Echoes from the abyss: Tentative evidence for Planck-scale structure at black hole horizons

Jahed Abedi,[1,2,3,*] Hannah Dykaar,[4,5] and Niayesh Afshordi[3,5,†]

[1]*Department of Physics, Sharif University of Technology, P.O. Box 11155-9161 Tehran, Iran*
[2]*School of Particles and Accelerators, Institute for Research in Fundamental Sciences (IPM), P.O. Box 19395-5531 Tehran, Iran*
[3]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada*
[4]*Department of Physics, McGill University, 3600 rue University, Montreal, Quebec H3A 2T8, Canada*
[5]*Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

(Received 5 March 2017; published 26 October 2017)

In classical general relativity (GR), an observer falling into an astrophysical black hole is not expected to experience anything dramatic as she crosses the event horizon. However, tentative resolutions to problems in quantum gravity, such as the cosmological constant problem, or the black hole information paradox, invoke significant departures from classicality in the vicinity of the horizon. It was recently pointed out that such near-horizon structures can lead to late-time echoes in the black hole merger gravitational wave signals that are otherwise indistinguishable from GR. We search for observational signatures of these echoes in the gravitational wave data released by the advanced Laser Interferometer Gravitational-Wave Observatory (LIGO), following the three black hole merger events GW150914, GW151226, and LVT151012. In particular, we look for repeating damped echoes with time delays of $8M\log M$ (+spin corrections, in Planck units), corresponding to Planck-scale departures from GR near their respective horizons. Accounting for the "look elsewhere" effect due to uncertainty in the echo template, we find tentative evidence for Planck-scale structure near black hole horizons at false detection probability of 1% (corresponding to 2.5σ.[1] significance level). Future observations from interferometric detectors at higher sensitivity, along with more physical echo templates, will be able to confirm (or rule out) this finding, providing possible empirical evidence for alternatives to classical black holes, such as in *"firewall"* or *"fuzzball"* paradigms.

DOI: 10.1103/PhysRevD.96.082004

I. INTRODUCTION

There is mounting, albeit controversial, theoretical evidence that quantum black holes might be significantly different from their classical counterparts, even in the regime where semiclassical gravity is expected to be valid. Such strong modifications may exist, not only due to nonperturbative quantum gravitational effects [1–4], but also at the level of semiclassical approximation [5,6]. In particular, modern versions of Hawking's black hole information paradox have led to exotic alternatives to classical black hole horizons such as fuzzball [2,3] and firewall paradigms [1,7]. These should form by Page time $\sim M^3$, but may emerge as early as the "scrambling time" $\sim M\log M$ [8,9], where M is the black hole mass in Planck units.

On more phenomenological grounds, it has been proposed that a wholesome solution to the (old and new) cosmological constant problems replaces the black hole horizons by a Planck-scale quantum barrier, which could naturally explain the observed scale of dark energy [10]. Furthermore, accretion into these "black holes" offers a possible origin for observed ultrahigh energy IceCube neutrinos [11].

In this *paper*, we search for possible signatures of quantum gravitational alternatives to black hole horizons in the gravitational wave data releases of black hole mergers observed by the advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) [12–14]. As a simple toy model, we replace the event horizon by a mirror (with Dirichlet boundary conditions) at $\sim$ Planck proper length outside the horizon. This picture is motivated by the realization that a thermal membrane on the stretched horizon, satisfying Israel junction conditions with Z_2 symmetry, happens to have a thermal entropy equal to the Bekenstein-Hawking area law [15]. Therefore, any horizonless microscopic model of the black hole which accounts for its entropy, should act as a mirror, at least for linear long wavelength perturbations. The mirror is not perfect for particles with $\omega \gg T_H$ (=Hawking temperature), as they can excite the microstates of the system, and thus be

*jahed_abedi@physics.sharif.ir
†nafshordi@pitp.ca

[1]In this paper, we use 2-tailed Gaussian probability to assign a significance to a p-value, e.g., $1-\text{p-value}=68\%$ and 95% correspond to 1σ and 2σ, respectively.

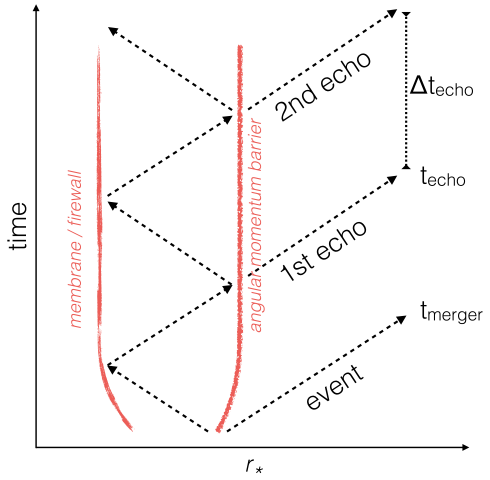

FIG. 1. Spacetime depiction of gravitational wave echoes from a membrane/firewall on the stretched horizon, following a black hole merger event.

absorbed by the membrane [16], but should be reflective at $\omega \lesssim T_H$ as these microstates cannot be excited. Incidentally, this is the frequency regime for gravitational waves in the ringdown phase of black hole mergers. In contrast, electromagnetic emissions from accretion into black holes are at much higher frequencies, where the membrane is expected to be highly absorbing, consistent with astrophysical observations [17,18] (but also see [11,19]).

In spite of its simplicity, this picture is remarkably robust: As first noticed in [20,21], introduction of structure near event horizon leads to late, repeating, echoes of the ringdown phase of the black hole merger, due to waves trapped between the near-horizon structure and the angular momentum barrier (Fig. 1). This is relatively insensitive to the nature of the structure, or indeed how one defines the Planck length, l_p, as the time for reflection from the stretched horizon is only logarithmically dependent on its distance from the event horizon, i.e. $\Delta t_{\rm echo} = 8M\log(M/l_p)$ (+spin corrections; see below). As a result, e.g., an order of magnitude change in this distance only affects the time of the echoes at 2%–3% level. While $\Delta t_{\rm echo}$ is determined by linear physics, the time between the main merger event and the first echo could be further affected by nonlinear physics during merger, i.e. $t_{\rm echo} - t_{\rm merger} = \Delta t_{\rm echo} + \mathcal{O}(M)$ (see Fig. 1), or equivalently,

$$\frac{t_{\rm echo} - t_{\rm merger}}{\Delta t_{\rm echo}} = 1 \pm \mathcal{O}(1\%), \tag{1}$$

where $\Delta t_{\rm echo}$ is predicted from the final (redshifted) mass and spin measurements for each event.

Quite surprisingly, we find statistical evidence for these delayed echoes in LIGO events: GW150914, GW151226, and LVT151012 at a false detection probability of 1% or combined significance of 2.5σ. We shall first describe our theoretical framework for the echoes, and then our statistical methodology and results.

II. ECHO TIME DELAYS

At the linear order, perturbed black holes are described by quasi-normal modes (QNM's) which satisfy the boundary conditions of purely outgoing waves at infinity and purely ingoing waves at the horizon. The transition (from ingoing to outgoing) takes place continuously at the peak of the black hole angular momentum potential barrier.

In our case, the ingoing modes of the ringdown reflect back from the membrane (e.g., fuzzball or firewall) near horizon and pass back through the potential barrier. Part of the wave goes to infinity with a time delay. We call this the first echo (see Fig. 1). This time delay corresponds to twice the tortoise coordinate distance between the peak of the angular momentum barrier ($r_{\rm max}$) and the membrane (which diverges logarithmically if the membrane approaches the horizon). The remaining part of the 1st echo returns back towards the membrane and the process repeats itself.[2] Assuming Dirichlet boundary conditions at the membrane (discussed above), the reflected waves must be phase inverted, i.e. even echoes have opposite phase with respect to the odd ones (a similar phase flip pattern is also observed in [21]).

For Kerr black hole with dimensionless spin parameter a, this implies

$$\begin{aligned}\Delta t_{\rm echo} &= 2\times r_*|_{r_+ +\Delta r}^{r_{\rm max}} = 2\times\int_{r_+ +\Delta r}^{r_{\rm max}} \frac{r^2+a^2M^2}{r^2-2Mr+a^2M^2}dr \\ &= 2r_{\rm max} - 2r_+ - 2\Delta r + 2\frac{r_+^2+a^2M^2}{r_+-r_-}\ln\left(\frac{r_{\rm max}-r_+}{\Delta r}\right) \\ &\quad -2\frac{r_-^2+a^2M^2}{r_+-r_-}\ln\left(\frac{r_{\rm max}-r_-}{r_+-r_-+\Delta r}\right),\end{aligned} \tag{2}$$

where $r_\pm = M(1\pm\sqrt{1-a^2})$, and Δr is the coordinate distance of the membrane and the (would-be) horizon.

The peak of the angular momentum barrier, $r_{\rm max}$, is given by the roots of a sixth-order polynomial [24]:

$$\begin{aligned}&(1-\mu^2)[(2-\mu^2)\hat{r}_{\rm max}^2 + 2(2+\mu^2)\hat{r}_{\rm max} + (2-\mu^2)]a^4 \\ &\quad + 4\hat{r}_{\rm max}^2[(1-\mu^2)\hat{r}_{\rm max}^2 - 2\hat{r}_{\rm max} - 3(1-\mu^2)]a^2 \\ &\quad + 2\hat{r}_{\rm max}^4(\hat{r}_{\rm max}-3)^2 = 0,\end{aligned} \tag{3}$$

where $\mu = m/(l+\frac{1}{2})$ and $\hat{r}_{\rm max} = r_{\rm max}/M$. For the dominant QNM, $r_{\rm max} < 3M$ and $(l,m) = (2,2)$ resulting in $\mu = 0.8$.

We further posit that the location of the membrane should be near a Planck proper length from the horizon. This assumption is required to explain the observed density of cosmological dark energy within the gravitational aether

[2]Also, note that due to the different boundary conditions near the horizon (compared to the classical picture) there exist a completely different QNM spectrum. A coherent superposition of a large number of these modes is responsible for creating echoes [20,22,23].

proposal [10], but is also expected from generic quantum gravity scalings, such as the brick wall model [25], or trans-Planckian effects [26,27]. This implies

$$\int_{r_+}^{r_+ + \Delta r} \sqrt{g_{rr}} dr|_{\theta=0} \sim l_p \simeq 1.62 \times 10^{-33} \text{ cm}, \quad (4)$$

which fixes the location of the membrane:

$$\Delta r|_{\theta=0} = \frac{\sqrt{1-a^2} l_p^2}{4M(1+\sqrt{1-a^2})}. \quad (5)$$

With this setup, we note that $\Delta t_{\rm echo} \simeq 8M \log(M/l_p) \times [1 + \mathcal{O}(a^2)]$ is comparable to the scrambling time: the time over which the black hole state is expected to thermalize [8,9,28,29].

Using the measurements of the final black hole (redshifted) mass and spin by the LIGO Collaboration, we can constrain $\Delta t_{\rm echo}$ for each merger event. Assuming Gaussian errors, we find (see Appendix D for details of calculations):

$$\Delta t_{{\rm echo},I}(\text{sec}) = \begin{cases} 0.2925 \pm 0.00916 & I = \text{GW150914} \\ 0.1013 \pm 0.01152 & I = \text{GW151226} \\ 0.1778 \pm 0.02789 & I = \text{LVT151012} \end{cases} \quad (6)$$

III. DATA AND THE ECHO TEMPLATE

In this analysis, we use four data sets for each event. The first two are the theoretical best-fit waveform for Hanford and Livingston detectors (in real time series) for the BH merger event, provided by the LIGO and Virgo Collaborations [30–33]. The other two are the observed strain datastream from the two detectors. We call these $M_{H,I}(t)$, $M_{L,I}(t)$, $h_{H,I}(t)$ and $h_{L,I}(t)$, respectively. We used the strain data at 4096 Hz and for 32 sec duration. The waveform model consists of three phases: inspiral, merger, and ringdown.

Following the numerical results of [20,21], we construct a phenomenological gravitational wave template for the echoes using five free parameters:

(1) $\Delta t_{\rm echo}$ is the time interval in between successive echoes, which we vary within the 1σ range, fixed by the uncertainties in (redshifted) mass and angular momentum of the final black hole [Eq. (6)].
(2) $t_{\rm echo}$ is the time of arrival of the first echo, which can be affected by nonlinear dynamics near merger [Eq. (1)].
(3) t_0 quantifies which part of the GR merger template is truncated to produce the subsequent echo templates[3] To do this, we introduce a smooth cutoff function with a free parameter t_0,

[3]Note that the wavelength of gravitational waves in the inspiral phase is much longer than the size of the black holes, which leads to an echo signal suppressed at 4PN order [34].

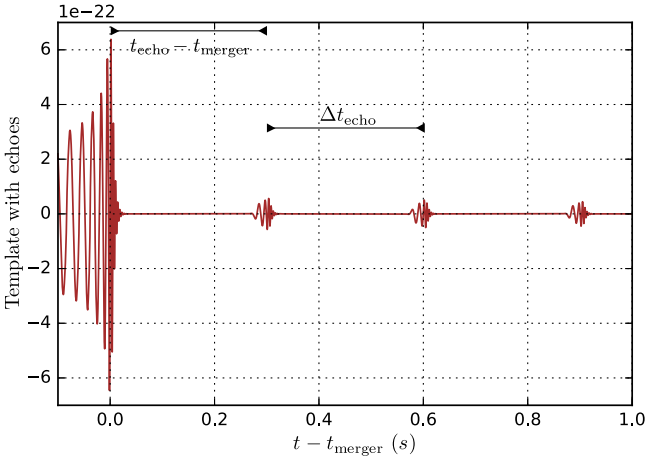

FIG. 2. LIGO original template for GW150914, along with our best-fit template for the echoes.

$$\Theta_I(t, t_0) \equiv \frac{1}{2}\left\{1 + \tanh\left[\frac{1}{2}\omega_I(t)(t - t_{\rm merger} - t_0)\right]\right\}, \quad (7)$$

where $\omega_I(t)$ is frequency of model as a function of time [35] and $t_{\rm merger}$ is the time at which the GR template peaks. As the intermediate region (merger) is before $t_{\rm merger}$, we assume t_0 is negative, and vary it within the range $t_0 \in (-0.1, 0)\overline{\Delta t}_{\rm echo}$. Using this definition, we can define the truncated template:

$$\mathcal{M}_{T,I}^{H/L}(t, t_0) \equiv \Theta_I(t, t_0)\mathcal{M}_I^{H/L}(t). \quad (8)$$

(4) γ is damping factor of successive echoes, which should be between 0 and 1. In our analysis, we vary this free parameter within the range (0.1,0.9) and look for the best fit.
(5) A is the over-all amplitude of the echo template (with respect to the main event) which we fit for, assuming a flat prior.

The truncated model with echoes and all the free parameters is then given by

$$M_{TE,I}^{H/L}(t) \equiv A\sum_{n=0}^{\infty}(-1)^{n+1}\gamma^n \mathcal{M}_{T,I}^{H/L}(t + t_{\rm merger} - t_{\rm echo} - n\Delta t_{\rm echo}, t_0), \quad (9)$$

where the term $(-1)^{n+1}$ is due to the phase inversion of the truncated model in each reflection. Figure (2) shows our best fit for this template for GW150914 within the parameter space described above, along with the main merger event. In the frequency domain we expect to see resonances of these echoes (Fig. 3).

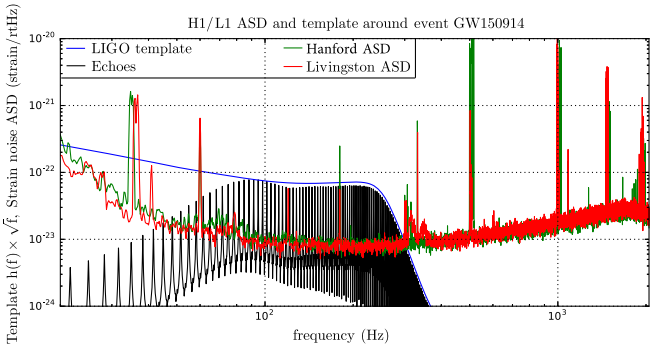

FIG. 3. Amplitude spectral densities (ASD's) of our best-fit echo template [Eq. (9)] and the main event, for GW150914. Since we have a quasi-periodic model, there are resonances in the spectrum. The ASDs are the square root of the power spectral densities, which are averages of the square of the fast Fourier transforms of the data. The noise spectra from Hanford and Livingston detectors are also shown.

IV. RESULTS

Our strategy is to search for the best fit for the echo template (9), by maximizing its signal-to-noise ratio, SNR, within the conjectured parameter space described above at fixed $x = (t_{\rm echo} - t_{\rm merger})/\Delta t_{\rm echo}$. We then identify the highest peak within the range $t_{\rm echo} - t_{\rm merger} = (1 \pm 0.01)\Delta t_{\rm echo}$ [Eq. (1)]. This range in $t_{\rm echo}$ is expected, e.g., due to a random phase in the complex echo template (see [36] and Appendix C). We quantify the significance of this peak by how often a higher SNR peak is achieved within an interval of duration $2\% \times \overline{\Delta t}_{{\rm echo},I}$, in the background (away from the main event) in the data stream, where $\overline{\Delta t}_{{\rm echo},I}$ is the mean of $\Delta t_{{\rm echo},I}$ for independent events in Eq. (6). It is worth noting that due to different angles and locations of each detector, a complex model is analyzed. Therefore, in the calculation of SNRs, we subtracted the phase of the main event from complex template and obtained two real templates corresponding to each detector (Hanford/Livingston). Then we set the original gravitational wave peak at $t = 0$ by removing the offset from SNRs (see [30–33]). We combine the SNRs of different detectors for each event by adding the χ^2 for two data sets, using the same echo model.[4]

We do the analysis once for GW150914 (LIGO's most significant detection), and repeat it for the three recorded events combined, by maximizing

$$\mathrm{SNR}^2_{\rm total} \equiv \sum_I \mathrm{SNR}^2_I. \tag{10}$$

[4]The proper way of combining data sets is through inverse weighting by noise[2]. We use the analysis packages provided by the LIGO Open Science Center (https://losc.ligo.org) to compute SNR for a given template and combine it for two detectors.

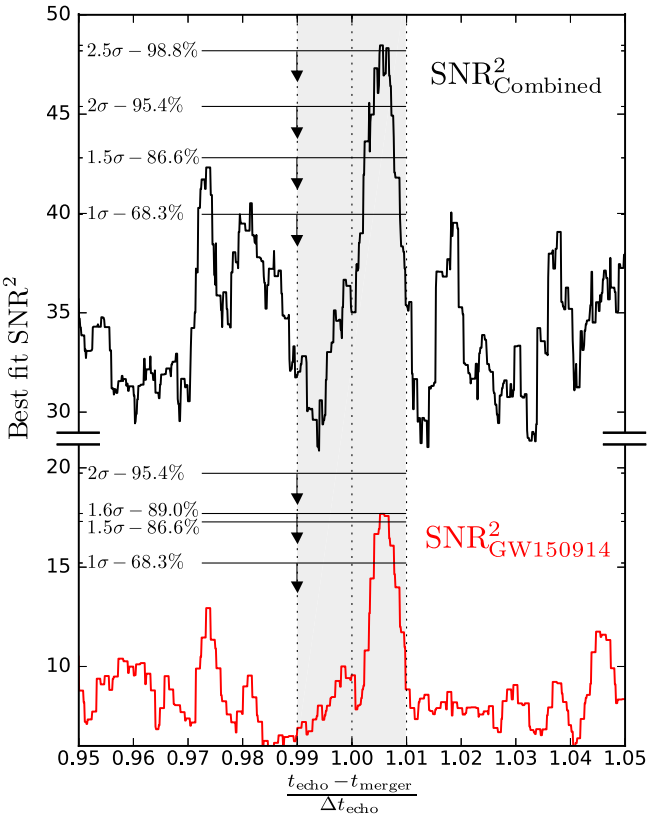

FIG. 4. Best-fit (or maximum) SNR^2 near the expected time of merger echoes (Eq's. (1) and (6), for the combined (top) and GW150914 (bottom) events. The significance of the peaks is quantified by the p-value of their $\mathrm{SNR}_{\rm max}$ within the gray rectangle (see Appendix E for detail of calculation).

In doing so, we assume the same γ and $t_0/\overline{\Delta t}_{\rm echo}$ for all three events, while keeping $\Delta t_{\rm echo}$ and A's independent. The results are shown in Figs. 4 and 5 and Tables I–II.

Figure 4 shows that there is indeed a significant peak with $\mathrm{SNR}_{\rm max} = 4.21(6.96)$ for echoes following the GW150914 (combined) merger event(s), within 0.54% of the predicted echo time delay. To find the significance of

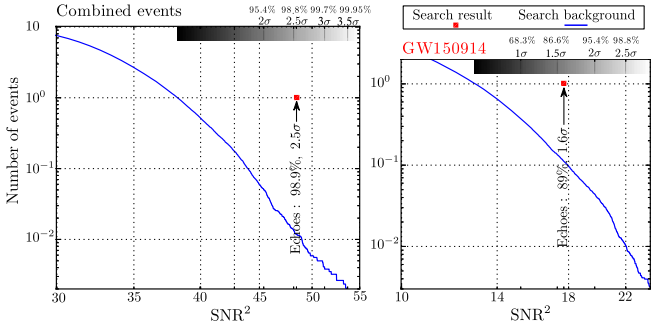

FIG. 5. Average number of noise peaks higher than a particular SNR-value within a time interval $2\% \times \overline{\Delta t}_{\rm echo}$ for combined (left) and GW150914 (right) events. The red dots show the observed SNR peak at $t_{\rm echo} = 1.0054\Delta t_{\rm echo}$ (Fig. 4). The horizontal bar shows the correspondence between SNR values and their significance.

TABLE I. Best-fit values for echo parameters of the highest SNR peak near the predicted $\Delta t_{\rm echo}$, and their significance.

	Range	GW150914	Combined
$(t_{\rm echo} - t_{\rm merger})/\Delta t_{\rm echo}$	(0.99,1.01)	1.0054	1.0054
γ	(0.1,0.9)	0.89	0.9
$t_0/\overline{\Delta t}_{\rm echo}$	$(-0.1, 0)$	-0.084	-0.1
Amplitude[a]		0.0992	0.124
$\rm SNR_{max}$		4.21	6.96
p-value		0.11	0.011
Significance		1.6σ	2.5σ

[a]The combined amplitude is given by $A_{\rm average} = \frac{\sum_I \frac{SNR_I^2}{|A_I|}}{\sum_I \frac{SNR_I^2}{A_I^2}}$.

TABLE II. Theoretical expectations for $\Delta t_{\rm echo}$'s of each merger event [Eq. (6)], compared to their best combined fit within the 1σ credible region, and the contribution of each event to the joint SNR for the echoes [Eq. (10)].

	GW150914	GW151226	LVT151012
$\Delta t_{\rm echo,pred}$ (sec)	0.2925	0.1013	0.1778
	$\pm 0.009\,16$	$\pm 0.011\,52$	$\pm 0.027\,89$
$\Delta t_{\rm echo,best}$ (sec)	0.300 68	0.097 58	0.190 43
$\lvert A_{\rm best,I}\rvert$	0.091	0.33	0.34
$\rm SNR_{best,I}$	4.13	3.83	4.52

finding this peak so close to the predicted value, we divide up the data steam (within the range $9\text{–}38 \times \overline{\Delta t}_{{\rm echo},I}$ after the merger) into intervals of $2\% \times \overline{\Delta t}_{{\rm echo},I}$, and compute the average number of points in the interval that exceed SNR (Fig. 5). This yields an estimate of the false detection probability and the significance of SNR peaks observed near the predicted echo times, at 0.11 (0.011) and 1.6σ (2.5σ) for the GW150914 (combined) merger event(s) respectively. More discussion of our statistical methodology, and possible alternatives, can be found in Appendix and [36,37].

V. CONCLUSIONS AND DISCUSSION

In this *paper*, we have searched advanced LIGO's public data release of the first observed gravitational wave signals from black hole merger events for signatures of Planck-scale structure near their event horizons. By building a phenomenological template for successive echoes from such exotic structures expected in e.g., *firewall* or *fuzzball* paradigms, after marginalizing over its parameters, we report a first tentative evidence for these echoes at false detection probability of 0.011 or 2.5σ significance in LIGO data proceeding its reported merger events. Future data releases from LIGO Collaboration at higher sensitivity will be able to definitively confirm or rule out this finding.

One may wonder how including GW151226 and LVT151012 may improve the significance of the echo signal, even though their ringdown phase was not detectable in LIGO data. We should note that while GW150914 and the two other events have similar numerical contributions to the significance, the nature of their contributions are quite different: GW150914 appears to pin down the echo template parameters, while the others help improve evidence for this template. Furthermore, the repeating nature of the echoes gives them a low frequency structure (Fig. 3) which may be detectable, even if the ringdown itself falls below the detector noise at high frequencies for GW151226 and LVT151012 (Fig. 1 in [12,13]).

We should note that the *ad hoc* nature of our echo template construction is not entirely satisfactory and could lead to some ambiguity in interpreting the statistical significance of our finding. In particular, the fact that the combined SNR is maximized on the edge of our parameter range (see Table I) points to a need for a better physical prior on parameters, or simply a more physical echo template. For instance, increasing the range of prior for t_0 adds a significant portion of the inspiral into the echo template, which may suggest a need for a less *ad hoc* truncation function. This does not change the statistical significance of our SNR peaks, but suggests better fits may lie beyond this range (see Appendix for more discussions). In addition, reliable extension of the analysis beyond this range (in particular $\gamma > 0.9$) requires analyzing a much larger portion of LIGO data, where one may also worry about the time variability and non-Gaussianity of the LIGO instrumental noise (see Figs. 14–15 in [38] and Appendix).

From a physical standpoint, a slowly damping echo, $\gamma \approx 1$, may not be unexpected and could be intimately related to the well-established instability of horizonless ergoregions [39,40]. Future numerical simulations of merging black holes with a membrane can sharpen the echo template, possibly increasing the detection significance. We thus predict that a synergy of improvements in observational sensitivity and theoretical modelling can provide conclusive evidence for quantum gravitational alternatives to black hole horizons.

ACKNOWLEDGMENTS

We thank Vitor Cardoso, Will East, Sabine Hossenfelder, Matt Johnson, Luis Lehner, Rafael Sorkin, Nico Yunes, Qingwen Wang, and Aaron Zimmerman for helpful comments and discussions. We especially thank the authors of [37] for their constructive criticism, as well as the anonymous referees for the valuable comments and suggestions to improve the clarity and the quality of this paper. We also thank all the participants in our weekly group meetings for their patience during our discussions, as well as Jonah Miller and Erik Schnetter for computational help in this project. J. A. would like to thank Perimeter Institute for Theoretical Physics (PI) for the great hospitality during the course of this work on his collaboration leave, as well as his research adviser Hessamaddin Arfaei for all his support. He

also thanks Mansour Karami, Seyed Faroogh Moosavian, and Yasaman Yazdi for their kind hospitality during his visit. This research has made use of data, software, and web tools obtained from the LIGO Open Science Center (https://losc.ligo.org), a service of LIGO Laboratory and the LIGO Scientific Collaboration. LIGO is funded by the U.S. National Science Foundation. This work has been partially supported by the Ministry of Science, Research and Technology of Iran, Institute for Research in Fundamental Sciences (IPM), University of Waterloo, Natural Sciences and Engineering Research Council of Canada, and Perimeter Institute for Theoretical Physics (PI). Research at PI is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

Note Added.—Since we submitted this paper, several authors have proposed alternative templates and/or search methodologies for "echoes from the abyss." For example, [41] used a Green's function method to find scalar echoes generated by a point mass on a marginally unstable circular orbit around a Schwarzschild black hole surrounded by a (partially) reflective wall. In contrast, [42] used the reflectivity of the angular momentum barrier for a Kerr black hole to construct a phenomenological echo template. While these treatments included more realistic elements of wave propagation (compared to our approach), they still miss two crucial pieces of physics, namely a frequency-dependent reflectivity for the fuzzball/firewall (see I above), and realistic initial conditions that should ideally come from nonlinear simulations of binary mergers. Therefore, it is not clear whether these templates are significantly more realistic than a phenomenological model such as the one adopted in III above.

Similar to our approach, Maselli *et al.* [43] provide a phenomenological echo template based on superposition of sine-Gaussians with free parameters, and forecast how well they can be measured using Advanced LIGO observations. It will be interesting to see the application of this method to real data and whether it can recover similar tentative evidence for echoes.

APPENDIX A: ESTIMATION OF THE TOTAL ECHO ENERGY EMISSION

With a simple assumption we may be able to estimate the total energy emitted in echoes. Due to physical reasons the inspiral part is removed and each echo has energy equivalent to $\gamma^{2n}E_{\text{Ringdown Merger}} = \gamma^{2n}\xi E_{\text{Total}}$. Where $0 < \xi < 1$ is the fraction of the energy in ringdown and merger part. With this given assumption we obtain

$$\frac{E^I_{\text{echoes}}}{E^I_{\text{Ringdown Merger}}} = \frac{A_I^2}{1-\gamma^2}. \tag{A1}$$

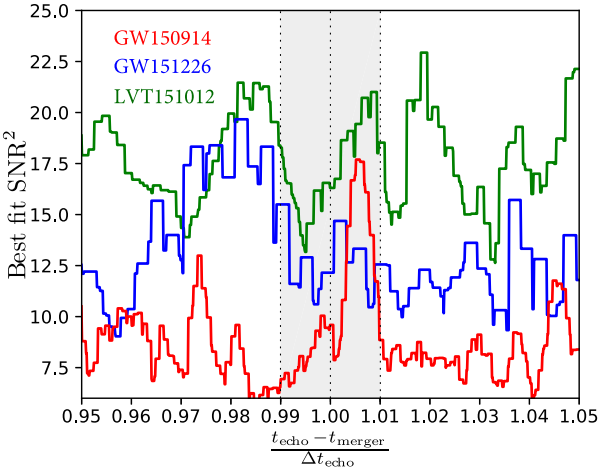

FIG. 6. Best-fit (or maximum) SNR^2 near the expected time of merger echoes (similar to the Fig. 4 in the main text), for all the three events.

Since the inspiral part for different events are not identical, the portion of the energy in inspiral part varies. Hence, we shall obtain different values of ξ_I for different events. Here we have the following best-fit values for different events,

$$\xi_I \simeq \begin{cases} 0.24 & I = GW150914 \\ 0.16 & I = GW151226 \\ 0.17 & I = LVT151012 \end{cases}. \tag{A2}$$

Therefore, the total echoes energy emission is given as follows,

$$E^I_{\text{echoes}}/(M_\odot c^2) = \begin{cases} 0.029 & I = GW150914, \\ 0.047 & I = GW151226, \\ 0.16 & I = LVT151012, \end{cases} \tag{A3}$$

which are the energies emitted in the source frame. One important point to consider is that since GW151226 and LVT151012 contribute to the significance differently, these results for energy emission may not be valid for them. As it is shown in Fig. 6 the background for GW151226 and LVT151012 is relatively higher than GW150914. Therefore, the amplitude used in this calculation is affected by background, systematically biasing the energy to higher values, especially for lower significance events.

APPENDIX B: DETECTION PROPERTIES OF ECHO TEMPLATE

Is the sensitivity of LIGO detectors enough to probe echoes from the ringdown in the data? In Fig. 9, we see that the *best-fit* echo templates are mostly concentrated around the minimum of the amplitude spectral distribution (ASD) of data, where the detector has highest sensitivity. In addition, due to the cumulative nature of echoes, they

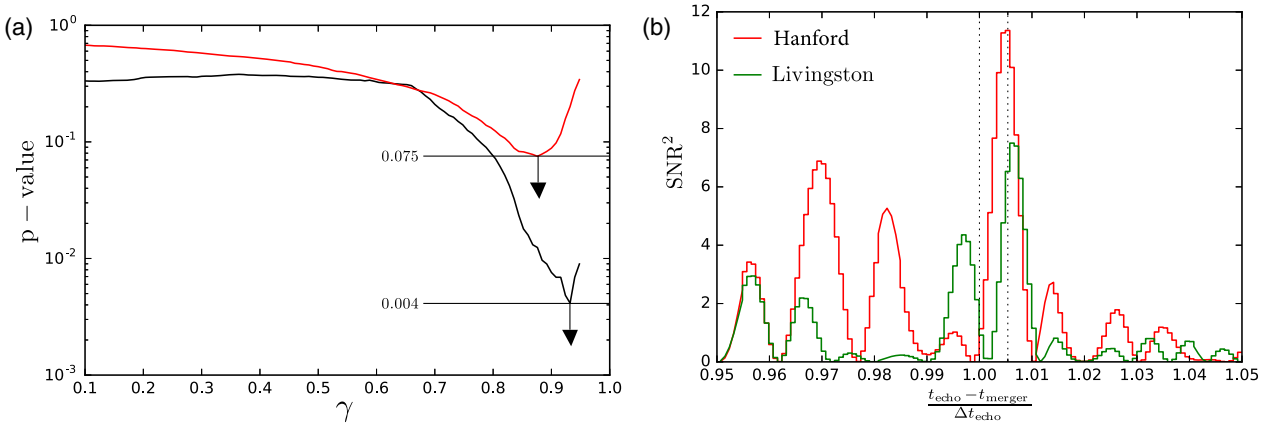

FIG. 7. (a) p-value for our maximum SNR, as a function of γ, computed within a time interval $(-1\%, 1\%) \times \overline{\Delta t}_{\rm echo}$ for GW150914 (top) and combined events (bottom). (b) SNR^2 near the expected time of best-fit merger echoes [Eq. (6)] for GW150914 in Hanford (red) and Livingston (green) detectors. Not only do we see that the two detectors see coincident SNR peaks, but also their ratio $2.42/3.49 = 0.69$ is comparable to the SNR ratio for the main events $13.3/18.6 = 0.72$ seen in the two detectors. Note that, unlike Fig. 4 in the main text, here we have fixed the echo parameters to their best-fit values for combined detectors [36].

produce resonance peaks comparable to the amplitude of the main event even for echoes with overall amplitude 10 times lower than the original event (see Fig. 9).

Given their best-fit values in Table I, one may wonder whether the prior ranges on the parameters γ and t_0 are too narrow. However, changing the prior for these parameters based on their best-fit values leads to *a posteriori* statistics and can bias the p-values. Furthermore, allowing the noise statistics to drive the priors for the model parameters, we may end up with an unphysical range that adversely affect echo searches in the future data releases.

For the damping factor $\gamma = 0.9$ at the boundary of our prior range, one may worry that it might pose a problem for our analysis (an issue that we discussed at length in the main text). This indeed would be the case if the goal was to measure these parameters. However, that has not been our goal, as the parameters only quantify a toy model for the echoes. The goal was rather to find whether the best-fit toy model, within the parameter range, is consistent with random noise. As we discussed in the main text, we find that has a probability of $<1\%$.

Figure 7(a) shows p-value as a function of the echo damping factor γ. We can see that for less damped echoes, p-value drops significantly, which provides substantial evidence for the existence of echoes of gravitational wave in the LIGO data. Here, p-value as a function of γ yields a p-value of 0.004 or a maximum significance of 2.9σ at $\gamma = 0.93$ for combined events. However, we should caution that, as we discuss in the main text, given the duration of data used in the analysis, the interpretation becomes less reliable for $\gamma > 0.9$. Furthermore, Fig. 7(a) shows that the p-value has a jump at $\gamma > 0.9$, signifying a jump in the best fit. This can be understood by noticing that as γ increases, the length of the template increases, eventually diverging in the limit $\gamma \to 1$. This is a singular limit, as arbitrarily high SNRs can be found by simply fitting the noise. Therefore, our method of maximizing SNR fails in this limit, and $\gamma \sim 1$ should be avoided in the prior.

As we argue in the main text, rather than pushing the parameters of a toy model to their extremes, in our opinion, it will be much more fruitful to find more physical echo templates, an effort that is already underway by several research groups.

Finally, one may wonder why combining events with poor individual evidence for echoes can strengthen the significance of the claimed echoes for GW150914. Here, the rationale is that in our model, each echo is quantified by 5 parameters that are poorly constrained. Given the marginal nature of the signal, there will be large degeneracies amongst these parameters from individual events. However, since echoes for different events have different frequency coverage (Fig. 9), the data constrains different combinations of the parameters. Therefore, they can be combined to (at least partially) break these degeneracies and reduce the error on amplitudes leading to a (marginal) detection.

APPENDIX C: THE HOLIDAY EDITION

While we report tentative evidence for the presence of echoes from Planck-scale modifications of general relativity, our statistical methodology was challenged by Ashton, *et al.* [37]. In this section, we summarize our response [36] which addresses these criticisms.

(1) Ashton *et al.* point out that we find a slightly higher $\mathrm{SNR}_{\rm best}$ for echoes in LVT151012, compared to GW150914, even though the SNR for the main event is lower by a factor of 2.4. Is this surprising? In fact, this is expected as constraints on final mass and spin of LVT151012 are significantly worse than

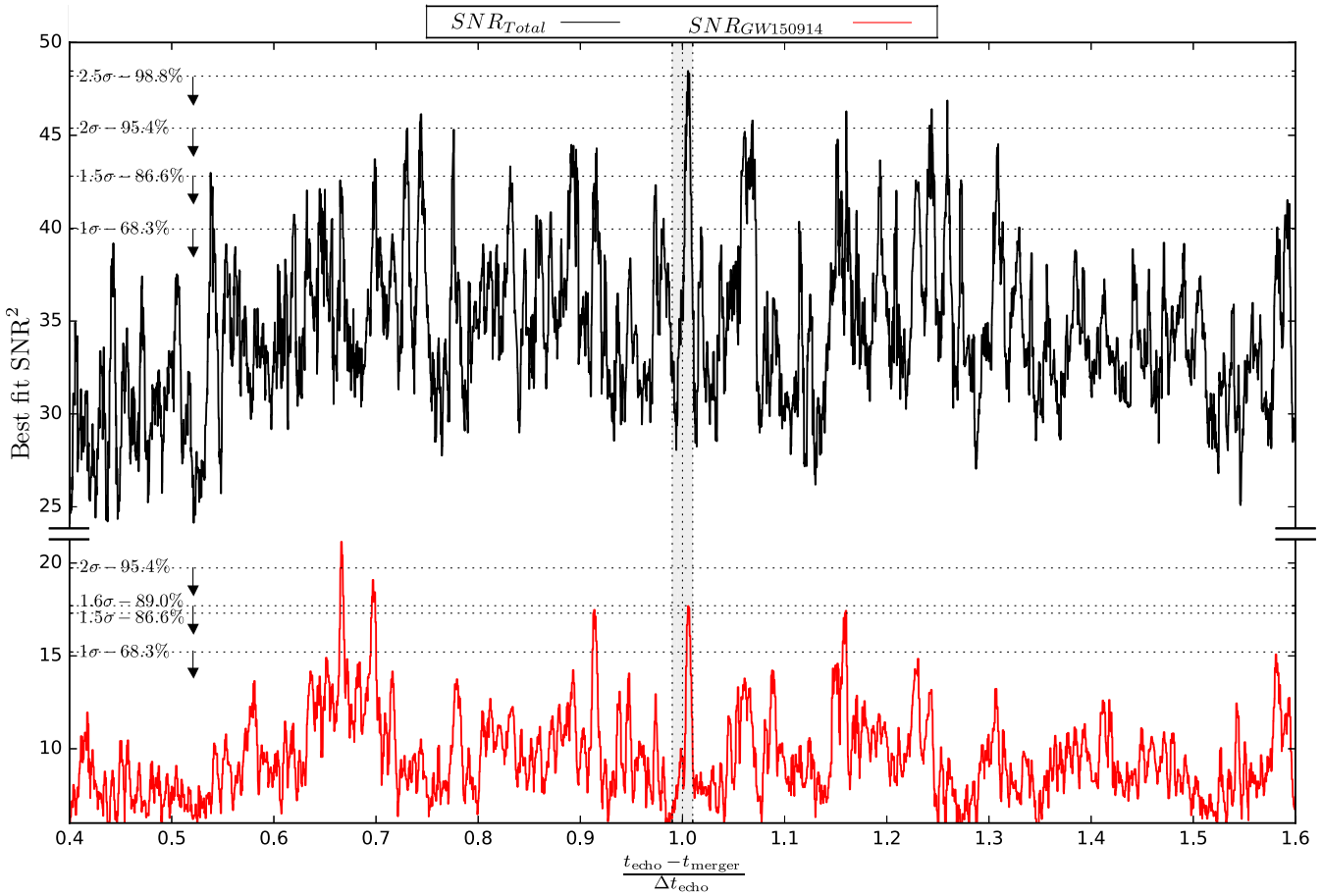

FIG. 8. Same as Fig. 4 in the main text, but over an extended range of $x = \frac{t_{echo} - t_{merger}}{\Delta t_{echo}}$. The SNR peaks at the predicted value of $x = 1$ have false detection probability of 0.11 (0.011) and significance of 1.6σ (2.5σ), for GW150914 (combined events) (See also [36]).

GW150914. As a result, the relative error on Δt_{echo} is 5 times higher for LVT151012, compared to GW150914. This leads to larger values of SNR_{best} across the board, as we are searching a larger region of parameter space. This, however, does not necessarily lead to increased significance, as the same would be true for all values of $x = \frac{t_{echo} - t_{merger}}{\Delta t_{echo}}$.

If there was no real echo signal in LVT151012 and GW151226, adding them to GW150914 would only dilute the significance of the peak near $x = 1$. The fact that the opposite happens suggests that, in spite of larger variations in SNR due to higher uncertainty in Δt_{echo}, there is still significant enhancement in SNR near $x = 1$.

We should also caution about comparing the significance of the echoes with that of the merger events, as they have very different frequency structures (see Fig. 9) leading to different SNR ratios, especially given the nontrivial frequency dependence of the LIGO detector noise.

Finally, we should warn about over-interpreting our quoted significances. Even though we gain comparable evidence for echoes by including LVT151012 and GW151226, i.e. $1.6^2 + 1.6^2 \simeq 2.5^2$, it doesn't mean that they have the same significance: A 1.6σ peak could be a 1-σ fluctuation of a 0.6-σ or a 2.6-σ underlying signal.

For completeness, the individual amplitudes of the best joint fit are listed in Table II in the main text. We note that, even though best-fit SNRs are comparable for the three events, the errors on the amplitude: $\Delta A = A_{best}/SNR_{best}$ is much smaller for GW150914, given that A_{best} is the smallest. Therefore, as expected, GW150914 which is the most significant of the 3 LIGO events, would also dominate the combined constraint on the echo amplitude.

(2) Perhaps the most serious objection of Ashton, *et al.* concerns our estimation of significance, or false-detection probability (p-value). As we outlined in the introduction, it is already clear from Fig. (8) that the p-value for our SNR peak near $x = 1$ should be $\lesssim 0.1$ and $\lesssim 0.01$, for GW150914 and combined events, respectively.

The main criticism of Ashton, *et al.* stems from us quantifying our p-values (in the original arXiv submission of the main text) by considering how often

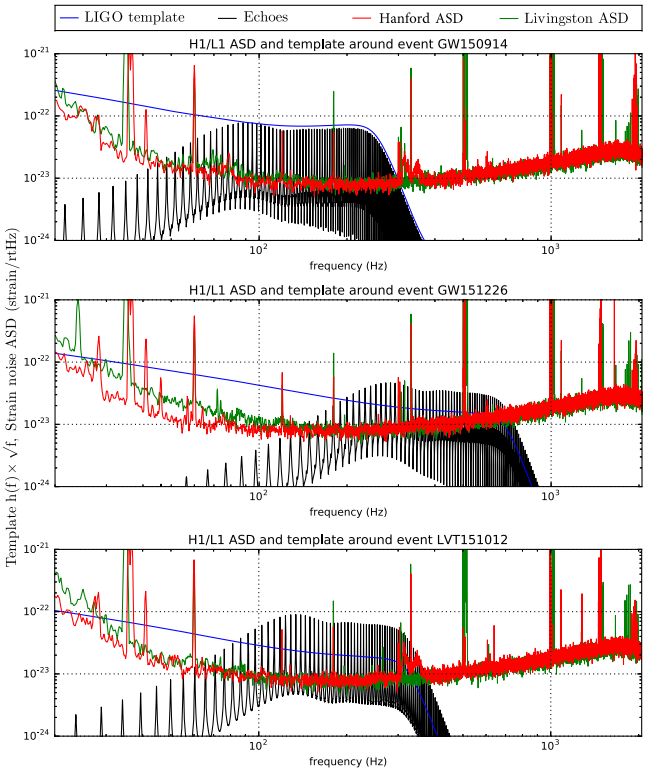

FIG. 9. Best-fit templates for LIGO main events and echoes (using the joint best fit described in the main text), in Fourier space (similar to Fig. 3 in the main text). The amplitude spectral distribution (ASD) for each detector is shown for comparison.

random intervals of size $\Delta x = 0.0054$ have an SNR bigger than the peaks we observe at $x = 1.0054$, while we should have actually allowed for different choices of Δx. This would depend on the prior for Δx: the larger the prior, the higher would be the p-value.

However, we already have a decent idea about this prior from Eq. (1) in the main text which suggests $\Delta x = \mathcal{O}(0.01)$, not far from what we used. We can get a more concrete handle on this prior by assuming that the echo template acquires a random phase (with respect to the main event) due to nonlinear propagation effects. Figure (10) shows the resulting prior on Δx, which we find by replacing the data in our SNR computation (for GW150914) by the echo template with a random phase, and finding the position of the peak. This results in a near top-hat prior with $-0.01 < \Delta x < 0.01$ (an interval of 0.02 rather than 0.0054), which slightly increases the p-value to 0.011 (or significance of 2.5σ).

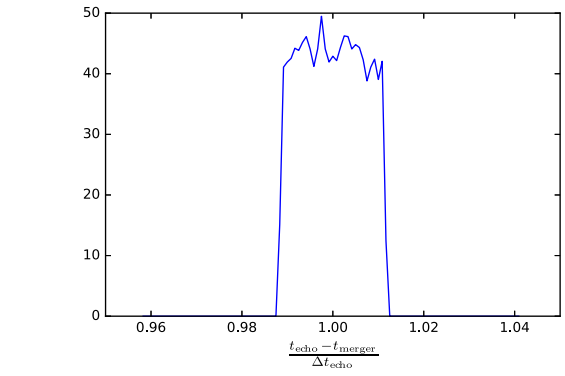

FIG. 10. Resulting prior distribution on $x = \frac{t_{\rm echo}-t_{\rm merger}}{\Delta t_{\rm echo}}$, assuming a random phase for the echo template.

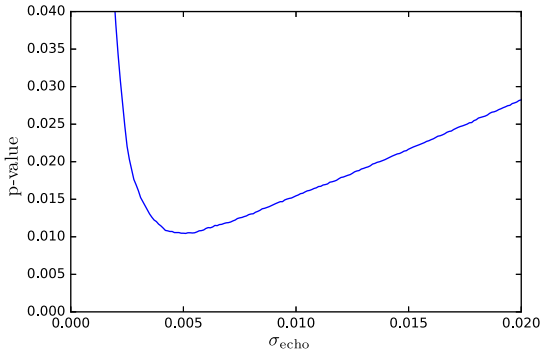

FIG. 11. An alternative false detection probability (p-value) as a function of uncertainty in $t_{\rm echo}$ defined in Eq. (C1).

Yet another way to quantify the significance would be to define a "loudness" function which averages the maximum likelihood for the echoes with a Gaussian prior $x = 1 \pm \sigma_{\rm echo}$, i.e.:

$$L(x,\sigma_{\rm echo}) \equiv \int \exp\left[\frac{\mathrm{SNR}^2_{\rm total}(x')}{2}\right] \times \frac{\exp\left[-\frac{(x-x')^2}{2\sigma^2_{\rm echo}}\right]}{\sqrt{2\pi\sigma^2_{\rm echo}}} dx'. \tag{C1}$$

We again use the LIGO data stream within the range 9–38 $\times$ $\overline{\Delta t}_{{\rm echo},I}$ after the merger event, to quantify how often $L(x,\sigma_{\rm echo})$ exceeds $L(1,\sigma_{\rm echo})$, for a given $\sigma_{\rm echo}$. This plotted in Fig. (11), and provides an alternative p-value (or probability of false detection). This is also minimized at $\sigma_{\rm echo} \simeq 0.5\%$, with p-value of 0.01 (or significance of 2.6σ).

(3) Ashton *et al.* are concerned that the range 9–38 $\times$ $\overline{\Delta t}_{{\rm echo},I}$ after the merger event, which we use to quantify false detection probability, might be contaminated by the echoes and somehow affect our significance estimation. Firstly, this is unlikely, as the evidence for echoes remains marginal and nearly all LIGO data (away from the merger event) is dominated by noise. Secondly, p-value quantifies the probability of null hypothesis, i.e. how often you see the echoes, assuming that there are none. As such, to find p-value one should assume that LIGO data, away from the main event, is pure noise and use that to quantify detection probability, which was what we did. Therefore, we find this criticism ill founded.

Ashton *et al.* further advocate using larger stretches of LIGO data (which is publicly available) to define p-value more precisely. While this is in

principle correct, LIGO noise is known to significantly vary and be very non-Gaussian over long time scales (see Figs. 14–15 in [35]), which makes the interpretation of p-value ambiguous. The 9–$38 \times \overline{\Delta t}_{\mathrm{echo},I}$ interval used is quite adequate to quantify the p-value for our signal, as otherwise we would see a sharp cutoff in our SNR cumulative distribution (Fig. 5 in the main text). We tested this by looking at other stretches of data within a minute of the main events. As can be seen in Fig. 12, we observe that the spread in the p-values in the tail of the distribution is much higher than expected from Poisson statistics of the SNR peaks. Interestingly, the smallest p-value is obtained within the range closest to the main event. This might be expected as the marginal LVT151012 detection is preferentially close to a minimum of the LIGO (combined) detector noise. We believe this justifies using a smaller stretch of data close to the main events, to obtain a faithful reflection of the p-value using the "instantaneous" LIGO noise properties.

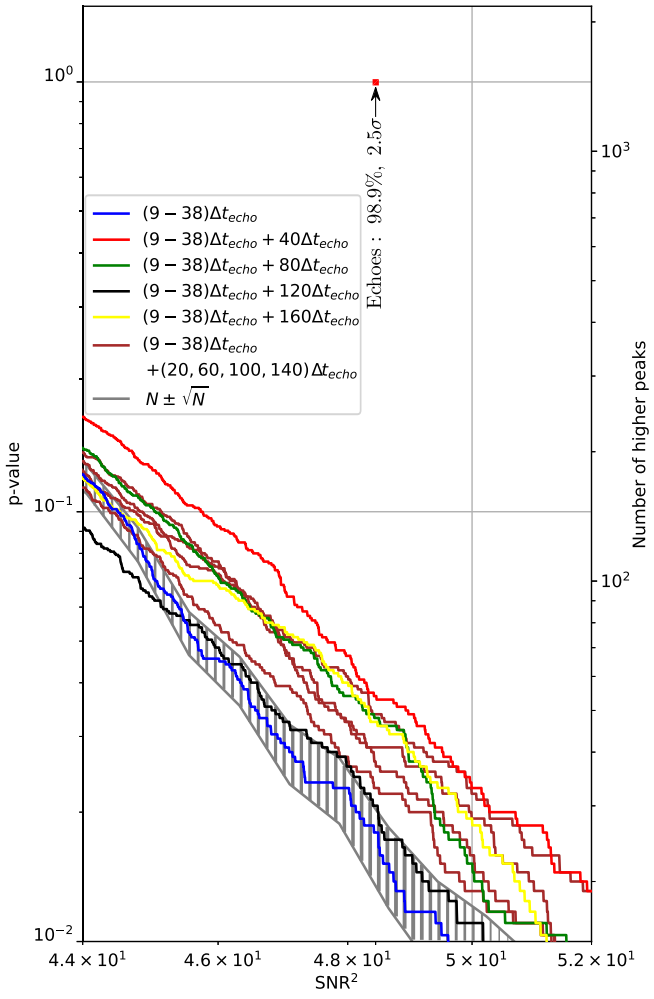

FIG. 12. Distribution of p-values for combined events for different stretches of data within 1 minute of the main events. Surprisingly, the blue line which is closest to the main event, and is used to define p-value in the main text (Fig. 5), happens to give the smallest p-value. The shaded region depicts the Poisson error range for blue histogram, showing that the variation in p-values is clearly much larger. We interpret this as non-Gaussianity and/or nonstationarity in the LIGO noise properties. Here the y-axis on the left (right) shows p-value (number of higher peaks) within the mentioned range of data. The total number of "peaks" considered in each histogram is $(38-9)/0.02 = 1450$.

APPENDIX D: LIGO PREDICTIONS FOR ECHO TIME DELAYS

We approximate the uncertainty in the final redshifted mass M and angular momentum a of the LIGO black hole merger events by a Gaussian probability distribution:

$$P(\Delta a, \Delta M) = \frac{\sqrt{\det M}}{2\pi} \exp\left(-\frac{1}{2}\alpha \Delta a^2 - \frac{1}{2}\beta \Delta M^2 - \gamma \Delta a \Delta M\right) \quad \text{(D1)}$$

where we assume $\langle M \rangle$ and $\langle a \rangle$ are the best-fit reported values, while their inverse covariance matrix is given by,

$$M_{ij} = \begin{bmatrix} \alpha & \gamma \\ \gamma & \beta \end{bmatrix}. \quad \text{(D2)}$$

Here

$$\Delta a = a - \langle a \rangle, \qquad \Delta M = M - \langle M \rangle. \quad \text{(D3)}$$

We can then obtain the probability distribution of Δt_{echo},

$$P(\Delta t) = \int \delta_D(\Delta t(a, M) - \Delta t_{\mathrm{echo}}) P(\Delta a, \Delta M) da dM = \int \frac{dM}{|\frac{\partial \Delta t}{\partial a}|} P(\Delta a, \Delta M) \quad \text{(D4)}$$

This leads to

$$P(\Delta t) \simeq \frac{\sqrt{\alpha\beta - \gamma^2}}{\sqrt{2\pi(\alpha\mu^2 + \beta + 2\gamma\mu)}|\partial \Delta t/\partial a|_{\bar{a},\bar{M}}} \times \exp\left(-\frac{1}{2}\frac{\alpha\beta - \gamma^2}{\alpha\mu^2 + \beta + 2\gamma\mu}\frac{(\Delta t - \Delta \bar{t})^2}{(|\partial \Delta t/\partial a|_{\bar{a},\bar{M}})^2}\right) \quad \text{(D5)}$$

where $\mu = -\frac{\partial \Delta t/\partial M|_{\bar{a},\bar{M}}}{\partial \Delta t/\partial a|_{\bar{a},\bar{M}}}$.

Using contour of 50% credible regions reported in [12], we can obtain the angles of the eigenvectors of the covariance matrix. This gives a relation between α, β, and γ:

$$\tan(2\theta_I) = \frac{2\gamma_I}{\alpha_I - \beta_I} = \begin{cases} 0.013848 & I = GW150914 \\ 0.0072280 & I = GW151226 \\ -0.0038272 & I = LVT151012 \end{cases} \quad \text{(D6)}$$

For the mean of the distribution, using the detector frame (or redshifted) masses we obtain (see [12] Table IV),

$$(\bar{M}_I/M_\odot, \bar{a}_I, \Delta t_{\rm pred}(\bar{a}_I, \bar{M}_I)) = \begin{cases} 67.8 & 0.68 & 0.29559s & I = GW150914 \\ 22.6 & 0.74 & 0.10246s & I = GW151226 \\ 42 & 0.66 & 0.17962s & I = LVT151012 \end{cases}, \quad \text{(D7)}$$

while $\langle \Delta a^2 \rangle$, and $\langle \Delta M^2 \rangle$ are 68% credible region (see [12] Table IV),

$$(\langle \Delta M_I^2 \rangle / M_\odot^2, \langle \Delta a_I^2 \rangle) = \left(\frac{\alpha_I}{\alpha_I \beta_I - \gamma_I^2}, \frac{\beta_I}{\alpha_I \beta_I - \gamma_I^2} \right) = \begin{cases} 4.6866 & 0.0012058 & I = GW150914 \\ 6.7632 & 0.0016633 & I = GW151226 \\ 36.091 & 0.0034056 & I = LVT151012 \end{cases} \quad \text{(D8)}$$

These can be combined with θ_I's (D6) to give

$$(\alpha_I, \beta_I/M_\odot^2, \gamma_I/M_\odot) = \begin{cases} 1019.1 & 0.26221 & 7.0546 & I = GW150914 \\ 634.92 & 0.15615 & 2.2940 & I = GW151226 \\ 305.49 & 0.028826 & -0.58452 & I = LVT151012 \end{cases} \quad \text{(D9)}$$

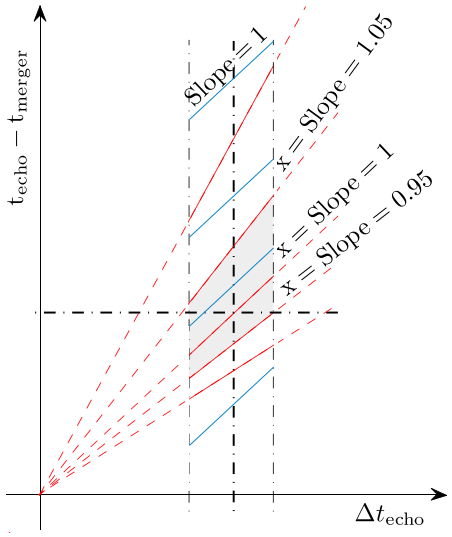

FIG. 13. SNR peaks and their significance: We maximize SNR along the red lines for fixed values of $x = \frac{t_{\rm echo} - t_{\rm merger}}{\Delta t_{\rm echo}}$, which is used to search for best-fit echo parameters, within 1σ region for $\Delta t_{\rm echo}$ (vertical lines). In particular, the grey trapezoid ($x = 1 \pm 0.05$) corresponds to Fig. 4 in the main text. After we find an SNR peak at some value of $x \simeq 1$, we quantify its significance by how often a comparable $\rm SNR_{max}$ can be found along the blue lines with unit slope for $x \gg 1$.

With these values we can obtain the Gaussian posterior for $\Delta t_{\rm echo}$'s,

$$P(\Delta t_I) = \begin{cases} \frac{77.187 s^{-1}}{\sqrt{\pi}} \exp(-5957.8 s^{-2} (\Delta t - 0.2925s)^2) & I = GW150914 \\ \frac{61.372 s^{-1}}{\sqrt{\pi}} \exp(-3766.5 s^{-2} (\Delta t - 0.1013s)^2) & I = GW151226 \\ \frac{25.357 s^{-1}}{\sqrt{\pi}} \exp(-643.00 s^{-2} (\Delta t - 0.1778s)^2) & I = LVT151012 \end{cases} \quad \text{(D10)}$$

APPENDIX E: DISCUSSION AND CONSISTENCY OF STATISTICAL METHODOLOGY

In this section we provide additional technical details that can be used to reproduce our results and test their robustness and consistency.

Figure (13) demonstrates how we find our SNR peaks and their significance. In searching for echoes, we consider $x = (t_{\rm echo} - t_{\rm merger})/\Delta t_{\rm echo} = {\rm constant}$ (red lines), and maximize SNRs varying all the free parameters within their priors (Tables I–II in main text). Figure 4 in the main text shows maximum SNR in the range $x = (0.95, 1.05)$ (grey trapezoid in Fig. 13). Here x depends on two variables: $t_{\rm echo}$ and $\Delta t_{\rm echo}$. As we discuss in the text, we seek the maximum SNR within $x = 1 \pm 0.01$ (see below), and quantify its significance by how often a higher $\rm SNR_{max}$ is achieved if the interval is shifted (e.g. the blue lines in Fig. 13) far from $x = 1$.

We used 4 times higher (4×4096 Hz $= 16384$ Hz) grid than the resolution of data for $\Delta t_{\rm echo}$ which is the most sensitive parameter in our search. For t_0 and γ which are the less sensitive ones we used 76 and 100 points respectively.

Figure 10 shows how the best-fit SNR peak for the echoes of GW150914 moves if the echo template is multiplied by a random phase, expected from nonlinear effects during the merger. We use this to fix the prior range for $t_{\rm echo}$, roughly corresponding to:

$$x = \frac{t_{\rm echo} - t_{\rm merger}}{\Delta t_{\rm echo}} \simeq 1 \pm 0.01. \quad \text{(E1)}$$

Figure 8 is the most clear demonstration of the significance of our results. This is the same as Fig. 4 in the main text, but over a larger range. For both GW150914 (the most significant reported LIGO event) and combined data from

all three events, there exists a peak at distance 0.54% from x = 1, which is inside the vertical grey bar [Eq. (E1) and Fig. 10]. For GW150914, the false detection probability is 11% or the significance is 1.6σ, meaning that comparable SNR peaks (from random noise) can be found within $\Delta x \simeq 0.02/0.1 = 0.2$, as can be seen with other peaks at $x \simeq 0.91$ and 1.16. For the combined events, the false detection probability is 1% or the significance is 2.5σ, i.e. comparable peaks can only be found within $\Delta x \simeq 0.02/0.011 = 1.8$. This is also consistent with Fig. 8 as no higher peak can be seen within $\Delta x = 1.2$ in the plot.

Yet another consistency test is presented in Fig. 7(b), where we show independent SNRs for Hanford and Livingston detectors for the best-fit parameters of GW150914 echoes. We can clearly see that not only the SNR peaks for echoes in different detectors coincide (after accounting for the event time delays), but also their ratio is consistent with SNR ratios for the main merger event reported by LIGO Collaboration.

[1] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, J. High Energy Phys. 02 (2013) 062.
[2] O. Lunin and S. D. Mathur, Nucl. Phys. **B623**, 342 (2002).
[3] O. Lunin and S. D. Mathur, Phys. Rev. Lett. **88**, 211303 (2002).
[4] J. Maldacena and L. Susskind, Fortschr. Phys. **61**, 781 (2013).
[5] J. Abedi and H. Arfaei, Classical Quantum Gravity **31**, 195005 (2014).
[6] J. Abedi and H. Arfaei, J. High Energy Phys. 03 (2016) 135.
[7] S. L. Braunstein, S. Pirandola, and K. Zyczkowski, Phys. Rev. Lett. **110**, 101301 (2013).
[8] P. Hayden and J. Preskill, J. High Energy Phys. 09 (2007) 120.
[9] Y. Sekino and L. Susskind, J. High Energy Phys. 10 (2008) 065.
[10] C. Prescod-Weinstein, N. Afshordi, and M. L. Balogh, Phys. Rev. D **80**, 043513 (2009).
[11] N. Afshordi and Y. K. Yazdi, Classical Quantum Gravity **33**, 235017 (2016).
[12] B. P. Abbott *et al.* (Virgo, LIGO Scientific Collaborations), Phys. Rev. X **6**, 041015 (2016).
[13] B. P. Abbott *et al.* (Virgo, LIGO Scientific Collaborations), Phys. Rev. Lett. **116**, 131103 (2016).
[14] S. B. Giddings, Classical Quantum Gravity **33**, 235010 (2016).
[15] M. Saravani, N. Afshordi, and R. B. Mann, Int. J. Mod. Phys. D **23**, 1443007 (2014).
[16] S. D. Mathur and D. Turton, J. High Energy Phys. 01 (2014) 034.
[17] A. E. Broderick, A. Loeb, and R. Narayan, Astrophys. J. **701**, 1357 (2009).
[18] A. E. Broderick, R. Narayan, J. Kormendy, E. S. Perlman, M. J. Rieke, and S. S. Doeleman (Perimeter Institute for Theoretical Physics), Astrophys. J. **805**, 179 (2015).
[19] U.-L. Pen and A. E. Broderick, Mon. Not. R. Astron. Soc. **445**, 3370 (2014).
[20] V. Cardoso, E. Franzin, and P. Pani, Phys. Rev. Lett. **116**, 171101 (2016); **117**, 089902(E) (2016).
[21] V. Cardoso, S. Hopper, C. F. B. Macedo, C. Palenzuela, and P. Pani, Phys. Rev. D **94**, 084031 (2016).
[22] R. H. Price and G. Khanna, arXiv:1702.04833.
[23] K. D. Kokkotas and B. G. Schmidt, Living Rev. Relativity **2**, 2 (1999).
[24] H. Yang, A. Zimmerman, A. Zenginolu, F. Zhang, E. Berti, and Y. Chen, Phys. Rev. D **88**, 044047 (2013); **88**, 044047 (2013).
[25] G. 't Hooft, Nucl. Phys. **B256**, 727 (1985).
[26] I. B. Khriplovich, *General Relativity*, 1st ed. (Springer-Verlag, New York, 2005).
[27] J. W. York, Phys. Rev. D **28**, 2929 (1983).
[28] D. Harlow, Rev. Mod. Phys. **88**, 015002 (2016); **88**, 015002 (2016).
[29] D. Harlow and P. Hayden, J. High Energy Phys. 06 (2013) 085.
[30] Binary Black Hole Signals in LIGO Open Data, https://losc.ligo.org/s/events/GW150914/LOSC_Event_tutorial_GW150914.html.
[31] Binary Black Hole Signals in LIGO Open Data, https://losc.ligo.org/s/events/GW151226/LOSC_Event_tutorial_GW151226.html.
[32] Binary Black Hole Signals in LIGO Open Data, https://losc.ligo.org/s/events/LVT151012/LOSC_Event_tutorial_LVT151012.html.
[33] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens, J. Phys. Conf. Ser. **610**, 012021 (2015).
[34] E. Poisson and M. Sasaki, Phys. Rev. D **51**, 5753 (1995).
[35] B. P. Abbott *et al.* (Virgo, LIGO Scientific Collaborations), Phys. Rev. Lett. **116**, 221101 (2016).
[36] J. Abedi, H. Dykaar, and N. Afshordi, arXiv:1701.03485.
[37] G. Ashton, O. Birnholtz, M. Cabero, C. Capano, T. Dent, B. Krishnan, G. D. Meadors, A. B. Nielsen, A. Nitz, and J. Westerweck, arXiv:1612.05625.
[38] D. V. Martynov *et al.* (LIGO Scientific Collaborations), Phys. Rev. D **93**, 112004 (2016).
[39] J. L. Friedman, Commun. Math. Phys. **63**, 243 (1978).
[40] V. Cardoso, P. Pani, M. Cadoni, and M. Cavaglia, Phys. Rev. D **77**, 124044 (2008).
[41] Z. Mark, A. Zimmerman, S. M. Du, and Y. Chen, arXiv:1706.06155.
[42] H. Nakano, N. Sago, H. Tagoshi, and T. Tanaka, Prog. Theor. Exp. Phys. **2017**, 071E01 (2017).
[43] A. Maselli, S. H. Vlkel, and K. D. Kokkotas, arXiv:1708.02217.