Parameter estimation for black hole echo signals and their statistical significance

Alex B. Nielsen,[1, 2, *] Collin D. Capano,[1, 2, †] Ofek Birnholtz,[3, ‡] and Julian Westerweck[1, 2, §]

[1] *Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany*
[2] *Leibniz Universität Hannover, D-30167 Hannover, Germany*
[3] *Center for Computational Relativity and Gravitation, Rochester Institute of Technology,*
170 Lomb Memorial Drive, Rochester, New York 14623, USA

Searching for black hole echo signals with gravitational waves provides a means of probing the near-horizon regime of these objects. We demonstrate a pipeline to efficiently search for these signals in gravitational wave data and calculate model selection probabilities between signal and no-signal hypotheses. As an example of its use we calculate Bayes factors for the Abedi-Dykaar-Afshordi (ADA) model on events in LIGO's first observing run and compare to existing results in the literature. We discuss the benefits of using a full likelihood exploration over existing search methods that used template banks and calculated p-values. We use the waveforms of ADA, although the method is easily extendable to other waveforms. With these waveforms we are able to demonstrate a range of echo amplitudes that is already is ruled out by the data.

I. INTRODUCTION

Black holes are defined by their horizons [1]. Although a large amount of astrophysical data is compatible with the existence of black holes [2], a number of theoretical models still predict dark compact objects without horizons or for which the horizon structure is significantly modified from classical vacuum general relativity [3–7]. These models are typically motivated by quantum effects or attempts to address issues related to black hole information and evaporation [8]. One possible observational signature of such structure is that infalling waves would not be entirely absorbed by the horizon as is generally expected in general relativity, but instead some amount of the infalling wave would be reflected.

Recent observations of gravitational waves from coalescences of binary black holes [9–14] by the LIGO [15] and Virgo [16] detectors have allowed for a number of new tests of the near horizon structure of black holes [17–19]. One such test involves searching for echo signals that could potentially be caused by reflective structure forming at or near the location of the black-hole horizon. A number of groups have searched for such signals in gravitational wave data with contrasting conclusions [20–22]. Here we propose a new method to search for these echo signals that provides an explicit probability for the compatibility of the data with the echoes hypothesis relative to Gaussian noise. We demonstrate this method on the binary black hole events detected during the first observing run of the Advanced LIGO detectors; these events were the subject of previous studies [20–22].

The general physical picture of echoes is that infalling radiation is reflected due to some mechanism near the putative horizon location. This radiation is then partially trapped between the near-horizon structure and the angular momentum light-ring barrier [23]. Some of the energy is transmitted away from the system by successive bounces, thereby forming a series of echoes. Generic parameters in the physical models are the amount of wave reflected by the boundary and the effective location where this reflection occurs. These in turn are related to the amplitude of the reflected echo signals and the time separation between the successive echoes. Bounds on the amplitude and time separation of echo signals derived from the data can be translated into bounds on the reflectivity and location of the near-horizon structure.

For illustrative purposes here, we focus on the explicit model of Abedi-Dykaar-Afshordi (ADA) [20], which has been the subject of discussion in the literature [21, 24–26]. However, we note that our methodology can just as well be applied to other, more detailed models with explicit waveforms, including those recently proposed in the literature [27, 28]. Efforts to search for echo templates using Bayesian model selection have been developed with LALInference [29] in parallel to our own work, and published concurrently with our own [30]. Other, model-agnostic searches [31], have also been ongoing, along with different techniques to constrain horizonless objects through their impact on the stochastic background [32].

The primary result of [20] is a p-value, calculated as the probability of observing a signal-to-noise ratio (SNR) in noise (assumed to be free of signal) at least as significant as that observed in the on-source data that potentially contains the signal. This by itself does not indicate the probability that the on-source data contains a signal. A probability that the data contains a signal can however be obtained using Bayes' theorem:

$$\mathrm{P(signal|data)} = \frac{\mathrm{P(data|signal)P(signal)}}{\mathrm{P(data)}}. \quad (1)$$

It is most convenient to compare this probability to an alternative hypothesis, for example that the data contains only Gaussian noise:

$$\frac{\mathrm{P(signal|data)}}{\mathrm{P(noise|data)}} = \frac{\mathrm{P(data|signal)}}{\mathrm{P(data|noise)}} \frac{\mathrm{P(signal)}}{\mathrm{P(noise)}}. \quad (2)$$

———————

[*] alex.nielsen@aei.mpg.de
[†] collin.capano@aei.mpg.de
[‡] ofek@mail.rit.edu
[§] julian.westerweck@aei.mpg.de

Echo param.	Prior range	GW150914 range	Injected value
Δt_{echo}	inferred	0.2825 to 0.3025 s	0.2925 s
t_{echo}	$\Delta t_{echo} \pm 1\%$	0.2795 to 0.3055 s	0.2925 s
t_0 trunc.	$(-0.1 \text{ to } 0)\Delta t_{echo}$	-0.02925 to 0 s	-0.02457 s
γ	0.1 to 0.9	0.1 to 0.9	0.8
A	unconstrained	0.00001 to 0.9	varying

TABLE I. Table of prior ranges and values used for injection studies. The ranges are adopted from [20] and the injected values are chosen to lie close to the parameter values found in that work, except for γ and t_0 trunc. which are chosen to lie within the prior range rather than at the boundary.

In the above, the first factor on the right hand side is the likelihood ratio and the second factor is the prior odds. Evaluating the prior odds is difficult without prior data (and in the case of a signal model that violates expected physics, might well be a very small factor) but the likelihood ratio can be calculated by exploring the likelihood function over the model parameters using a stochastic sampling algorithm, such as a Markov chain Monte Carlo (MCMC). To obtain a final Bayes factor, the model parameters must be marginalized over using their respective prior distributions.

The example we consider here is based on the hypothesis of ADA [20]; we refer the reader to that work for more detail on the model and the meaning of the various model parameters. The most important of these parameters are the overall amplitude of the echoes relative to the original signal's peak A, the relative amplitude between successive echoes γ, and the time separation between successive echoes Δt_{echo}; these and the other parameters t_{echo} and t_0 trunc. are explained more fully in [20]. Table I gives the prior ranges we use for the relevant parameters. These are adapted for our purposes from the template bank search performed in [20].

In the ADA model the range for Δt_{echo} is inferred from the published parameters of GW150914 [11], using 50% ranges, and assuming Gaussian distributions. The Kerr metric formula is used for the light travel time between the light ring and a perfectly reflecting surface. This surface is assumed to be at a proper distance one Planck length along Boyer-Lindquist time slices from the Kerr metric event horizon. The parameter γ was chosen to reflect the physical expectation that the amplitude of successive echoes should decrease due to energy loss through one or both of the boundaries. We allow the parameter t_{echo} to vary independently from Δt_{echo} within 1% of its maximum values, and choose an explicit prior for the relative amplitude.

Since the value of the amplitude will have a direct influence on the signal strength, and hence the signal likelihood, its prior range is of central importance to our results. In the template bank search of [20] a prior for the amplitude is not explicitly given. Instead, it is maximised over the template bank. To replicate as closely as
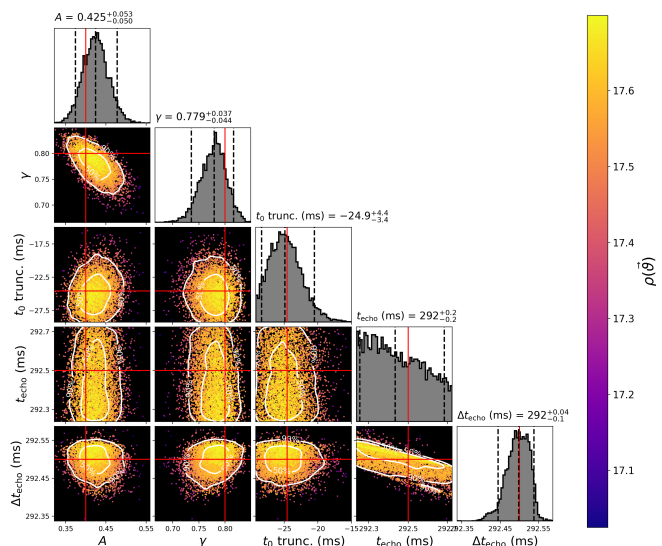

FIG. 1. Posterior on the echo parameters for a loud (SNR ~ 17) simulated signal. The signal has GW150914-like parameters at a fiducial distance of $400\,\mathrm{Mpc}$. An amplitude factor of 0.4 is used for the echoes. Off-diagonal plots show 2D marginal posteriors; the white contours show the 50% and 90% credible regions. Each point represents a random draw from the posterior, colored by the SNR (ρ) at those parameters. The diagonal plots show the 1D marginal posteriors, with the median and 90% credible intervals indicated by the dashed lines. The reported values are the median of the 1D marginal posterior plus/minus the 5/95 percentiles. We see that the injected parameter values, shown by the red lines, are all within the 90% credible intervals. The log Bayes factor for this signal is 140.57.

possible the method of [20] we choose a flat amplitude prior from 10^{-5} to 0.9. This ensures we are sensitive to relatively quiet amplitude signals, although not arbitrarily quiet, and implements the reasonable assumption that the first echo should not be louder than the main signal.

For simplicity we choose to fix the number of echoes to 30. In principle this could be allowed to vary, but for values of γ less than 0.9, 30 echoes capture the main part of the signal that influences the SNR. In testing, we found that varying this number did not change the results substantially.

To establish that our method can correctly identify echo signals in the data, we first test it on simulated echo signals with a variety of different amplitudes. These simulations are added to real detector data, which is made available by the Gravitational Wave Open Science Center (GWOSC) [33, 34], 100 seconds after GW150914. The 100-second delay makes it unlikely that the data at that time is contaminated by a real astrophysical signal [35]. We then apply our method directly to the three binary black hole events in O1: GW150914, LVT151012 and GW151226. Finally, we show how these results can be used to place bounds on the reflectivity of structure that has formed a given distance from the location of the
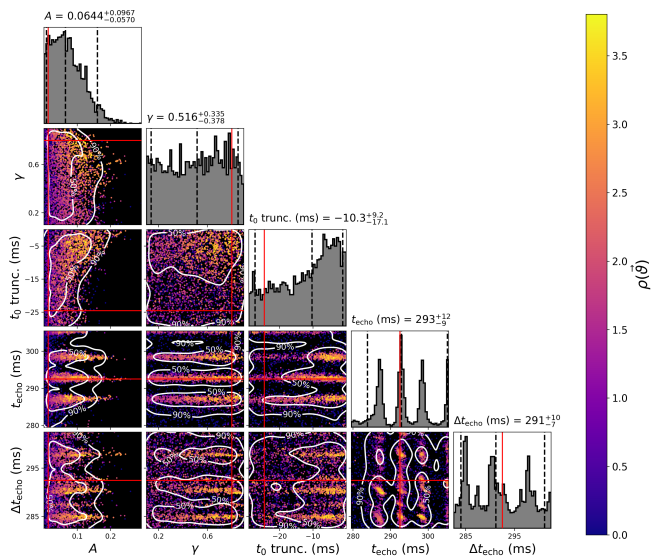
FIG. 2. Posterior on the echo parameters for a quiet simulated signal. The signal has GW150914-like parameters at a fiducial distance of 400 Mpc. An amplitude factor of 0.0125 is used for the echoes. Again, the injected values are shown by the red lines, while points are colored by the SNR at that point in the parameter space. The log Bayes factor for this injection is -1.55, thus indicating what we would expect when the signal is not distinguishable from noise. The prior ranges are largely saturated and lines appear in t_{echo}.

FIG. 3. The 90% credible regions of the 2D marginal posteriors of Δt_{echo} and γ for GW150914-like simulated signals. Shown are a range of echo amplitudes (relative to the peak amplitude of the original signal) A. The injected values are given by the horizontal and vertical red lines. For small values of A, the 90% contour covers most of the prior range, whereas for larger amplitudes the contours narrow down onto the injected values.

would-be horizon.

II. METHODOLOGY AND ANALYSIS PIPELINE

The pipeline we use is based on pycbc_inference [36]. It employs a parallel-tempered MCMC algorithm, emcee_pt [40, 41], to sample the likelihood function for a hypothesis based on the existence of a signal in the data. The likelihood function is chosen to be compatible with the assumption that the underlying noise is Gaussian with a given power spectral density. Once the likelihood has been mapped, the marginalization over the model parameters is performed using thermodynamic integration to obtain a probability for the hypothesis given the data. Although it is known that LIGO data is not Gaussian over long periods of time, over shorter periods it is approximately Gaussian [35]. To account for the non-Gaussianities without a model hypothesis for them, it is possible to sample the Gaussian Bayes factor over many realisations of the true detector noise.

In the results presented here we used 100 Markov chains to sample the likelihood. We require that each chain run for at least five auto-correlation lengths (ACL) beyond 1000 iterations of the sampler. The ACL is measured by averaging parameter samples over all chains, then taking the maximum ACL over all parameters. For the thermodynamic integration of the likelihood function,
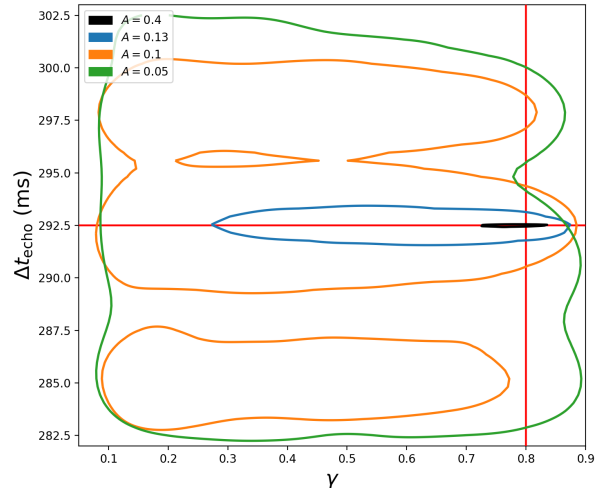
care has to be taken that it is sufficiently sampled both near its peak, but also at lower values of the likelihood. In tests we found that using 16 different temperatures, each placed by inspection, was sufficient to guarantee a consistent value of the Bayes factor. Convergence of this result was checked by running with double the number of temperatures and ensuring that the results were consistent. The posterior distributions are constructed from the coldest temperature chain.

III. INJECTIONS BASED ON GW150914

To test our method we choose to examine simulated echo signals based on GW150914. This is, to date, the loudest binary black hole signal that has been observed via gravitational waves, and should play a central role in constraints derived from the data. Following ADA for simplicity, we choose to fix the base inspiral-merger-ringdown (IMR) waveform to be echoed for both injections and for the search templates. The parameters for these base IMR waveforms are given in the appendix and are obtained from the maximum likelihood results of [36]. The waveforms are constructed using the phenomenological IMR waveform family IMRPhenomPv2 [37, 38] which is freely available as part of LALSuite [39]. These IMR signals are then used to produce echo signals with echo parameters given in Table I. The simulated echo injections are added linearly at varying amplitudes to real de-
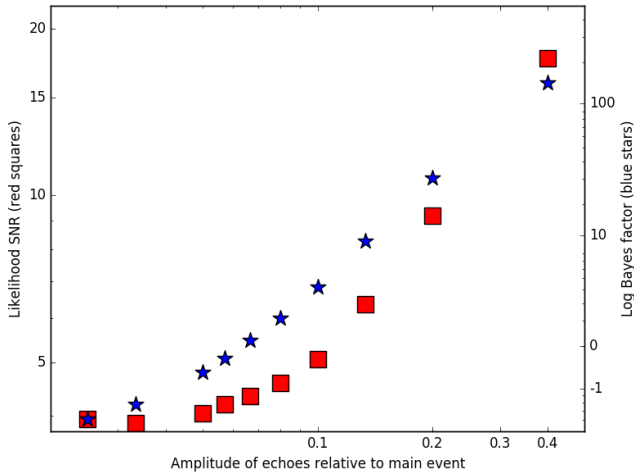
FIG. 4. Values of the maximum likelihood SNR and log Bayes factors for GW150914-based injections with amplitudes from 0.025 to 0.4 at a distance of 400Mpc. A linear fit is possible through the SNR points down to an amplitude around 0.1. The log Bayes factor is negative for amplitude values below ∼ 0.07 (indicating formal preference for Gaussian noise over the echoes hypothesis, although at low absolute values of the log Bayes the data is uninformative). For amplitudes larger than 0.08, the log Bayes factor is greater than 1, indicating positive preference for echoes (the injected signal) over Gaussian noise, by the nomenclature of [43].

FIG. 5. Corner plot for ADA echoes templates in data just after the merger of GW150914. The log Bayes factor for this data is −1.81, indicating a preference the Gaussian noise hypothesis over the Echoes hypothesis. Lines are visible in the t_echo subplots, but the SNR associated with these is still not high. These lines are also seen in tests of the pipeline on simulated Gaussian noise.

tector noise (chosen to be 100 seconds after GW150914, far enough away to be uncontaminated by echo signals or any pre-merger signal). We then attempt to recover them with our analysis pipeline. Example results are shown in Figs 1 and 2.

Figure 1 shows a very loud injection with a relative amplitude of 0.4 and a maximum likelihood SNR of ∼ 17.7. The log Bayes factor for this injection is 140.57, showing a strong preference for the echoes hypothesis over the pure Gaussian noise hypothesis. In this case the echo parameters are well recovered, with the injected values lying within the 90% credible intervals of the marginalised one-dimensional posterior distributions.

Figure 2 shows a much quieter injection with a relative amplitude of 0.0125 and a maximum likelihood SNR of only 3.8. The log Bayes factor for this injection is −1.55 showing a preference for the pure Gaussian noise hypothesis. In this case most echo parameters are not well recovered and their posterior distributions are close to the original prior distributions.

Figure 3 shows the recovery of γ and $\Delta t_{\rm echo}$ for a range of different injection amplitudes. As the amplitude is increased, the recovered value is increasingly constrained to the injected value.

The recovery of signals with different amplitudes is shown in Fig. 4. This figure can be compared with Fig. 4 of [21], which shows the recovery of amplitudes relative to the injected amplitudes. In that work it was found that below a certain injection strength, the recovered echo am-
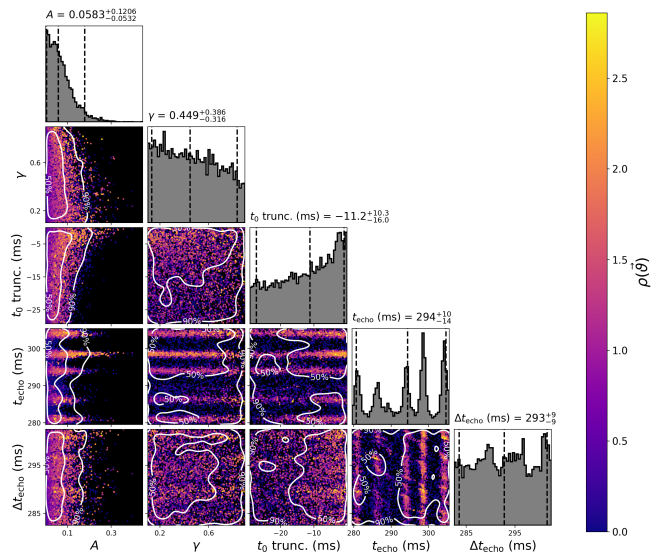
plitude was no longer reliable using the template bank method. Our results here are consistent with that finding. Here we find that below an amplitude of ∼ 0.1 the recovered maximum likelihood SNR no longer falls off linearly and flattens out to an approximately constant value of ∼ 4. At amplitudes below ∼ 0.07 the log Bayes factor becomes negative.

IV. EVENTS IN THE FIRST OBSERVING RUN

The developed pipeline can be run directly on data immediately after the observed GW events (without injections). We show results for the three events of the first LIGO observing run in Table II. This shows that Gaussian noise is favoured over the echoes hypothesis for GW150914 with a log Bayes factor of −1.81. GW150914 is the loudest binary black hole merger yet detected. A corner plot of the posterior distributions for the echo parameters for GW150914 is shown in Fig. 5. The 90% credible interval for the marginalised posterior of the parameter γ is almost as wide as the prior range. The posterior of the amplitude, A, prefers lower values of the amplitude. The posterior for $t_{\rm echo}$ shows distinct lines at certain values of time. These lines are unlikely to be associated with an astrophysical signal and are also seen in tests on simulated Gaussian noise with the same pipeline.

As seen in Table II, both GW151226 and LVT151012 prefer the echoes hypothesis over Gaussian noise, but only marginally. These two events have lower amplitudes

for the main signal than for GW150914 and thus echoes signals with the same relative amplitude would have a lower absolute amplitude relative to the ambient noise [21]. The detector noise is known not to be truly Gaussian for the LIGO detectors [42]. We performed 20 background tests on off-source data that lies before or after the time of LVT151012 at intervals of 50 seconds. Each of these tests is sufficiently separated in time from the others that it will not be contaminated by a common signal. In these background tests, two examples were found with a Bayes factor larger than the result for LVT151012 shown in Table II. A total of four intervals returned Bayes factors that favoured the echo hypothesis over Gaussian noise. Backgrounds for similar (but not identical) echoes hypotheses were also studied in [30] who found evidence for significant tails in the distribution of Bayes factors in real detector noise versus simulated Gaussian noise.

While it is interesting to speculate whether a signal model could be developed that postdicts echo signals for certain events, such as LVT151012, but not for others, such as GW150914, we do not pursue that here. The argument that LVT151012 should be accepted as a genuine binary black hole merger was given recently in [44], however we do not feel that the echoes data for LVT151012 is sufficiently strong to seriously entertain a model where LVT15012-like events display echoes, but GW150914-like events do not.

The SNR values found for the maximum likelihood templates in Table II are comparable, although not identical to those found in [20] and [21]. The values computed here use a slightly modified echo waveform and the finite template spacing in the template banks of [20] and [21] also causes a minor difference. The main differences are the different base IMR waveform employed and the different power spectral density (PSD) used to calculate the matches. The work of [20] and [21] used a PSD directly from [33] whereas here we have used a PSD computed in pycbc [45, 46] using Welch's method. We estimate the PSD by taking the median value over 64 8 second-long segments (each overlapped by 4 seconds), centered on the main event.

With the simplistic hypothesis that all three binary black hole events should show evidence for echo signals in the range of parameters assumed, we can simply add the log Bayes factor together to obtain an overall log Bayes factor for this model relative to Gaussian noise of $-1.81 + 1.25 + 0.42 = -0.14$. This is negative, indicating a preference for Gaussian noise, but not by much. It is worth noting that this simplistic combination assumes that the values for the echo parameters can lie anywhere in their prior ranges for any of the three events. This is slightly different from the hypothesis of [20] that assumes certain echo parameters should have the same value in all three events. With a hypothesis that fixes the values of certain echo parameters to be the same in all cases, it is possible that the overall Bayes factor would be different from our result. But this issue also raises the question of how these common parameters should be fixed; a simple

Event	Log Bayes factor	Max SNR
GW150914	-1.8056	2.86
LVT151012	1.2499	5.5741
GW151226	0.4186	4.07

TABLE II. Table of Bayes factor results. Negative values indicate that the Gaussian noise hypothesis is preferred. Positive values indicate that the echoes hypothesis is preferred after marginalization over parameters. Log Bayes values with magnitude < 1 are "not worth more than a bare mention" in the nomenclature of [43].

maximization of the sum of the squares of the template SNRs as in [20], or as a maximization or marginalization of the likelihood function introduced here. We defer investigation of these subtle issues to future work.

V. DISCUSSION AND CONCLUSIONS

With knowledge of how sensitive our pipeline is from the injection test runs of Sec. III we can determine the amplitude of echoes that would have been detectable had they been present in the data. This allows us to place a bound on the amplitude of echoes emitted from the events considered here. We remind the reader that bounds from our search only relate to the family of echo waveforms considered here. These are based on the model proposed in [20] and adopting the prior ranges of Table I.

As shown in Fig. 5, the posterior amplitude recovery has a 90% confidence interval from $0.0583 + 0.1206 = 0.1789$ to $0.0583 - 0.0532 = 0.0051$. For this realization of the noise, amplitudes above 0.1789 are ruled out at 90% confidence. This is consistent with the injection studies depicted in Fig. 4 which show that (for noise at a different time, 100 seconds after the main event) echo signals with amplitudes $\gtrsim 0.15$ would have been unambiguously identified in the data.

Echo signals of amplitudes 0.1 relative to GW150914 would correspond to approximately 0.1 solar masses of energy being reflected from near the black horizon [24]. Although this value of the amplitude is not conclusively ruled out with the current data, an amplitude as high as 0.2 is conclusively ruled out by our results.

For numerical simulations of systems similar to GW150914 within general relativity, it is estimated that approximately 4 solar masses of gravitational energy flows across the horizon [47]. Our constraints here on the amplitude of echoes within the model of [20] suggest that no more than 5% of this energy is being reflected by near-horizon structure and re-emitted as echoes.

We have seen that there is little evidence of ADA echo-like signals in the data of GW150914. Although there is some evidence of echoes in LVT151012 and GW151226, as both show positive log Bayes factors, this evidence is not very strong. Sampling the true detector noise by

running over off-source times, shows that the log Bayes factor found for LVT151012 is not unusual. A number of improved echo waveform models have been proposed; we defer running with these on further events to future work.

VI. ACKNOWLEDGMENTS

[1] S. W. Hawking and G. F. R. Ellis, doi:10.1017/CBO9780511524646

[2] R. Narayan and J. E. McClintock, arXiv:1312.6698 [astro-ph.HE].

[3] M. Visser, C. Barcelo, S. Liberati and S. Sonego, PoS BHGRS (2008) 010 doi:10.22323/1.075.0010 [arXiv:0902.0346 [gr-qc]].

[4] 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** (2016) no.8, 084031 doi:10.1103/PhysRevD.94.084031 [arXiv:1608.08637 [gr-qc]].

[5] V. Cardoso and P. Pani, "The observational evidence for horizons: from echoes to precision gravitational-wave physics," [arXiv:1707.03021 [gr-qc]].

[6] V. Cardoso, E. Franzin, A. Maselli, P. Pani and G. Raposo, "Testing strong-field gravity with tidal Love numbers," Phys. Rev. D **95**, no. 8, 084014 (2017) Addendum: [Phys. Rev. D **95**, no. 8, 089901 (2017)] doi:10.1103/PhysRevD.95.089901, 10.1103/PhysRevD.95.084014 [arXiv:1701.01116 [gr-qc]].

[7] B. Holdom and J. Ren, "Not quite a black hole," Phys. Rev. D **95** (2017) no.8, 084034 doi:10.1103/PhysRevD.95.084034 [arXiv:1612.04889 [gr-qc]].

[8] J. Polchinski, doi:10.1142/9789813149441_0006 arXiv:1609.04036 [hep-th].

[9] B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], "Observation of Gravitational Waves from a Binary Black Hole Merger," Phys. Rev. Lett. **116** (2016) no.6, 061102 doi:10.1103/PhysRevLett.116.061102 [arXiv:1602.03837 [gr-qc]].

[10] B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], "GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence," Phys. Rev. Lett. **116** (2016) no.24, 241103 doi:10.1103/PhysRevLett.116.241103 [arXiv:1606.04855 [gr-qc]].

[11] B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], "Binary Black Hole Mergers in the first Advanced LIGO Observing Run," Phys. Rev. X **6** (2016) no.4, 041015 doi:10.1103/PhysRevX.6.041015 [arXiv:1606.04856 [gr-qc]].

[12] B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], "GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2," Phys. Rev. Lett. **118**, no. 22, 221101 (2017) doi:doi:10.1103/PhysRevLett.118.221101 [arXiv:1706.01812 [gr-qc]].

[13] B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], "GW170608: Observation of a 19-solar-mass Binary Black Hole Coalescence," Astrophys. J. **851**, no. 2, L35 (2017) doi:10.3847/2041-8213/aa9f0c [arXiv:1711.05578 [astro-ph.HE]].

[14] B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], "GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence," Phys. Rev. Lett. **119**, no. 14, 141101 (2017) doi:10.1103/PhysRevLett.119.141101 [arXiv:1709.09660 [gr-qc]].

[15] J. Aasi *et al.* [LIGO Scientific Collaboration], "Advanced LIGO," Class. Quant. Grav. **32**, 074001 (2015) doi:10.1088/0264-9381/32/7/074001 [arXiv:1411.4547 [gr-qc]].

[16] F. Acernese *et al.* [VIRGO Collaboration], "Advanced Virgo: a second-generation interferometric gravitational wave detector," Class. Quant. Grav. **32**, no. 2, 024001 (2015) doi:10.1088/0264-9381/32/2/024001 [arXiv:1408.3978 [gr-qc]].

[17] B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], "Tests of general relativity with GW150914," Phys. Rev. Lett. **116** (2016) no.22, 221101 doi:10.1103/PhysRevLett.116.221101 [arXiv:1602.03841 [gr-qc]].

[18] M. Cabero, C. D. Capano, O. Fischer-Birnholtz, B. Krishnan, A. B. Nielsen, A. H. Nitz and C. M. Biwer, "Observational tests of the black hole area increase law," Phys. Rev. D **97** (2018) no.12, 124069 doi:10.1103/PhysRevD.97.124069 [arXiv:1711.09073 [gr-qc]].

[19] A. B. Nielsen and O. Birnholtz, "Testing pseudocomplex general relativity with gravitational waves," Astron. Nachr. **339** (2018) no.4, 298 doi:10.1002/asna.201813473 [arXiv:1708.03334 [gr-qc]].

[20] 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) doi:10.1103/PhysRevD.96.082004 [arXiv:1612.00266v2 [gr-qc]].

[21] J. Westerweck, A. B. Nielsen, O. Fischer-Birnholtz, M. Cabero, C. Capano, T. Dent, B. Krishnan,

G. D. Meadors and A. Nitz, "Low significance of evidence for black hole echoes in gravitational wave data," Phys. Rev. D **97**, no. 12, 124037 (2018) doi:10.1103/PhysRevD.97.124037 [arXiv:1712.09966 [gr-qc]].

[22] R. S. Conklin, B. Holdom and J. Ren, Phys. Rev. D **98** (2018) no.4, 044021 doi:10.1103/PhysRevD.98.044021 [arXiv:1712.06517 [gr-qc]].

[23] V. Cardoso, E. Franzin and P. Pani, "Is the gravitational-wave ringdown a probe of the event horizon?," Phys. Rev. Lett. **116** (2016) no.17, 171101 Erratum: [Phys. Rev. Lett. **117** (2016) no.8, 089902] doi:10.1103/PhysRevLett.117.089902, 10.1103/PhysRevLett.116.171101 [arXiv:1602.07309 [gr-qc]].

[24] G. Ashton, O. Birnholtz, M. Cabero, C. Capano, T. Dent, B. Krishnan, G. D. Meadors, A. B. Nielsen, A. Nitz and J. Westerweck, "Comments on: "Echoes from the abyss: Evidence for Planck-scale structure at black hole horizons"," [arXiv:1612.05625 [gr-qc]].

[25] J. Abedi, H. Dykaar and N. Afshordi, "Echoes from the Abyss: The Holiday Edition!," [arXiv:1701.03485 [gr-qc]].

[26] J. Abedi, H. Dykaar and N. Afshordi, "Comment on: "Low significance of evidence for black hole echoes in gravitational wave data"," [arXiv:1803.08565 [gr-qc].

[27] Z. Mark, A. Zimmerman, S. M. Du and Y. Chen, "A recipe for echoes from exotic compact objects," Phys. Rev. D **96** (2017) no.8, 084002 doi:10.1103/PhysRevD.96.084002 [arXiv:1706.06155 [gr-qc]].

[28] H. Nakano, N. Sago, H. Tagoshi and T. Tanaka, "Black hole ringdown echoes and howls," PTEP **2017** (2017) no.7, 071E01 doi:10.1093/ptep/ptx093 [arXiv:1704.07175 [gr-qc]].

[29] J. Veitch *et al.*, "Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library," Phys. Rev. D **91**, no. 4, 042003 (2015) doi:10.1103/PhysRevD.91.042003 [arXiv:1409.7215 [gr-qc]].

[30] R. Ka Lok Lo *et al.*, "Template-based Gravitational-Wave Echoes Search Using Bayesian Model Selection," [arXiv:1811.07431 [gr-qc].

[31] K. W. Tsang *et al.*, "A morphology-independent data analysis method for detecting and characterizing gravitational wave echoes," Phys. Rev. D **98**, no. 2, 024023 (2018) doi:10.1103/PhysRevD.98.024023 [arXiv:1804.04877 [gr-qc]].

[32] E. Barausse, R. Brito, V. Cardoso, I. Dvorkin and P. Pani, Class. Quant. Grav. **35** (2018) no.20, 20LT01 doi:10.1088/1361-6382/aae1de [arXiv:1805.08229 [gr-qc]].

[33] LIGO Scientific Collaboration, "LIGO Open Science Center - Data Releases for Observed Transients", 2017, doi:10.7935/K5MW2F23, 10.7935/K5CC0XMZ, 10.7935/K5H41PBP, 10.7935/K53X84K2 http://losc.ligo.org/events

[34] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein and B. Stephens, "The LIGO Open Science Center," J. Phys. Conf. Ser. **610**, no. 1, 012021 (2015) doi:10.1088/1742-6596/610/1/012021 [arXiv:1410.4839 [gr-qc]].

[35] A. B. Nielsen, A. H. Nitz, C. D. Capano and D. A. Brown, "Investigating the noise residuals around the gravitational wave event GW150914," [arXiv:1811.04071 [astro-ph.HE].

[36] C. M. Biwer, C. D. Capano, S. De, M. Cabero, D. A. Brown, A. H. Nitz and V. Raymond, "PyCBC Inference: A Python-based parameter estimation toolkit for compact binary coalescence signals," [arXiv:1807.10312 [astro-ph.IM].

[37] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jimnez Forteza and A. Boh, "Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era," Phys. Rev. D **93** (2016) no.4, 044007 doi:10.1103/PhysRevD.93.044007 [arXiv:1508.07253 [gr-qc]].

[38] M. Hannam, P. Schmidt, A. Boh, L. Haegel, S. Husa, F. Ohme, G. Pratten and M. Pürrer, "Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms," Phys. Rev. Lett. **113** (2014) no.15, 151101 doi:10.1103/PhysRevLett.113.151101 [arXiv:1308.3271 [gr-qc]].

[39] LIGO Scientific Collaboration (2018), LIGO Algorithm Library, https://doi.org/10.7935/GT1W-FZ16, https://git.ligo.org/lscsoft/lalsuite .

[40] D. Foreman-Mackey, D. W. Hogg, D. Lang and J. Goodman, "emcee: The MCMC Hammer," Publ. Astron. Soc. Pac. **125** (2013) 306 doi:10.1086/670067 [arXiv:1202.3665 [astro-ph.IM]].

[41] W. D. Vousden, W. M. Farr, and I. Mandel, Mon. Not. Roy. Astron. Soc. **455** (2016) 1919 doi:10.1093/mnras/stv2422 [arXiv:1501.05823 [astro-ph.IM]].

[42] B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], "Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914," Class. Quant. Grav. **33** (2016) no.13, 134001 doi:10.1088/0264-9381/33/13/134001 [arXiv:1602.03844 [gr-qc]].

[43] R. E. Kass and A. E. Raftery, "Bayes Factors". Journal of the American Statistical Association. 90 (1995) 791. doi:10.2307/2291091.

[44] A. H. Nitz, C. Capano, A. B. Nielsen, S. Reyes, R. White, D. A. Brown and B. Krishnan, "1-OGC: The first open gravitational-wave catalog of binary mergers from analysis of public Advanced LIGO data," [arXiv:1811.01921 [gr-qc].

[45] T. Dal Canton *et al.*, "Implementing a search for aligned-spin neutron star-black hole systems with advanced ground based gravitational wave detectors," Phys. Rev. D **90** (2014) no.8, 082004 doi:10.1103/PhysRevD.90.082004 [arXiv:1405.6731 [gr-qc]].

[46] S. A. Usman *et al.*, "The PyCBC search for gravitational waves from compact binary coalescence," Class. Quant. Grav. **33** (2016) no.21, 215004 doi:10.1088/0264-9381/33/21/215004 [arXiv:1508.02357 [gr-qc]].

[47] A. Gupta, B. Krishnan, A. Nielsen and E. Schnetter, "Dynamics of marginally trapped surfaces in a binary black hole merger: Growth and approach to equilibrium," Phys. Rev. D **97** (2018) no.8, 084028 doi:10.1103/PhysRevD.97.084028 [arXiv:1801.07048 [gr-qc]].

Appendix: Fiducial IMR waveform parameters

Parameter	GW150914	LVT151012	GW151226
mass1	39.03	22.87	18.80
mass2	32.06	18.67	6.92
spin1x	-0.87	0.12	0.44
spin1y	-0.43	0.19	0.59
spin1z	-0.06	-0.20	0.33
spin2x	-0.11	0.018	0.00
spin2y	-0.03	-0.019	-0.017
spin2z	-0.15	0.062	0.0033
distance	477	751	315
ra	1.57	0.65	2.23
dec	-1.27	0.069	0.98
tc	1126259462.42	1128678900.46	1135136350.66
polarization	5.99	5.64	1.43
inclination	2.91	2.32	0.68
coa_phase	0.69	4.44	1.64
phase_shift	-0.92	-0.91	1.86

We list here the parameters of the base IMR waveforms used to construct the echo templates both for injections and for the searches. These values are obtained from the maximum likelihood values of [36].