}$  and  $2 \times 10^{-12}$  respectively.

We could expect that the individual p-values become somewhat less trustworthy as the signal value  $x_{\text{sig}}$  goes further out onto the tail of the distribution, since the analytical  $P_{\text{gg}}(x)$  is being used to extrapolate into this region where there are few background values. Thus the lowest p-values Table II may have relatively larger uncertainty. To consider the impact of this on the combined p-value, we can for example add a fixed amount to *all* the p-values, which relatively impacts the lowest ones the most. For example when we add 0.005, the combined p-values become  $1 \times 10^{-8}$  and  $2 \times 10^{-10}$ . This is just for illustration, and 0.005 seems far more than needed to account for the PDF uncertainty.

<sup>6</sup> The Stouffer method may be the less well known, but it is easy to describe: convert each p-value  $p_i$  to  $x_i$  using  $p_i = \int_{x_i}^{\infty} dx P_{\text{norm}}(x)$ , get the combined  $x_c = \sum_i x_i / \sqrt{N}$ , and then convert  $x_c$  back to the combined p-value  $p_c$ .

In contrast to our simple p-values, we could consider obtaining p-values by repeating our *entire* analysis, many times, on background data. Our search would have to be automated and be applicable to both signal and background regions. p-values may depend on how this is accomplished. Although we do not accomplish this here, it would entail specifying the ranges of quantities, over which they are varied to find the signals. This includes  $\Delta t$  itself as well as the frequency bandpass limits  $f_{\min}$  and  $f_{\max}$  and the integer  $N_E/10$ . We could then find the probability for finding a background correlation as high as our signal correlation, for each event, and thus a set of  $N$  p-values (here  $N = 11$ ). The combined p-value would then be the probability of observing correlations as extreme as our signal correlations, over the ranges specified and assuming background only.

But there is more to our signals than the correlation size. For a particular background that can generate a signal-size correlation, the above mentioned quantities now take whatever random values the background produces. One then needs to consider the probability that these random values could display the regularities actually observed in the signals. For  $\Delta t$  we have taken the regularity to be the  $2\sigma$  version of the narrow band in Fig. 4. If this corresponds to a fraction  $p_1$  of the  $\Delta t$  range actually explored for each event, then the probability would be  $\sim p_1^N$ . For the bandpass, it is the probability that each of  $f_{\min}$  and  $f_{\max}$  lies in the range that is compatible with our theoretical spectra, as happens for the signals. We write this probability as  $\sim p_2^N p_3^N$ . For  $N_E$ , it is the probability that each event has correlation peaks for a range of  $N_E$  (even though we have chosen only one in each case) and that these ranges largely overlap between events, as happens for the signals (other than GW151226). We write this probability as  $\sim p_4^{N-1}$ . The final probability of interest would then be the combined p-value times  $\sim (p_1 p_2 p_3)^N p_4^{N-1}$ . In fact we could consider increasing the search ranges until the combined p-value is no longer very small. But then the  $p_i$ 's will be small, and we see that the  $p_i$ 's do not have to be very small to imply by themselves a very small final probability.

The point is that in addition to evidence against the background hypothesis, we are seeing evidence for the signal hypothesis. The signal hypothesis involves a simple theoretical model that is predicting how the various quantities in each event should be related, and the data is largely conforming to these predictions. Returning to our simple p-values, they are being defined for each event by holding  $f_{\min}$ ,  $f_{\max}$  and  $N_E$  fixed to their values from the signal region. We are ignoring both the search ranges of these quantities and how the observed regularities in these quantities support the signal hypothesis. These are compensating effects and so our

simple p-values are a sensible first step to assess the significance of our results. Our simple p-values also ignore the additional evidence for the signal hypothesis coming from GW170729, namely that the two different bandpasses have produced the same  $\Delta f$ .

## CONCLUSIONS

LIGO is sensitive to the cavity resonance structure of not quite black holes. This particular sensitivity to Planck-scale physics is under appreciated. The full signal waveform produced by a not quite black hole may have too much information and model dependence for a standard matched-filter analysis. On the other hand the evenly-spaced resonance pattern is robust and relatively easy to search for. This is the approach taken here.

Evidence is already accumulating from the 10 BH merger events reported in [13]. We have described the consistency of the measured  $\Delta t$  values with the mass, spin and red-shift dependence predicted from the simple truncated Kerr BH model. Event GW170729 in particular is sufficiently different from the other events to provide a good test of the model. We have highlighted the strength of its signal and the way it may already be showing evidence of the expected two component structure of the resonance pattern [11]. We have also displayed the consistency of the data-driven bandpasses with the spectrum predicted for each event in the same model. The measured  $\Delta t$  values point to a common distance scale  $\delta r$ , with our measured value being larger than the proper Planck length,  $\delta r \approx 10^{12} \bar{\ell}_{\text{Pl}}$ .

The LIGO data certainly displays non-Gaussian noise, some of it not understood, and some of it displaying comb-like features reminiscent of our signal. But given the observed regularities of the signals, and in addition the low p-values, it is now very unreasonable to expect that this could explain our signals. Our simple p-values are both transparent and able to provide a sensible first estimate of the significance of our results. The combined p-value is found to be very small.

We have stressed that our search strategy is uniform across all events. But to define the strategy, some choices and parameters were fixed once and for all. With the strategy now frozen, it will be important to apply it to new data. Many new BBH merger events have already been discovered in the LIGO-Virgo O3 run and it only remains to look at this data. We have one hope concerning the data. In their continuing efforts to reduce noise with respect to their merger-signal target, it is to be hoped that LIGO-Virgo does not inadvertently remove

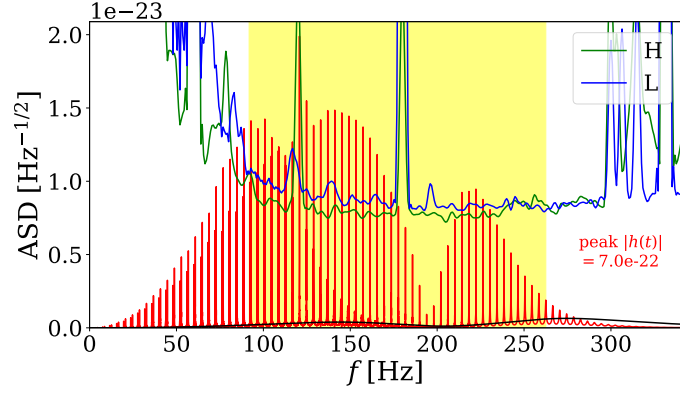
poorly understood non-Gaussian noise that, as seen here, could harbour new physics.<sup>7</sup>

---

<sup>7</sup> In this connection we note that a strong resonance signal for event GW170104 as described in [10] was based on an early O2 data release. This signal weakens considerably when the subsequent GWTC-1 data release [13] is used instead, and now GW170104 gives our weakest signal. For all events in the present analysis we use only the GWTC-1 data release.



FIG. 6. GW150914



$N_E$	200
$\Delta f$	3.991 Hz
$\frac{\Delta t}{M}$	806
$\chi$	2/3

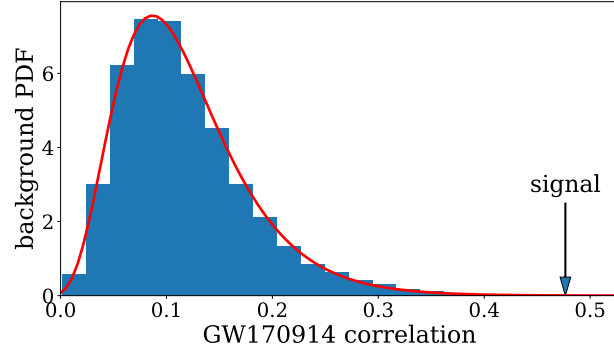
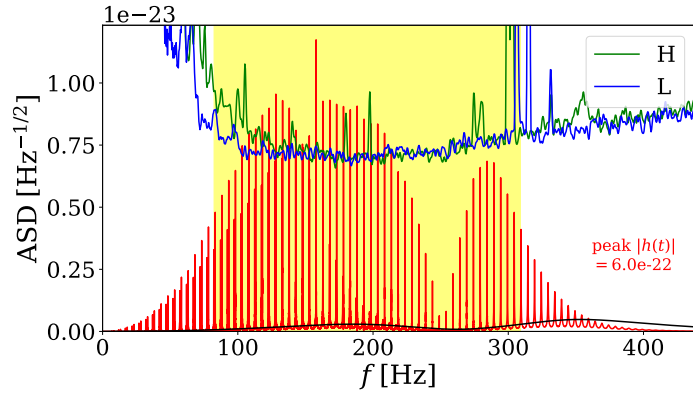


FIG. 7. GW170104



$N_E$	150
$\Delta f$	4.976 Hz
$\frac{\Delta t}{M}$	831
$\chi$	2/3

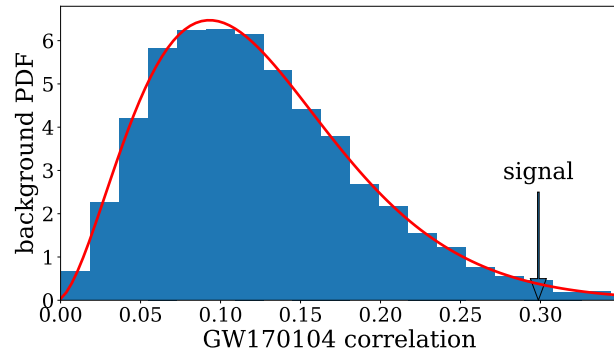
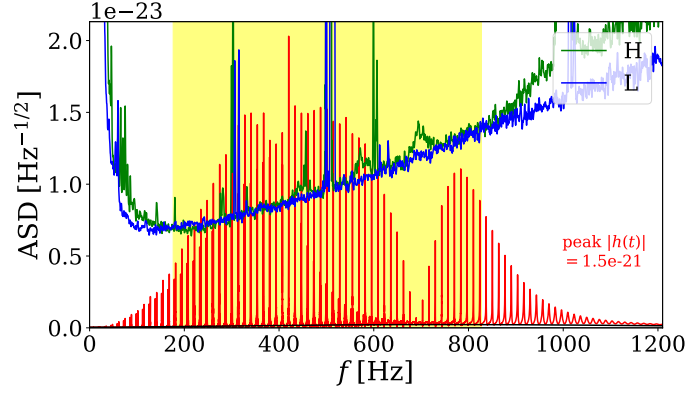


FIG. 8. GW170608



$N_E$	200
$\Delta f$	13.226Hz
$\frac{\Delta t}{M}$	862
$\chi$	2/3

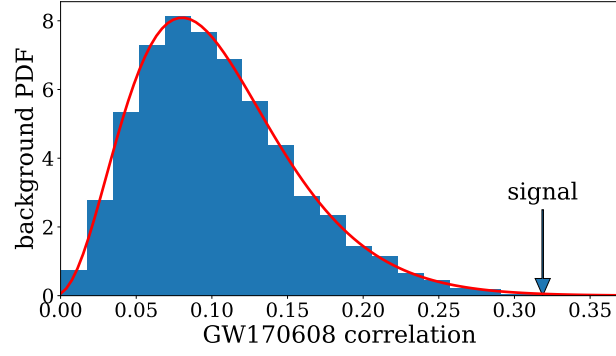
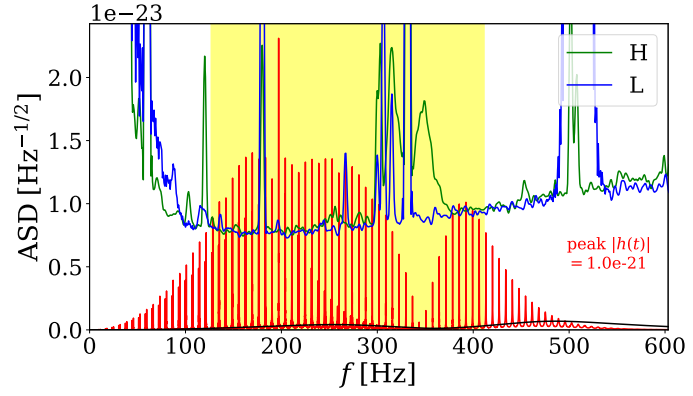


FIG. 9. GW151012



$N_E$	160
$\Delta f$	6.8802Hz
$\frac{\Delta t}{M}$	826
$\chi$	2/3

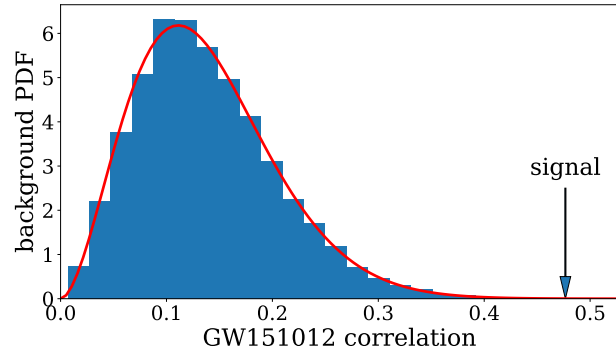
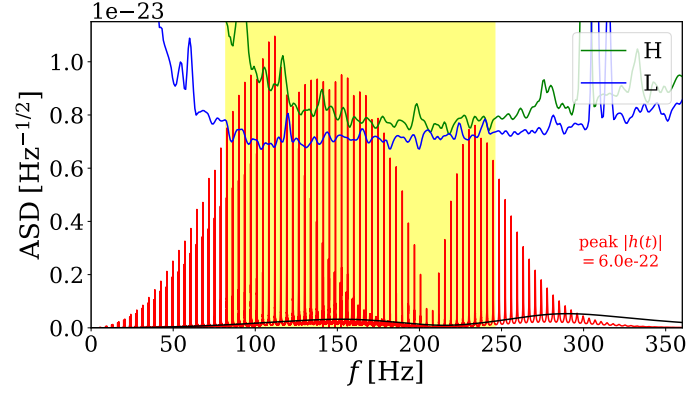


FIG. 10. GW170818



$N_E$	140
$\Delta f$	3.638Hz
$\frac{\Delta t}{M}$	933
$\chi$	2/3

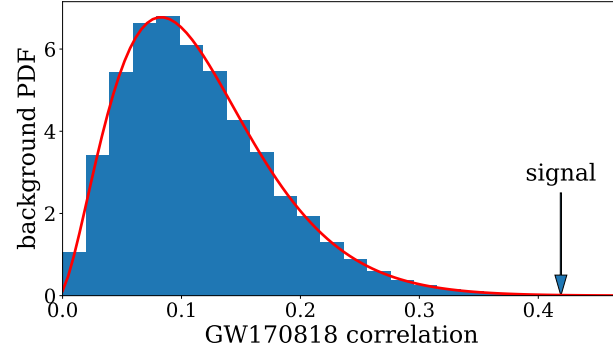
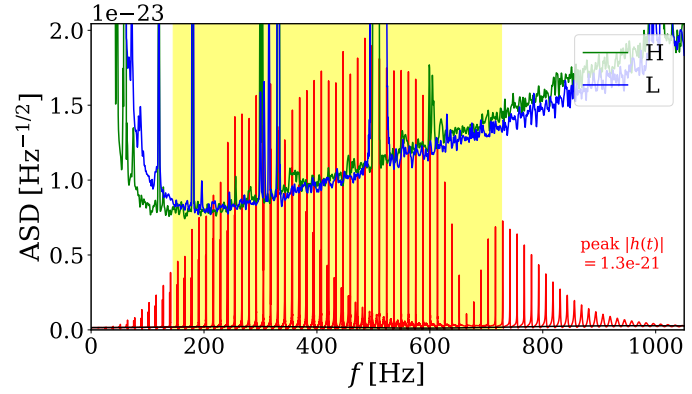


FIG. 11. GW151226



$N_E$	270
$\Delta f$	12.644Hz
$\frac{\Delta t}{M}$	783
$\chi$	0.72

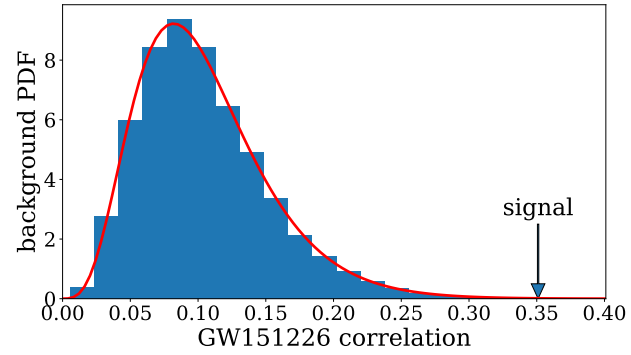
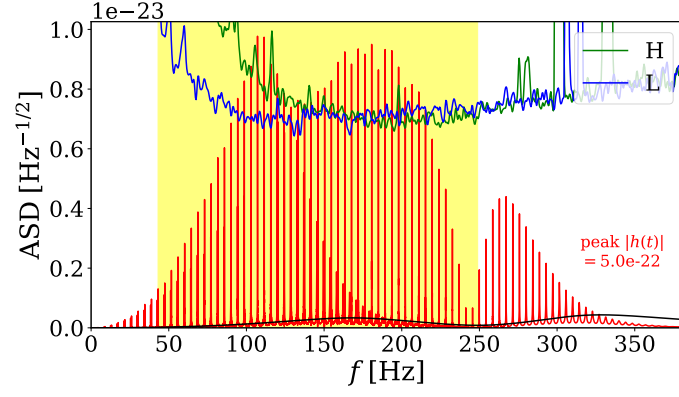


FIG. 12. GW170809



$N_E$	170
$\Delta f$	4.26Hz
$\frac{\Delta t}{M}$	845
$\chi$	0.72

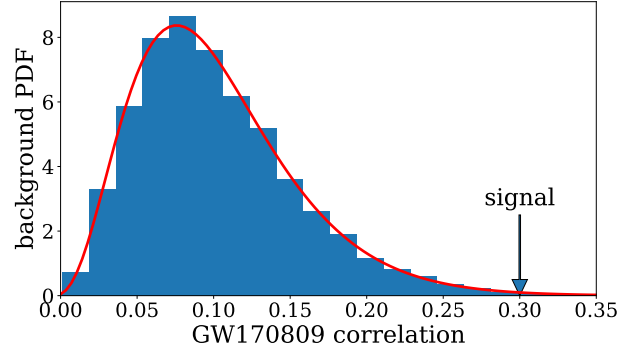
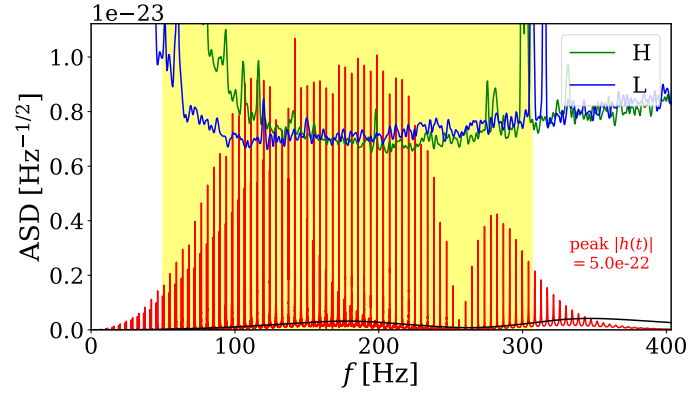


FIG. 13. GW170814



$N_E$	200
$\Delta f$	4.329Hz
$\frac{\Delta t}{M}$	878
$\chi$	0.72

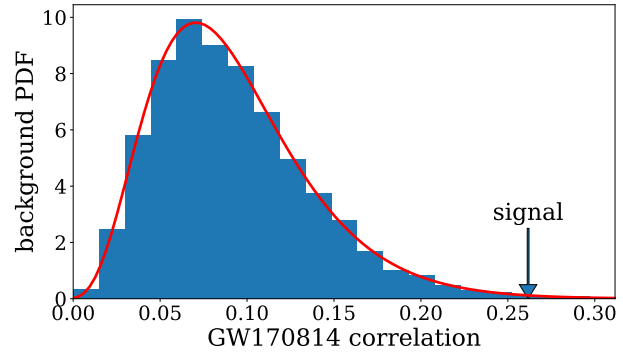
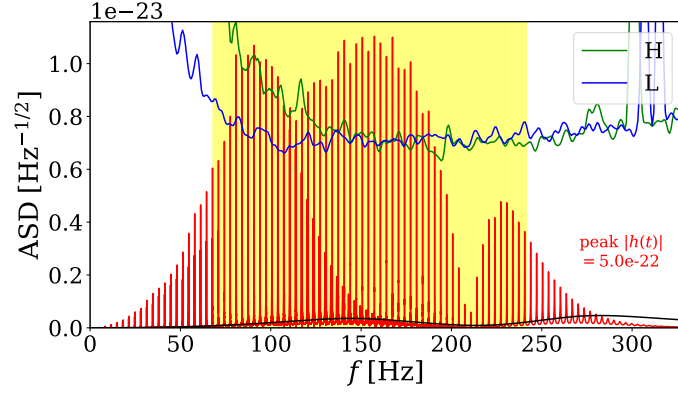


FIG. 14. GW170823



$N_E$	200
$\Delta f$	3.283Hz
$\frac{\Delta t}{M}$	942
$\chi$	0.72

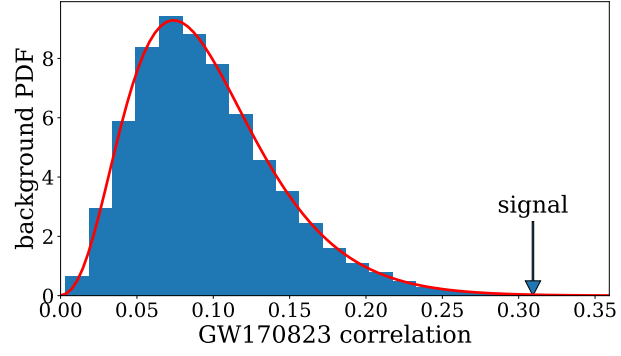


FIG. 15. GW170729

$N_E$	180/170
$\Delta f$	2.0441/2.0443Hz
$\frac{\Delta t}{M}$	1240
$\chi$	0.81

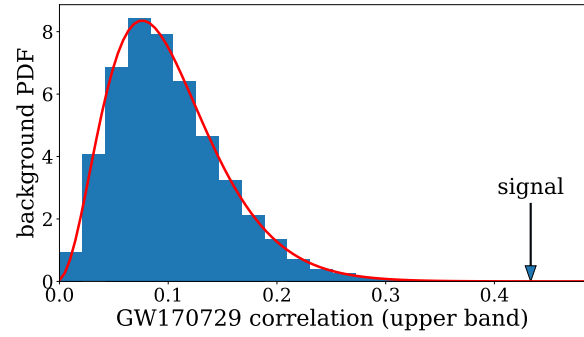
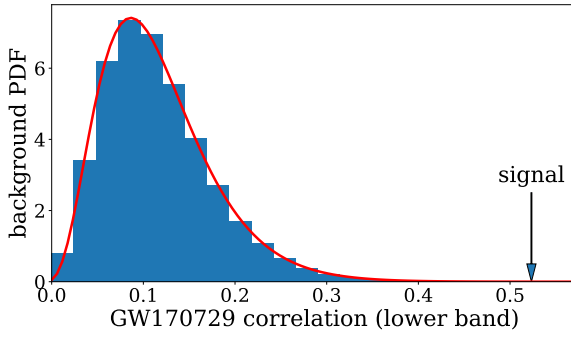
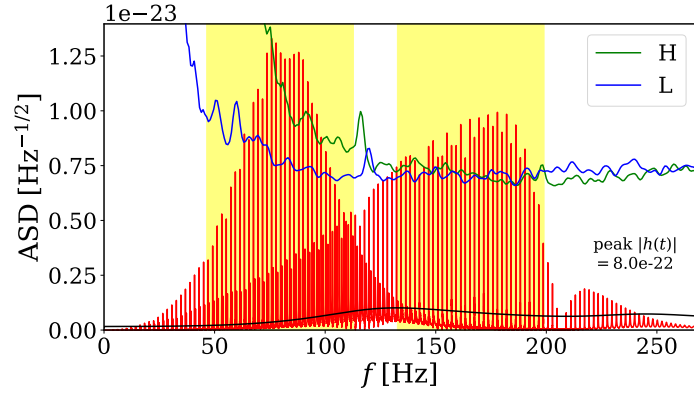
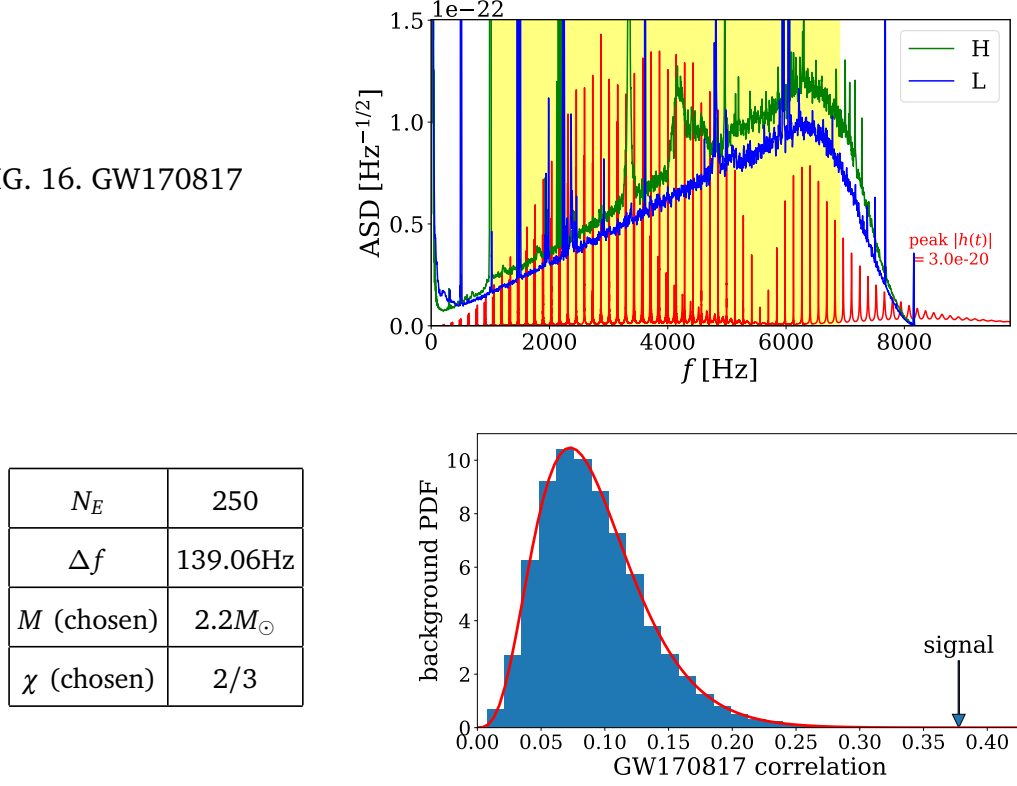


FIG. 16. GW170817



### ACKNOWLEDGMENTS

We thank the organizers of Cosmological Frontiers in Fundamental Physics 2019, at the Perimeter Institute September 3-6, where a version of this work was presented. This research was supported in part by the Natural Sciences and Engineering Research Council of Canada.

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (<https://www.gw-openscience.org>), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by Polish and Hungarian institutes.

- 
- [1] B. Holdom, “On the fate of singularities and horizons in higher derivative gravity”, Phys. Rev. D **66**, 084010 (2002) [hep-th/0206219].
  - [2] B. Holdom and J. Ren, “Not quite a black hole”, Phys. Rev. D **95**, no. 8, 084034 (2017).
  - [3] K. S. Stelle, Phys. Rev. D **16**, 953 (1977).
  - [4] B. Holdom, “A ghost and a naked singularity; facing our demons,” arXiv:1905.08849 [gr-qc].
  - [5] J. Ren, “Anatomy of a thermal black hole mimicker,” Phys. Rev. D **100**, no. 12, 124012 (2019) [arXiv:1905.09973 [gr-qc]].
  - [6] B. Holdom and J. Ren, “QCD analogy for quantum gravity”, Phys. Rev. D **93**, no. 12, 124030 (2016) [arXiv:1512.05305 [hep-th]].
  - [7] V. Cardoso, E. Franzin and P. Pani, “Is the gravitational-wave ringdown a probe of the event horizon?,” Phys. Rev. Lett. **116**, no. 17, 171101 (2016) Erratum: [Phys. Rev. Lett. **117**, no. 8, 089902 (2016)] [arXiv:1602.07309 [gr-qc]].
  - [8] V. Cardoso, S. Hopper, C. F. B. Macedo, C. Palenzuela and P. Pani, “Gravitational-wave signatures of exotic compact objects and of quantum corrections at the horizon scale”, Phys. Rev. D **94**, no. 8, 084031 (2016) [arXiv:1608.08637 [gr-qc]].
  - [9] J. Abedi, H. Dykaar and N. Afshordi, “Echoes from the Abyss: Tentative evidence for Planck-scale structure at black hole horizons”, Phys. Rev. D **96**, no. 8, 082004 (2017) [arXiv:1612.00266 [gr-qc]].
  - [10] R. S. Conklin, B. Holdom and J. Ren, “Gravitational wave echoes through new windows”, Phys. Rev. D **98**, no. 4, 044021 (2018) [arXiv:1712.06517 [gr-qc]].
  - [11] R. S. Conklin and B. Holdom, “Gravitational wave echo spectra,” Phys. Rev. D **100**, no. 12, 124030 (2019) [arXiv:1905.09370 [gr-qc]].
  - [12] R. S. Conklin, “Gravitational Wave Perturbations on a Kerr Background and Applications for Echoes,” Phys. Rev. D **101**, no. 4, 044045 (2020) [arXiv:1911.07122 [gr-qc]].
  - [13] B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], “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**, no. 3, 031040 (2019) [arXiv:1811.12907 [astro-ph.HE]].
  - [14] M. Vallisneri *et al.* “The LIGO Open Science Center”, proceedings of the 10th LISA Symposium, University of Florida, Gainesville, May 18-23, 2014; also arXiv:1410.4839.

- [15] V. Cardoso and P. Pani, “The observational evidence for horizons: from echoes to precision gravitational-wave physics”, *Nat. Astron.* 1 (2017) 586 [arXiv:1707.03021 [gr-qc]].