Uncorrected Proof

The first direct detection of gravitational waves opens a vast new frontier in astronomy

F. J. Raab[1,]* and D. H. Reitze[2]

[1]LIGO Hanford Observatory, P.O. Box 159, Richland, WA 99352, USA
[2]LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, USA

Keywords: Astronomy, direct detection, gravitational waves, space-time.

Introduction

THE first direct detection of gravitational waves (GWs), announced on 11 February 2016, has opened a vast new frontier in astronomy. Albert Einstein predicted the existence of these waves about a century ago as a consequence of his general theory of relativity. Radio astronomy observations of the binary pulsar system PSR 1913 + 16 over a 20 year period beginning in 1975 provided strong observational evidence that gravitational waves carried energy away from the orbits of neutron stars at precisely the level predicted by general relativity (GR). This relentless conversion of orbital energy into gravitational wave energy causes binary orbits to decay until the objects eventually collide and merge. The frontier of precision measurement science, using laser interferometers, was pushed for more than four decades to achieve this first direct detection, marking a milestone in experimental physics and engineering. Even more significantly, this milestone also opens a new window onto our universe and a completely new kind of astronomy to explore.

Gravitational waves are dynamic distortions in space–time produced by accelerating masses. Two concepts from special relativity are helpful to understand these waves and how they are detected. Although space and time are relative concepts, the speed of light is not and thus measurements of light define space and time. Also, the speed of light is the maximum speed for information transfer. If one conducts an illustrative (and unphysical) thought experiment in which the Sun disappears instantaneously, the Earth would continue in its orbit for more than 8 min before earthlings noticed the disappearance of the Sun, simultaneously accompanied by the Earth leaving its traditional orbit. One by one the planets would leave their orbits as this front of changing gravity, called a gravitational wave, passed by.

Einstein's GR describes the physics of gravitational waves. In GR, the presence of matter or energy induces curvature in space. The ratio of curvature to density is proportional to (G/c^4), where G is the strength of gravity and c is the speed of light. In SI units G/c^4 is of order 10^{-44}, which indicates that space is very stiff, requiring large concentrations of matter or energy to warp it by a measureable amount.

A closed system of masses cannot have a fluctuating mass dipole moment due to conservation of momentum, but can have a fluctuating quadrupole moment which describes the stretching and squeezing of the mass distribution. Orbiting binary stars have a time-varying quadrupole moment, which produces a gravitational wave of amplitude h. A pair of 1.4 $M_\odot$ neutron stars, orbiting with frequency f_{orb} = 400 Hz, separated by R = 20 km, at a distance r of 15 Mpc in the Virgo cluster of galaxies, would produce a gravitational wave strain on the Earth of approximately:

$$h \approx \frac{4\pi^2 GMR^2 f_{\mathrm{orb}}^2}{c^4\ r} \rightarrow h \sim 10^{-21}. \tag{1}$$

Physically, a passing gravitational wave produces a quadrupolar strain, $h = 2\Delta L/L$ in the plane of space transverse to its propagation direction, where ΔL is the stretching or shrinking of the length L along the perpendicular direction. This can occur in two polarizations illustrated in Figure 1.

As conceptualized by Gershenstein and Pustovoit[1] and investigated in detail by Weiss[2], Michelson interferometers are well-suited for detecting gravitational-wave strains. A Michelson interferometer measures the time of

*For correspondence. (e-mail: fred.raab@ligo.org)

Uncorrected Proof

flight in the form of a phase shift between light beams travelling along perpendicular distances marked by suspended mirrors (referred to as test masses) in the centre (beam splitter) and along the circumference (end mirrors) of a large circular cross-section of space as shown in Figure 2 (centre). The isolated test masses can be thought of as survey stakes planted in inertial space. On the plus (left) and minus (right) cycles of the GW distortion, the light in each path travels at the same speed but takes differing amounts of time to return to the central beam splitter. The resultant phase shift between the light beams is converted into an electrical signal by the "detector" in the figure. Controllably displacing the mirrors by a known amount allows that electrical signal to be calibrated in terms of strain.

The ability to measure gravitational wave signals unambiguously requires understanding and mitigating sensing and other background noises which can mask a true signal. Sensing noise refers to effects that limit the ability to resolve very small mirror displacements. Shot noise, due to the quantum nature of light, fluctuations in the residual gas molecules in the light paths, and light scattering produce sensing noise. Background noises refers to real displacements of the mirrors. Fluctuations in local gravity near the mirrors, radiation pressure fluctuations on the mirrors (also due to the quantum nature of light), as well as motions of the atoms at temperature T (referred to as thermal noise) in the suspended mirrors are a few examples of background. Typically a laser interferometer is limited at low frequencies by vibrations transmitted from the local environment and fluctuations in local gravity. Quantum fluctuations in the electromagnetic vacuum produce shot noise that limits sensitivity at high frequencies, as well as quantum radiation pressure fluctuations that combine with thermal noise to limit sensitivity at intermediate frequencies[3].

Ultimately backgrounds drive terrestrial GW detectors to kilometre length scales. At this scale, strains in the range 10^{-21} to 10^{-22}, expected at the Earth from the strongest astrophysical sources, will produce displacements of only billionths of the size of an atom. This is tiny compared to the typical motion of an atom in a solid at room temperature; however, the laser beam averages over all of the atoms at the mirror surface, allowing interferometers to achieve high measurement precision.

The initial proposal for construction of LIGO (the Laser Interferometer Gravitational-Wave Observatory)[4] envisioned a pair of 4 km detector facilities in the US with continental-scale separation that would be operated as components of an evolving international network of detector facilities. Many mechanisms were expected to produce random transients in a single detector that would be larger and more frequent than the GW signals being sought. Detection of identical signals in at least two independent detectors would be a minimal requirement to positively identify a true GW signal among these transients. As a further reason for requiring multiple detectors, although individual GW detectors have poor directional resolution, differences in the time of arrival of GW signals among widely spaced detectors provide information on source direction. A network of 3–4 distant detectors on a curved Earth can provide good resolution of directional and polarization information. As more distant detectors are added, this resolution and the network robustness continue to improve.

LIGO was proposed to be built in two phases. The initial LIGO phase envisioned construction of two separated observatory facilities, which could support successive generations of installed detectors as technology advanced, with a first (pathfinder) detector installed using the best technology available at the end of the 20th century. The two facilities would be optimally aligned for detection confidence, at the expense of polarization resolution. GW detection in this first phase was considered possible, but not probable.

Lessons learned from operating this detector, combined with a technology development programme drawing lessons from observatory operations, would inform the design of an Advanced LIGO detector that would make GW detections routine.

Figure 3 illustrates the sensitivities of these detector concepts, along with the best estimates of the GW signals that could be produced by different astrophysical events at a number of distances. Of course, little was known about the existence, distance and frequency of these events when the LIGO Construction Proposal was submitted[5].

The LIGO Laboratory, jointly proposed and managed by the California Institute of Technology (Caltech) and the Massachusetts Institute of Technology (MIT), USA, ensures uniform operation of the two observatories in coincidence with the support facilities at Caltech and MIT, uniform configurations of the detectors and uniform design, construction and installation of detector upgrades. Realizing that reaping the best science from LIGO operations would require much greater expertise than existing in the LIGO Laboratory, the LIGO Scientific Collaboration[6] was formed as an open collaboration in 1997.

The first phase of LIGO construction began in 1992 and the first detectors operated in six observing runs between 2000 and 2010. Although the design sensitivity targets were met by the initial detector[7], no GW detections were made. However, more than 100 journal articles reported upper limits on abundance or lower limits on strengths of a variety of astrophysical events or sources[8]. Data analyses searched for GW emission from: inspirals, mergers and ringdowns of binaries composed of neutron stars and black holes, for which there are explicit waveform models; transient burst sources such as supernovae, for which the waveforms are poorly known; quasi-periodic sources such as from deformed rotating neutron stars, and GW stochastic background of cosmological or

Uncorrected Proof

astrophysical origin. Partnerships were developed with astronomers for electromagnetic[9] and neutrino detector[10] follow-ups of GW triggers, in anticipation of multi-messenger astronomy associated with GW detections.

Research and development for an Advanced LIGO detector continued throughout the initial phase of LIGO construction and operation. The Advanced LIGO detectors were constructed and installed between 2008 and 2015 (ref. 11). Advanced LIGO detector improvements included: higher power lasers to reduce shot noise; improved, active vibration isolation systems and multi-stage suspensions to improve robustness and extend operation to much lower frequency; development of monolithic fused silica mirror/suspension stages to reduce thermal noise; thermal wavefront tuning and incorporation of a signal recycling cavity to allow tuning of the frequency response of the detector. A simplified optical layout of Advanced LIGO is shown in Figure 4. These detectors are highly complex, described by approximately 15,000 drawings, with multiple coupled optical cavities, controlled by approximately 350 high-performance servo systems, many of which are multi-input, multi-output, with blended sensors and actuators, with hundreds of thousands of named data channels.

By August 2015, commissioning of the advanced LIGO detectors had reached sufficient sensitivity to embark on a pre-observing engineering run. In the morning hours of 14 September 2015, a few days before the official announcement of the first observing run, an on-line burst analysis code found a strong GW candidate event in the Livingston detector (L1), with a matching strong trigger from the Hanford detector (H1) 7 ms later. This began an intense campaign of scrutiny by the LIGO and Virgo collaborations that confirmed this as the first direct detection of gravitational waves[12], designated as GW150914. Figure 5 shows the signals seen in the time-series data.

The frequency evolution of these signals is shown in Figure 6 the time–frequency plane. Although the full power of numerical solutions general relativity are needed to accurately extract the full information in the waveforms, inspection of Figures 5 and 6 indicates the likely origin to be the merger of two black holes[13]. The rise in signal frequency as time progresses indicates that the signal is generated by a binary inspiral. The detailed evolution of the frequency and amplitude indicates that these are massive objects (of order tens of solar masses). The highest frequency indicates that the objects are very compact approaching a centre of mass separation of only a couple of hundred kilometres. Since no observed states of matter could have this much mass with this centre of mass separation, we identify the merging objects as black holes. The merger should produce an initially deformed spinning black hole, with a decaying ringdown.

Data were analysed for first the observing run of Advanced LIGO[14] (O1) from 12 September 2015 from which time the detectors were in a stable state at the close of engineering run ER10, until 19 January 2016.

These produced three coincident triggers with high significance and waveforms expected from binary black hole mergers as shown in Figure 8. For GW150914 and GW151226, the probability of astrophysical origin of the signal is better than 99.9999%, and they are confidently claimed as gravitational-wave detections. The event LVT151012 has an 86% probability of astrophysical origin; this did not meet our standard to claim confidently as a gravitational wave detection. Figure 7 shows different times to merger when measured from entering the Advanced LIGO detector frequency band, with binaries containing a smaller mass component radiating less strongly and taking longer to merge.

In addition to proving the existence of binary black hole systems, these waveforms were used to provide significant tests of GR in the most extreme space–times ever observed. All tests agreed with GR predictions to within the precision of the data. To appreciate how remarkable this is, consider that the ratio of the speed of the binary neutron stars used for the best previous tests of Einstein's GR was moving at a thousandth the speed of light. In the merger producing GW150914, the two black holes were moving at roughly half the speed of light when their event horizons met.

Prior to these detections it was unknown whether black-hole binaries would form in Nature. We now know that they do and that such mergers are likely to occur somewhere in the universe hourly.

Even as we try to refine that rate and explore the statistics of black hole properties, experimentalists are working on refinements, upgrades and new generations of detectors that could eventually observe black hole mergers across the universe. A first priority is to improve the sensitivity of the detectors already in operation. GW detectors detect the strain, which decreases as distance from a source, rather than the energy, which decreases as distance squared. Thus improving strain sensitivity by a factor of two allows a given source to be detected at twice the distance. The number of galaxies, and thus rate of detected events, should increase as the cube of the sensitivity increase. Figure 8 compares the sensitivity of the Advanced LIGO detectors in the O1 run[16], compared to the best sensitivity achieved by the first-generation detector in the S6 observing run, as well as the design sensitivity that fully commissioned Advanced LIGO detectors can be expected to achieve by 2019–2020. The figure plots the signal to noise ratio (SNR) that a detector can achieve for an optimally aligned source (directly over head, with 'stretch' and 'shrink' axes along the detector axes) as a function of the redshift, due to expansion of the universe while the wave travelled to Earth. With further commissioning, we expect Advanced LIGO should improve its range by more than a factor of 2.5, improving the event rate by a more than factor of 15.

Uncorrected Proof

Expanding the global network of gravitational wave detectors will allow dramatic improvements in extracting source position and source polarization information. The Virgo detector[17] in Italy is expected to begin observations by 2017. The KAGRA detector[18] in Japan currently under construction is expected to be observing before 2020, followed by LIGO-India[19] in 2024.

One important target of achieving the full design sensitivity of Advanced LIGO and the expanding the global network is to detect and resolve the mergers of binaries containing one or two neutron stars. Black holes binaries are expected to be electromagnetically dark due to the absence of matter. However, if the extreme gravity gradients encountered by a neutron star just before merger shred the star, then accelerating charges would be shining brightly across the electromagnetic spectrum. A network of three or more detectors should provide telescopes with orders of magnitude more precise coordinates than the two LIGO facilities alone could provide[20].

Farther in the future, we anticipate a third generation of gravitational wave detectors that can: observe black hole mergers back to the first generation of stars; measure the neutron-star equation of state; unravel the workings of the supernova engine; perform precision cosmology and evolution studies, and perform high-precision tests of general relativity. The major fundamental barriers to the needed sensitivity improvements are thermal noise from the currently available mirror coatings and quantum noise.

Quantum noise in GW detectors originates from the fluctuating electromagnetic vacuum[3,21], which in gravitational wave interferometers produces both phase noise at higher frequencies (above 100 Hz) and radiation pressure fluctuations at low frequencies (less than 100 Hz). Vacuum squeezing to suppress phase noise has already been implemented to reduce phase noise in GW detectors[22,23]. Frequency-dependent squeezing, capable of reducing both phase noise and radiation-pressure noise, has been demonstrated at frequencies of interest for GW detectors[24].

Coating thermal noise can be mitigated by reducing mechanical losses in the multilayer dielectric mirror coatings, which historically have been optimized for low optical losses[25]. Molecular modelling to better understand the current silica/tantala coatings has been making progress recently[26], leading to possible prescriptions for producing coatings with lower mechanical loss currently being investigated.

Alternatively, thermal noise can be reduced by reducing the temperature of the test masses. Research into cooling test masses continues, which requires new test-mass substrates and perhaps new coatings and laser sources. KAGRA plans to operate with sapphire test masses cooled to 20 K. Research is also being done toward detectors that may operate in current or new facilities at 120 K using silicon test masses, new crystalline mirror coatings and new laser sources in the range of 1.5–2 μm wavelength.

New facilities with longer arm lengths, in the 10–40 km length range, are a possible way to extend the range of GW detectors. A given gravitational-wave strain produces a larger displacement in metres in a longer detector. Scaling of noise is more complicated, but generally favourable to longer detectors, especially if test-mass size and mass are increased[27,28].

In summary, the first direct detections of gravitational waves from BBH inspirals and mergers have opened up a vast new frontier for astronomy and physics. These first signals, captured in the first observing run of second-generation Advanced LIGO detectors, came from more than a billion light years away. Other detectors such as Virgo, KAGRA and LIGO-India will soon come on-line to fill out a global network of GW detectors working with electromagnetic and neutrino detectors for multi-messenger explorations of this frontier. Farther into the future, we expect to push GW observations farther and farther into our past, back to remnants of the first generation of stars.

1. Gertsenschtein, M. E. and Pustovoit, V. I., On the Detection of low frequency gravitational waves. *Sov. Phys. JETP*, 1963, **16**, 605–607.
2. Weiss, R., Electromagnetically coupled broadband gravitational antenna. Quarterly Report No. 105, MIT Radiation Laboratory, 1972; https://dcc.ligo.org/P720002/public.
3. Caves, C. M., Quantum mechanical noise in an interferometer. *Phys. Rev. D*, 1981, **23**, 1693–1708.
4. Abramovici, A. *et al.*, LIGO – the Laser–interferometer–gravitational–wave-observatory. *Science*, 1992, **256**, 325–333.
5. Vogt, R., Drever, R., Thorne, K., Raab, F. and Weiss, R., The construction, operation, and supporting research and development of a Laser Interferometer Gravitational-Wave Observatory; https://dcc.ligo.org/LIGO-M890001/public
6. LIGO Scientific Collaboration; http://www.ligo.org
7. Abbott, B. P. *et al.*, LIGO: the Laser Interferometer Gravitational-Wave Observatory. *Rep. Prog. Phys.*, 2009, **72**, 076901.
8. A list of LIGO Scientific Collaboration articles can be found at https://www.lsc-group.phys.uwm.edu/ppcomm/Papers.html
9. Aasi, J., *et al.*, First searches for optical counterparts to gravitational-wave candidate events. *Astrophys. J. Suppl.*, 2014, **211**, 7.
10. Aartsen, M. G. *et al.*, Multimessenger search for sources of gravitational waves and high-energy neutrinos: initial results for LIGO-Virgo and IceCube. *Phys. Rev. D*, 2014, **90**, 102002.
11. Aasi, J. *et al.*, Advanced LIGO. *Class. Quantum Grav.*, 2015, **32**, 074001.
12. Abbott, B. P., Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett*., 2016, **116**, 061102; https://doi.org/10.1103/PhysRevLett.116.061102.
13. Abbott, B. P. *et al.*, The basic physics of the binary black hole merger GW150914. *Ann. Phys.*, 2016.
14. Abbott, B. P. *et al.*, Binary Black Hole Mergers in the First Advanced LIGO Observing Run. *Phys. Rev. X*, 2016, **6**, 041015; https://doi.org/10.1103/PhysRevX.6.041015.
15. Veitch, J. *et al.*, Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LAL Inference software library. *Phys. Rev. D*, 2015, **91**, 042003.

Uncorrected Proof

16. Abbott, B. P. *et al.*, GW150914: the Advanced LIGO detectors in the era of first discoveries. *Phys. Rev. Lett.*, 2016, **116**, 131103; https://doi.org/10.1103/PhysRevLett.116.131103.
17. Acernese, F. *et al.*, Advanced Virgo: a second-generation interferometricgravitational wave detector. *Class. Quant. Grav*., 2015, **32**, 024001.
18. Aso, Y. *et al.*, ‘Interferometer design of the KAGRA gravitational wave detector. *Phys. Rev. D*, 2013, **88**, 043007.
19. Article by Sendhil and Tarun.
20. Sathyaprakash, B. S., Fairhurst, S., Schutz, B. F., Veitch, J., Klimenko, S., Reitze, D. H. and Whitcomb, S. E., Scientific benefits of moving one of LIGO Hanford detectors to India; https://dcc.ligo.org/LIGO-T1200219/public.
21. Caves, C. M., Quantum-mechanical radiation-pressure fluctuations in an interferometer. *Phys. Rev. Lett*., 1980, **45**, 75–79.
22. Abadie, J. *et al.* (The LIGO Scientific Collaboration), A gravitational wave observatory operating beyond the quantum shot-noise limit. *Nat. Phys.*, 2011, **7**, 962–965.
23. Aasi, J. *et al*. (The LIGO Scientific Collaboration), Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light. *Nature Photonics*, 2013, **7**, 613–619.
24. Oelker, E., Isogai, T., Miller, J., Tse, M., Barsotti, L., Mavalvala, N. and Evans, M., Audio-band frequency-dependent squeezing for gravitational-wave detectors. *Phys. Rev. Lett*., 2016, **116**, 041102.
25. Harry, G. M., *et al*., Thermal noise in interferometric gravitational wave detectors due to dielectric optical coatings. *Class. Quantum Grav*., 2002, **19**, 897–917.
26. Hamden, R., Trinastic, J. P. and Cheng, H., Molecular dynamics study of the mechanical loss in amorphous pure and doped silica. *J. Chem. Phys*., 2014, **141**, 054501.
27. Dwyer, S., Sigg, D., Ballmer, S. W., Barsotti, L., Mavalvala, N. and Evans, M., Gravitational wave detector with cosmological reach. *Phys. Rev. D*, 2015, **91**, 082001.
28. Einstein Telescope design study; http://www.et-gw.eu/etdsdocument.

ACKNOWLEDGEMENTS. We gratefully acknowledge the National Science Foundation award PHY-0757058 in supporting the LIGO Laboratory and the writing of this paper.

doi:

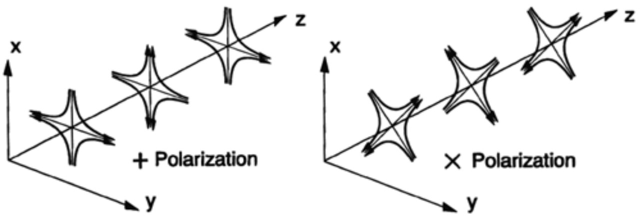

Figure 1. Stretching and squeezing of space transverse to the direction of a gravitational wave. The wave is traveling in the +*z* direction, and the stretching and squeezing occur in the *x* and *y* dimensions.

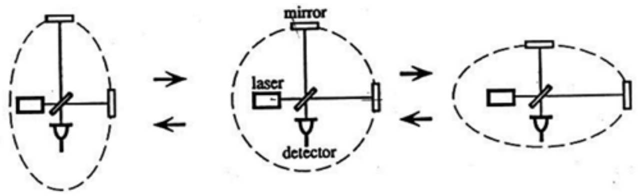

Figure 2. Sketch of oscillating stretching and shrinking of space; the effect of the wave on a free-mass laser interferometer is superposed. Here, the wave is travelling perpendicular to the plane of the interferometer.

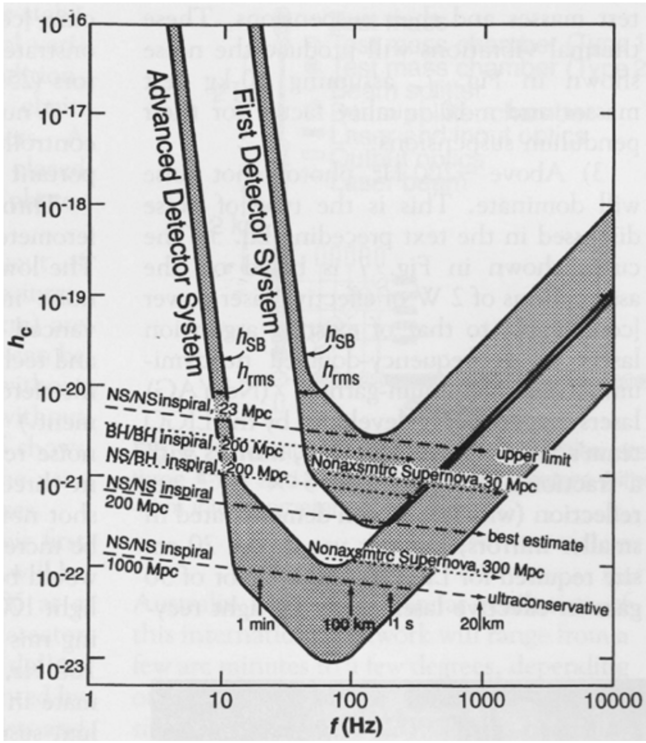

Figure 3. Illustration of concept of LIGO detector evolution for various burst sources. The curves labeled h_{rms} denote root mean square (rms) noise of a single detectors for a one-cycle-long burst at a particular frequency. The curves labelled h_{SB} are the average level of strain for each detector that could be confidently detected from rare events in arbitrary directions. Characteristic strengths of various sources are given by h_c, which is the signal amplitude at a particular frequency weighted by the square root of the number of cycles of the wave near that frequency. The labels NS and BH are for neutron star and black hole respectively (taken from ref. 4).

Uncorrected Proof

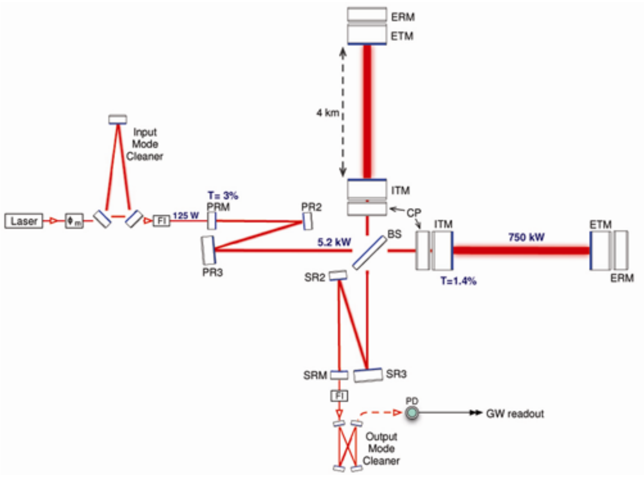

Figure 4. Advanced LIGO optical configuration.ITM, Input test mass; ETM, End test mass; ERM, End reaction mass; CP, Compensation plate; PRM, Power recycling mirror; PR2/PR3, Power recycling mirror 2/3; BS, 50/50 beam splitter; SRM, Signal recycling mirror; SR2/SR3, Signal recycling mirror 2/3; FI, Faraday isolator; φ_m, Phase modulator and PD, Photodetector. The laser power numbers correspond to full-power operation. All of the components shown, except the laser and phase modulator, are mounted in the LIGO ultra-high vacuum system on seismically isolated platforms.

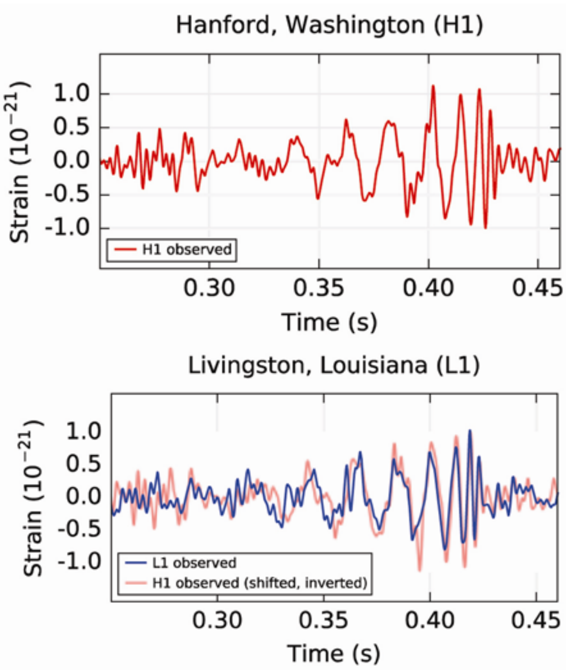

Figure 5. The gravitational wave event GW150914 observed by the LIGO Hanford (H1, top) and Livingston (L1, bottom) detectors. Times are shown relative to 14 September 2015 at 09 : 50 : 45 UTC. For visualization, all time series were filtered with a 35–350 Hz bandpass filter to suppress large fluctuations outside the most sensitive frequency band of the detector and band-reject filters to remove the strong instrumental spectral lines. Because of their different orientations, the signal from H1 should be the nearly the same as that from L1, but inverted because of the detector orientation and should arrive within ±10 ms. The pink curve in the bottom panel is the H1 signal flipped in sign and shifted by 7 ms (taken from ref. 12).

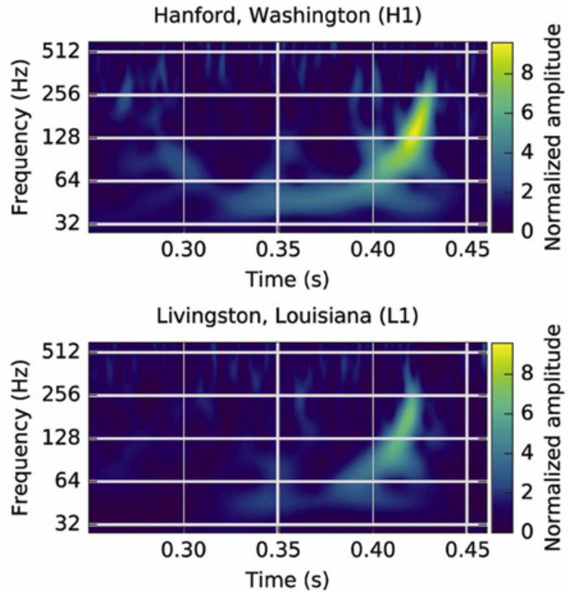

Figure 6. Time–frequency evolution of GW150914 in the H1 and L1 detectors (taken from ref. 12).

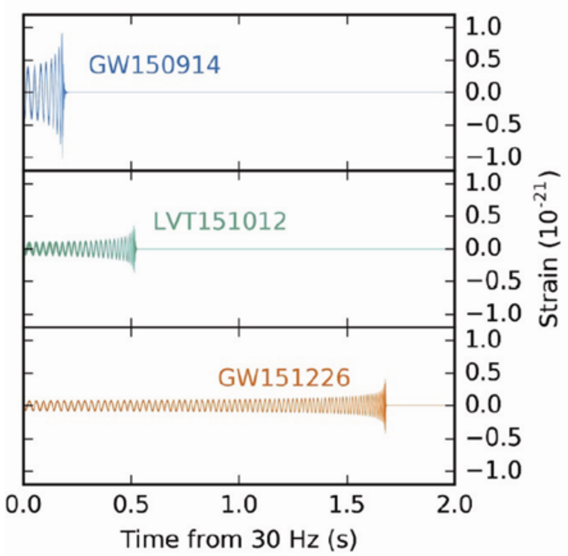

Figure 7. Time evolution of the recovered signals from O1, starting when they entered the sensitive band of the detectors at 30 Hz. Shown are the 90% credible regions of the LIGO Hanford signal reconstructions from a coherent Bayesian analysis using a non-precessing spin waveform model[15] (taken from ref. 14).

Uncorrected Proof

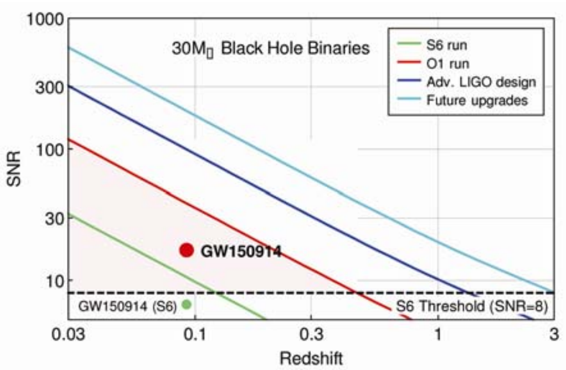

Figure 8. The single detector signal-to-noise ratio (SNR) under optimal orientation as function of redshift z for two merging black holes each 30 solar masses, shown for the first LIGO detector (green line), Advanced LIGO during the O1 run (red line), Advanced LIGO at design sensitivity (blue line) and Advanced LIGO following a possible upgrade (cyan line). GW150914 (red dot) was not optimally oriented and was detected with a single detector SNR of 13–20 at $z = 0.09$; had this event occurred in S6, it would have been below the S6 detection threshold (dashed line; taken from ref. 16).