

The first five years of gravitational-wave astrophysics

Salvatore Vitale^{1,2*}

¹LIGO Laboratory, Massachusetts Institute of Technology, 185 Albany St, Cambridge, MA 02139, USA

²Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA

November 10, 2020

Gravitational waves are ripples in the fabric of spacetime generated by the acceleration of astrophysical objects. A direct consequence of Einstein’s general relativity, they were directly observed for the first time in 2015 by the twin LIGO observatories. Five years later, tens of gravitational waves have been detected, emitted by pairs of merging compact objects such as neutron stars and black holes. These signals yield insights about the formation of compact objects and their progenitor stars; they enable stringent tests of general relativity and reveal the behavior of matter at supranuclear densities. Bright mergers, emitting both gravitational and electromagnetic waves, are unique probes of the formation of short gamma ray bursts and of heavy elements, and can be used to measure the local expansion rate of the universe.

^{**}To whom correspondence should be addressed; E-mail: salvatore.vitale@ligo.org

Introduction

Einstein's theory of general relativity (GR), published in 1915 (1), revolutionized our understanding of what gravity is and how it works. Rather than a force, gravity must be thought of as the effect of spacetime curvature on anything with mass-energy; that is, on everything. At the same time, mass and energy deform the spacetime in which they move and hence *are* the sources of gravity. The delicate interplay between such sources and geometry was famously captured in a quote by physicist J. A. Wheeler: "Spacetime tells matter how to move; matter tells spacetime how to curve" (2).

One of the new phenomena predicted by GR are gravitational waves (GWs) (3); ripples in the fabric of spacetime generated by astrophysical masses as they accelerate. While changes in the gravitational potential propagate instantaneously in Newton's theory, GWs in GR travel at the speed of light. Although the existence of wave-like solutions is a direct mathematical consequence of Einstein's field equations, scientists debated for decades about whether GWs are physically measurable (4). It was only in 1957 at the Chapel Hill conference that the debate was settled: R. Feynman and F. Pirani showed that a gravitational-wave train would deposit energy in a detector. These simple proof-of-principle ideas led to the development of resonant-bar GW detectors in the 60s. Detections were reported by J. Weber in 1969 and 1970 (5), but his findings could not be verified or replicated by independent researchers.

Ultimately, the first conclusive piece of evidence for the existence of GWs was an indirect one. Within GR, two objects in a binary system will lose energy through emission of GWs. This energy loss makes the orbit shrink, further enhancing the emission of GWs. This runaway process will ultimately end when the two objects collide, Fig. 1. In 1974 R. Hulse and J. Taylor reported the discovery of a binary neutron star system, with one of the two objects visible in the radio band as a pulsar. Data collected over the next 8 years showed that the orbit was

shrinking in perfect agreement with the predicted energy loss due to GW emission (6). Another significant prediction of GR, black holes, was verified in the same years: by the mid 1970s most astronomers were convinced that the X-ray source in Cygnus X-1 was most likely a black hole of roughly 15 solar masses (henceforth, $M_{\odot}$). If black holes existed, they could theoretically live in binary systems and emit even more GW energy than neutron star binaries. The sky was suddenly full of potential sources of GWs that Einstein could not have imagined.

Interferometric detectors

The thought experiments discussed at the Chapel Hill conference by Feynman and Pirani, and later by H. Bondi, made it clear that to reveal GWs one must monitor the distance between freely falling objects. Due to the quadrupole nature of GWs, objects along perpendicular directions are affected by the waves in the opposite ways: e.g. distances decrease along the x direction and increase along the y direction. While the idea is quite straightforward, the technological hurdles appeared insurmountable in the 1960s: folding in values for the realistic strength of potential GW sources it was found that the relative displacement induced by GWs – and hence the required precision of measurement – were of the order of 10^{-21} .

By the early 1970s, work by R. Weiss, R. Drever, and others had shown that the most promising way of detecting gravitational waves on Earth was by using laser interferometry (7). It would take another 3 decades of research and development before the Laser Interferometer Gravitational-wave Observatory (LIGO), which consists of two detectors in Louisiana and Washington, started taking data in 2002, followed by the Italian-French detector Virgo in 2007. These detectors took data until 2010, without reporting any detections, before being decommissioned and upgraded to their advanced configuration, which promised a factor of ten improvement in sensitivity.

Fig. 2 shows a schematic of the Advanced LIGO detectors (8). The light from a high-power

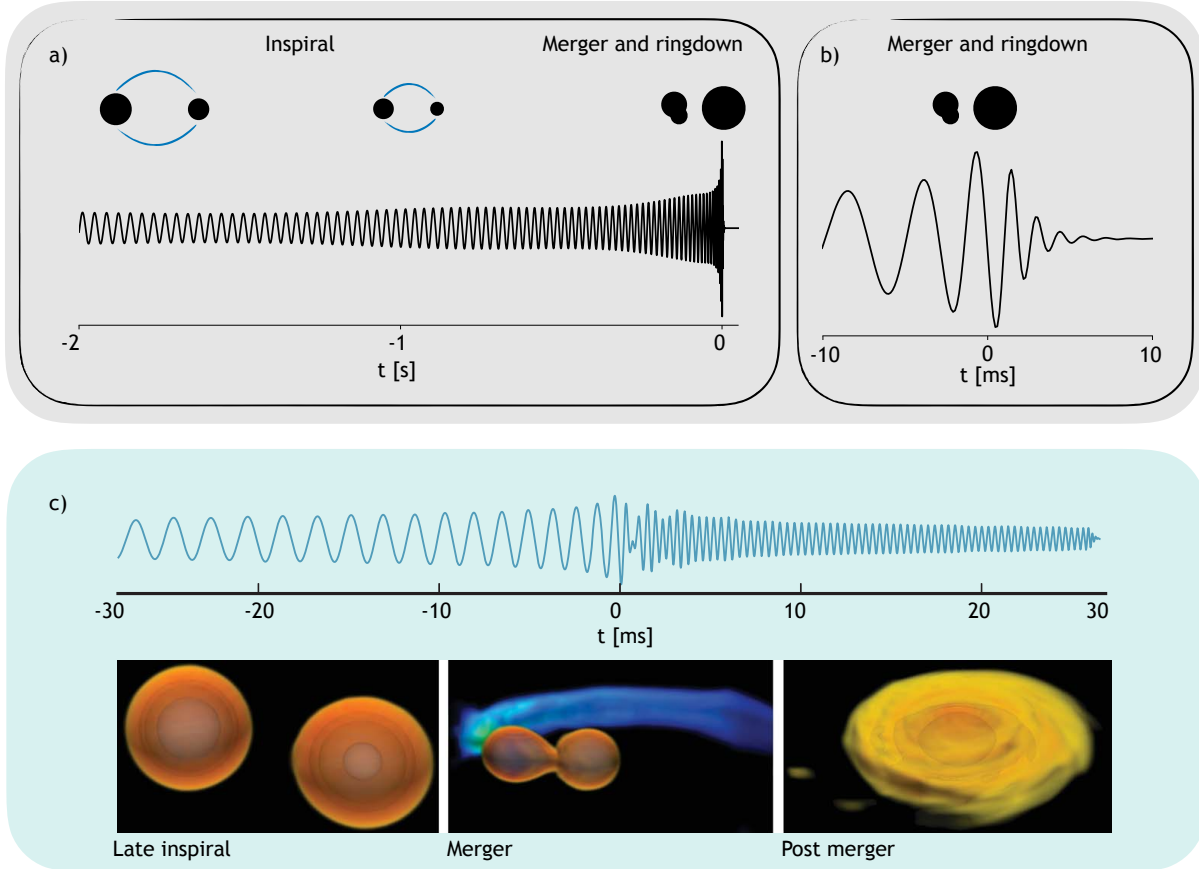


Figure 1: a) The last two seconds of the gravitational waveform produced by a binary black hole, together with a cartoon representation of the two compact objects. The clearly visible amplitude modulation in the inspiral phase is due to spin-induced precession of the orbital plane b) A zoom in the same signal in the 10 milliseconds around the merger time, showing the newly formed black hole ringing down to equilibrium. c) 30 milliseconds around the merger time of a gravitational wave produced by a binary neutron star merger. Post-merger high-frequency GW emission is clearly visible. The neutron stars are first tidally deformed and then disrupted in the late inspiral. A disk of matter is formed around the final compact object. (Credits for panel c, T. Dietrich)

laser is split and sent along the two perpendicular arms of a Michelson interferometer with Fabry-Perot arms. Each arm contains two mirrors, seismically isolated from the ground by 4-stage pendulum suspensions, which act as test masses, as well as by active and passive isolation systems. The transmittivity and reflectivity of the mirrors are chosen such that photons bounce back and forth ~ 300 times, before being recombined into a photodetector. Additional power recycling and signal recycling mirrors help to increase the sensitivity of the instrument. The detectors are operated such that, in the absence of a gravitational-wave signal, light from the orthogonal arms interferes (nearly, but not exactly) destructively at the photodetector. With a 4 km arm length, in order to detect a relative length variation of the order of 10^{-22} , the position of the mirrors must be known to better than the size of a proton. As unbelievable as this might sound, it can be done. The Advanced Virgo detector (9) works in a similar manner, with some different design choices (notably for the suspension of the mirrors), and a shorter arm length (3 km).

The discovery of gravitational waves

The Advanced LIGO detectors were scheduled to start their first observing run on the 18th of September 2015 (11). However, a few days before, on September 14 2015, an extremely loud and coherent signal was detected by the two LIGO detectors during the engineering run preceding the observing run. Extensive follow-up excluded the possibility that the signal was unintentionally or maliciously added into the data stream. Nothing that could have caused the signal was found in the hundreds of thousands of auxiliary channels that monitor everything from seismic activity to winds and cosmic rays. It was a real gravitational wave. On one side, the discovery successfully ended a scientific and technological quest that at times seemed to be far beyond the horizon: GWs had finally been directly detected, 99 years after Einstein derived the equations that govern them. From the other, it marked the beginning of a new era for physics,

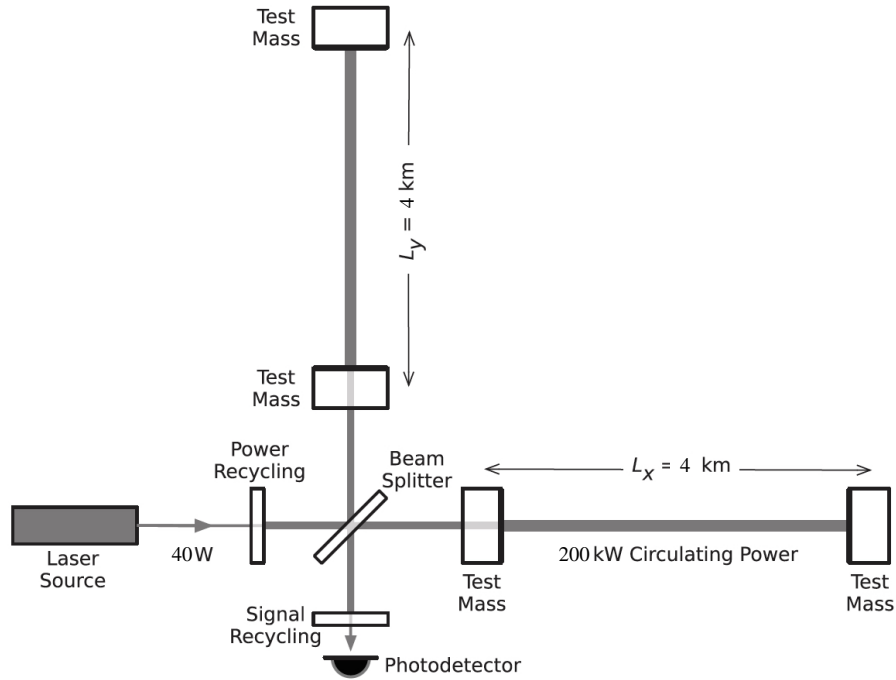


Figure 2: A schematic of the Advanced LIGO detector, adapted from Ref. (10)

astrophysics and cosmology.

GW150914, as this first source has been called, was quickly identified as the merger of two black holes, each of roughly $30 M_{\odot}$, colliding at half of the speed of light to form a new black hole, 1.4 billion light-years away from Earth. It was the first time that a binary black hole system was observed (12). Furthermore, black holes of $30 M_{\odot}$ came as surprise to many in the scientific community. Decades of observations following the discovery of Cygnus X-1 had painted a pretty clear picture: black holes in X-ray binaries seemed to exist in the range between roughly 5 and 15 solar masses, whereas the black holes found at the center of most galaxies had masses of millions or billions of solar masses. If the component black holes of GW150914 were the end-product of the lives of massive stars, then their progenitors must have been born in an environment very poor in metals (13) (note: using common astrophysical jargon, we refer to any element heavier than helium as metal). In fact, metal-rich stars are expect to lose more

of their initial mass during their life, owing to higher opacities and hence higher stellar winds, and might leave behind lighter black holes.

Two more binary black hole (BBH) mergers, one of which resulted in a weak signal, were detected in the first observing run, that concluded in January 2016. This showed that the detection of GW150914 was not a rare and fortunate accident (*14*). The BBH merger rate inferred from these discoveries, together with the expected increase in sensitivity of the advanced detectors (*11*), implied that LIGO and Virgo would detect many BBHs in their following observing runs. This is quite a dramatic change of perspective: it was clear that the transition from the first-ever GW detection to the era of routine detections was not going to be spread over a long period of time, but only take a few years.

Indeed, in the two following observing runs – O2 (2016-2017) and O3 (2019-2020) – the number of detections has continued to grow. Fig. 3 shows the cumulative number of confirmed and candidate GW sources that the LIGO-Virgo Collaboration (LVC) has found in the first three observing runs. The clear change of slope between O2 and O3 was driven by improvements in the sensitivity of the instruments, including the implementation of quantum squeezing in LIGO (*15*) and Virgo (*16*).

The most recent LVC catalog reports 39 significant detections (*17*) found in the first half of the O3 data (O3a), most of which are BBHs. Additional detections, usually with low signal-to-noise ratios, have been reported by groups not affiliated with the LVC, using public data from the first two observing runs (*18–21*).

The life and death of stellar mass black holes

Where do the BBHs that LIGO and Virgo have detected come from, and how did they form?

From a theoretical point of view, the two most likely avenues to form compact binary systems are isolated formation in galactic fields (*23, 24*) and dynamical encounters in dense envi-

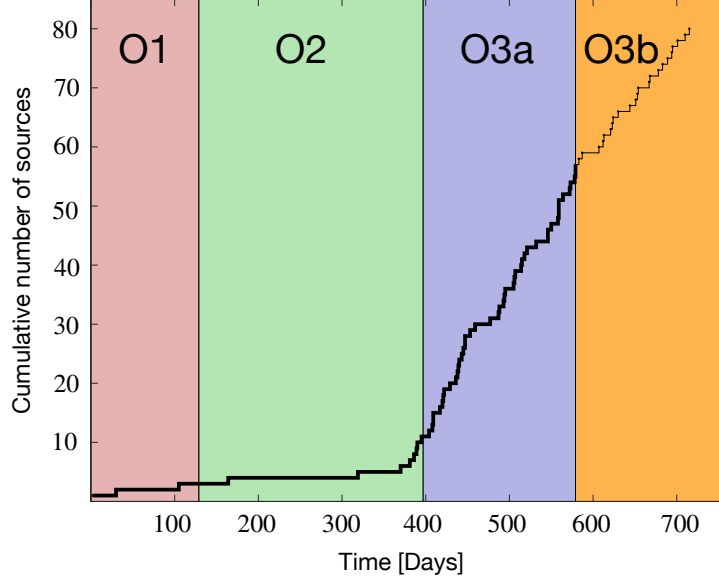


Figure 3: The cumulative number of detections (thick line) and non-retracted public alerts (thin line) in the first three observing runs (the third observing run, O3, is split in two parts, O3a and O3b). The significant change of slope between O2 and O3 is due to the increased sensitivity of the detectors. (Adapted from Ref. (22)).

ronments such as globular clusters (25, 26).

In the first scenario, the progenitors of the two black holes are two massive stars in a stellar binary. After the first star becomes a giant and overfills its Roche lobe, a common envelope of gas is formed around both stars, which is expected to shrink the binary's orbit. When the stars in the binary run out of fuel, the two black holes are left in a tight orbit, which can merge in less than a Hubble time. The two black holes in binaries formed this way are expected have similar masses, and spin vectors nearly aligned with the orbital angular momentum.

Dynamical formation instead assumes that the two black holes that eventually merge do not share a common past: rather, they formed in a region so densely populated by black holes that the probability of randomly meeting a partner and forming a bound binary is not small.

Globular clusters might be the ideal arena for this to occur, with stars near the center of the cluster separated by distances comparable to the size of our solar system. As the cluster evolves from its initial stellar population, black holes form with timescales of millions of years and, owing to dynamical friction, migrate toward the center of the cluster, where two-body and three-body interactions are common. The two black holes in a binary formed through this channel are also expected to have mass ratios close to unity, but their spin vectors should be randomly oriented, owing to the random nature of the encounter.

The mass distribution of stars heavier than the Sun is well matched by a power law, $p(m) \sim m^\alpha$, with a slope $\alpha \in [-2.7, -2.3]$ (27), and it is expected that the masses of the black holes they generate will have the same functional form. However, the positions of both the lower and upper ends of this distribution are not known, and are the subject of intense research.

Theoretical predictions based on nuclear physics suggest that the relation between progenitor mass and black hole mass is not simply monotonic. If the progenitor star mass is in the range $\sim [100 - 130] M_\odot$, a runaway production of electron-positron pairs at the core triggers a series of contractions and expansions which eject the outer layers of the star. The ultimate stage of these stars' lives is a supernova that will leave behind a black hole of $\lesssim 50 M_\odot$, irrespective of the progenitor mass. If the progenitor is larger than $\sim 130 M_\odot$ but smaller than $\sim 250 M_\odot$, the explosion triggered by the first compression is so violent that it entirely disrupts the star, leaving no black hole behind. Finally, for stars more massive than $\sim 250 M_\odot$, nothing can prevent the gravitational collapse and a black hole is formed without any explosion: the resulting black hole will be as massive as the core of the progenitor. If these models are correct, there should exist a pair-instability supernovae (PISN) “mass gap” in the mass range $\sim [50 - 150] M_\odot$, where black holes cannot form from stellar collapse (28).

At the other end of the mass spectrum, black holes discovered in X-ray binaries have masses larger than $\sim 5 M_\odot$, which provides some observational evidence for a mass gap between

neutron stars and stellar mass black holes (29,30). If the gap exists, it could provide information about the physics of supernova explosions and stellar binary evolution (31).

Confirmations and surprises

Are these theoretical expectations representative of the true properties of black holes in nature?

The GW sources detected in the first two observing runs (32) and the first part of the third observing run (17) have started to provide answers. In their most recent analysis (33), the LVC has considered many analytical models for the mass distribution of black holes in binaries, with different levels of complexity. The model that appears to be favored by the data describes the mass distribution of the primary (i.e. most massive) black hole in the detected binaries as the sum of two components: a power-law distribution and a Gaussian distribution ¹. The power-law component is included to model the black holes that are generated by the death of massive stars, whereas the Gaussian component can describe the black holes that might pile up at the bottom end of the pair-instability gap. In total, this model depends on 8 unknown parameters, including the minimum and maximum mass generated by the power-law component and its slope, and the mean and standard deviation of the Gaussian component and the fraction of BBHs that belong to it. The measured posterior distribution of the overall mass distribution, as well as the contributions from the two components, are shown in Fig. 4.

The LVC measures $\alpha = -2.62^{+0.62}_{-0.73}$ (median and 90% credible intervals unless stated otherwise) for the slope of the primary black hole mass distribution power law, consistent with that of the stellar population. Moreover, they find that in most BBHs, the two components objects have similar masses, consistent with what is expected for the main formation channels described above. The Gaussian component has a mean and standard deviation equal to

¹Unless otherwise stated, I will be quoting the results of Ref. (33) from the “power-law + peak” analysis which excludes GW190814.

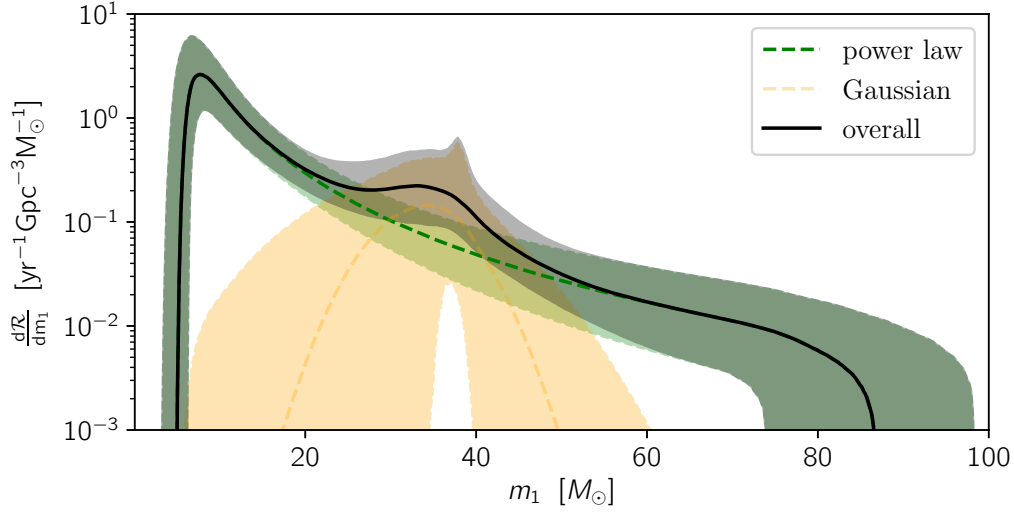


Figure 4: Posterior distribution of the primary (i.e. most massive) black hole mass distribution, from the LVC data release (33). This measurement models the underlying overall population as the sum of a power law and a Gaussian component, it includes GW190521, and it accounts for selection effects. The thick lines represent the medians, whereas the shaded areas report the 90% credible intervals. See Ref. (33) for details on the analysis.

$\mu_G = 33.49^{+4.54}_{-5.51} M_\odot$ and $\sigma_G = 5.09^{+4.28}_{-4.34} M_\odot$ respectively, and it contains $10^{+14}_{-7} \%$ of the BBHs.

Interestingly, models that do not allow for the existence of a Gaussian component are disfavored by the data, with odds of 1 to ~ 4 or worse depending on the specific alternative model. Using the model favored by the data, the LVC measures the value of the 99th percentile of the mass distribution of the underlying black hole population, finding $m_{99\%,\text{BH}} = 59.6^{+14.5}_{-16.1} M_\odot$.

The measurement of the maximum mass is significantly driven by GW190521 (34), perhaps the most consequential BBH detection after GW150914. With component masses of $m_1 = 90.9^{+29.1}_{-17.3} M_\odot$ and $m_2 = 66.3^{+19.3}_{-20.3} M_\odot$, GW190521 is the heaviest BBH discovered to date. Both of its component objects are consistent with being in the PISN mass gap, and the black hole formed by the merger constitutes the first ever clear evidence for the existence of intermediate-mass black holes (IMBHs), the missing link between stellar mass black holes and the supermassive black holes found at the centers of most galaxies (35).

Though the GW data do not unequivocally prove this is the case, it is possible that one or both of the component objects of GW190521 were in turn the result of a previous merger (36), which could explain their unusually high masses. As mentioned above, dynamical formation can happen in regions of the universe with a high density of black holes. If the black hole created by a first-generation merger does not acquire a recoil velocity larger than the escape velocity of its environment, it can be retained and can merge again. This might happen more easily in high escape-velocity environments such as dense star clusters and the disks of gas around supermassive black holes. Measuring how frequently this hierarchical growth can lead to IMBHs would have dramatic consequences on our understanding of structure formation in the universe. The repeated merger of IMBHs could be the first step of the formation of supermassive black holes of billions of solar masses, which already existed when the universe was only a few hundred million years old (35). Future GW detections will certainly shed light on the origin and properties of IMBHs. If one believes that GW190521 is the result of a second-generation merger, or in any way an outlier, then it makes sense to remove it while analyzing the properties of the rest of the BBH population: doing so yields $m_{99\%,\text{BH}} = 52.1^{+13.8}_{-7.3} M_{\odot}$ (33), close to the edge of the putative PISN mass gap.

The measurement of the lower edge of the black-hole mass distribution is similarly driven by one source, GW190814 (37). GW190814 was produced by the merger of a $23.2^{+1.1}_{-1.0} M_{\odot}$ black hole with a $2.6^{+0.1}_{-0.1} M_{\odot}$ compact object. This latter object is either the lightest black hole or the heaviest neutron star ever detected. If GW190814 is the merger of a black hole and a neutron star, then it would be the first ever discovered. Unfortunately, it is hard to be sure: owing to the large mass ratio of GW190814, even if the lighter object were in fact a neutron star, it would not be tidally deformed before crossing the event horizon of its companion. In turn, that implies that such a merger would not be accompanied by electromagnetic (EM) emission, which requires matter to be left out from the event horizon (EM counterparts to GW sources are further

discussed below). Therefore, the facts that the LVC could not constrain the tidal deformation of the lighter object, and that no EM counterparts have been reported to date, do not reveal the nature of the lighter object. A $2.6^{+0.1}_{-0.1} M_{\odot}$ neutron star is likely too heavy to be supported by a realistic equation of state of nuclear matter (37), and only future GW detections might clarify if such an object can exist. By treating both of the objects in GW190814 as black holes, the LVC measures the 1st percentile of the black hole mass distribution to be $m_{1\%,\text{BH}} = 2.5^{+0.3}_{-0.3} M_{\odot}$, narrowing or eliminating the putative mass gap between neutron stars and black holes. On the other hand, if GW190814 is treated as a neutron star black hole merger, and excluded from the BBH analysis, they obtain $m_{1\%,\text{BH}} = 6.0^{+1.0}_{-1.6} M_{\odot}$, remarkably consistent with what observed in X-ray binaries (33).

A litte rotation

The spins of black holes in binaries carry information about their formation pathways. Unfortunately, measurement of the component spins is usually less precise than that of the component masses (38, 39). Luckily, one can still draw significant astrophysical inferences using auxiliary parameters: χ_{eff} (40) is related to the component of the total spin along the orbital angular momentum and is usually measured much more precisely than either component spin. On the other hand, χ_p (41) contains information about the component of the total spin along the orbital plane and is usually poorly constrained. For BBHs formed in galactic fields, the individual spins should be preferentially aligned with the angular momentum, meaning that χ_{eff} should be positive, whereas BBHs formed dynamically have random spin orientations, meaning that χ_{eff} can have both signs.

While the results from LIGO-Virgo’s first two observing runs were consistent with a distribution of χ_{eff} centered at zero², adding the new sources from O3a suggests a different scenario.

²However, see Ref. (19) which reports a highly spinning BBH in LIGO’s public data, though the signifi-

The latest LVC analysis (33) reports a median for the χ_{eff} distribution of $\mu_{\text{eff}} = 0.06^{+0.05}_{-0.05}$, indicating that not all of the black holes have vanishing spin, though spins are typically small. This is particularly interesting, as recent studies suggested that most black holes are born with negligible spins (44). Even more significantly, the LVC measures that 27^{+17}_{-15} % of the BBHs have negative χ_{eff} . This is a major development in our understanding of binary black hole formation, as it might indicate that a significant fraction of them are formed dynamically (45).

While a precise measurement of the level of misalignment between the orbital angular momentum and the component spins remains elusive, these latest results show that a configuration where all of the spins are aligned with the angular momentum is excluded at 99% credibility. Misalignment between the spin and the orbital angular momentum results in orbital precession, which induces a characteristic phase and amplitude modulation in the GW signal (46), see Fig. 1 panel a. However, the evidence for misalignment in the current dataset is driven by weak unresolved features in many sources, rather than by significant precession in a few of them.

Future detections will certainly help increase our understanding of the main formation channels and the properties of the BBHs they generate. At the same time, there will be more oddballs that challenge our understanding in new and exciting ways. GW190412 is a remarkable representative of this category, for many reasons (47). The spin of the most massive black hole in the binary is large and positive, $a_1 \in [0.4, 0.6]$ ³ (depending on the specific waveform model used), and some evidence exists for orbital precession. Its large mass ratio, $\sim 4 : 1$, also sets it apart from most of the other BBHs detected to date (17, 32, 48), fueling speculations about its origin, e.g. (49–52).

cance (42) and parameters (43) of the source are still being debated.

³We quote the dimensionless spin parameter, defined in the range 0 (no spin) to 1 (maximally spinning).

Let there be light

Detecting GWs and light from the same source has historically been one of the most exciting prospects for gravitational-wave astrophysics. The strong tidal pull of the companion is expected to deform and disrupt the neutron star, generating a tidal tail of debris which travels at mildly relativistic speeds. Shortly after the merger, an accretion disk forms around the newly created compact object, see Fig. 1 panel c. BNS mergers have long been believed to generate short gamma ray bursts (sGRBs), powered by matter accretion from the disk onto the central compact object and narrowly beamed along the direction of its magnetic field (53). As the gamma ray jet loses energy to the interstellar medium, the beaming angle increases and the EM emission moves to lower frequencies (54). Meanwhile, the neutron-rich material surrounding the compact object (referred to as mass ejecta) is predicted to produce large amounts of radiation, generated by the decay of heavy nuclei, and powers a *kilonova* visible in the optical band for the hours following the merger (55).

All of these predictions have been verified with the momentous discovery of GW170817 (10, 56), the first BNS merger detected by the LIGO and Virgo observatories. The GW signal was first found by a real-time search pipeline as a significant BNS candidate in the LIGO Hanford data only, due to the presence of a transient instrumental noise artifact at around the time of the event in the LIGO Livingston data and issues with the transfer of Virgo data. Six minutes earlier, the *Fermi* Gamma-ray Space Telescope had issued an alert reporting the discovery of an sGRB roughly 1.7 seconds after the merger time of the putative BNS event (57). 5 hours later, the LVC produced an updated three-detector sky map, having cleaned the LIGO Livingston data and having acquired Virgo data. The BNS sky map produced by the LVC was entirely contained in the GRB sky map produced by the *Fermi* collaboration, and it constrained the location of the source to an ellipse of $\sim 31\text{deg}^2$, at a distance of 130 million light years (10, 58).

Virtually all of the telescopes on and orbiting around the planet started observing that patch of sky, and less than 11 hours after the merger time a bright optical source was identified near the galaxy NGC 4993, which was not present in archival data (58). Extensive follow-up detected emission in the ultraviolet, infrared, X-rays and radio (58–63), and subsequent analyses showed that the evolution of the EM emission was unlike that of any other known astrophysical transient (64): it was the first time humans observed light and gravitational waves from the same source.

GW17017 has truly been a gold mine for nuclear physics, astrophysics and cosmology. From the gravitational-wave side, signals from BNSs are morphologically more rich than those from black hole binaries (65), see Fig. 1. In the late phase of the inspiral, tidal deformation deposits energy into the neutron stars, affecting the phasing evolution of the gravitational wave (66). The compact object formed by the collision can be a stable neutron star, an unstable neutron star that will collapse into a black hole after a timescale of milliseconds to seconds, or directly into a black hole (67). Each possible outcome results in a different set of observables in the post-merger gravitational-wave signal (68) (as well as in the EM emission). Both the magnitude of the inspiral dephasing and the details of the post-merger GW emission are critically dependent on the unknown equation of state (EOS) of cold nuclear matter (69). This implies that compact binary mergers can be used to measure the EOS of neutron stars in a way which is entirely orthogonal to what can be done with EM-based methods (e.g. by the Neutron Star Interior Composition Explorer - NICER - telescope (70)).

Unfortunately, tidal effects are larger at high frequencies (above a few hundred Hertz) where the sensitivity of LIGO and Virgo is low. Therefore, unless additional constraints are imposed, only an upper bound can be set on the tidal deformability of the compact objects in GW170817 (10, 56). This means that based on the GW data alone one cannot rule out that both objects have zero tidal deformability, i.e., that they are black holes. Nevertheless, the detection

and energetics of the EM counterparts and the estimated compact object masses are strongly suggestive of a BNS merger. The measurement of the tidal parameters can be improved if one *assumes* that both objects are neutron stars, governed by the same equation of state. With those assumptions, the LVC obtains a measurement of $\Lambda_{1.4M_\odot} = 190^{+390}_{-120}$ for the tidal deformability of a $1.4 M_\odot$ neutron star (71), which disfavors large neutron star radii (72). Indeed, using empirical relations between the tidal deformability, the masses, and the radius (73), the constraint on the tidal deformability can be recast into a measurement of the radii of the two compact objects in GW170817, $R_1 = 11.8^{+2.7}_{-3.3}$ km and $R_2 = 10.8^{+2.9}_{-3.0}$ km (71). The measurements can be improved by requiring that the EOS supports NS masses as large as $1.97 M_\odot$, a conservative estimate for the mass of the heaviest NS known at the time of the analysis, obtaining $R_1 = R_2 = 11.9^{+1.4}_{-1.4}$ km (71) (the fact that the two radii are equal arises from the analysis and is not assumed by the analysts). Finally, these constraints can be converted into a measurement of the pressure inside the neutron star, obtaining a value of $p(2\rho_{\text{nuc}}) = 3.5^{+2.7}_{-1.7} \times 10^{34}$ dyn/cm² for the pressure at twice the nuclear saturation density (71).

A clear testament of the truly multimessenger nature of GW170817 is the fact that the constraints on tidal deformability can be also improved by exploiting EM information. The basic idea is that modeling the EM emission at various frequencies provides an estimate of the amount of mass ejecta produced during the merger. That, in turn, depends on physical properties of the merger, such as the neutron star masses and, critically, the tidal deformability. By requiring that enough ejecta is produced to match the observed EM emission, one can obtain a lower limit for the tidal deformability (74). Recently, several groups have reported improved constraints obtained by combining GW and NICER data (75, 76), and accounting for the large mass of a newly discovered pulsar, likely above $2 M_\odot$ (77). While the results presented in these analysis are somewhat model-dependent, they clearly show the potential for precise and accurate measurement of the NS EOS in the forthcoming years.

The list of significant results associated with the successful EM follow-up campaign of GW170817 is too long to be explored here (see e.g. (78) and references therein). We will only stress that these observations have proven that at least a fraction of the sGRBs are produced by merging BNSs and have offered insights into the geometry of the relativistic jet (57). They have shown that BNS mergers generate kilonova emission, whose description might require two or more components, with different energetics and neutron-fractions (63). In turn, this implies that a significant fraction of the heavy elements in the universe is produced by BNS mergers, rather than SNe explosions (79).

The LVC has reported the detection of a second BNS merger, discovered on the 25th of April 2019, during the third observing run (80). Unfortunately, the fact that the source was very far and poorly localized made the quest for an EM counterpart extremely challenging, and none has been reported. In spite of that, and of its quite low signal-to-noise ratio, the BNS that generated GW190425 is a remarkable source since its total mass, $M = 3.4_{-0.1}^{+0.3} M_{\odot}$ (80), makes it the heaviest BNS ever discovered, including known galactic binary pulsars, whose total mass distribution ends at $M \sim 2.88 M_{\odot}$ (81). The high mass of GW190425 is not easy to explain within the traditional formation channels. The classic picture of formation in an isolated binary struggles to explain how the system survived the supernovae that generated the neutron stars, since these explosions are more disruptive for more massive progenitor stars. Meanwhile, dynamical formation in dense stellar environments seems unlikely: while black holes sink toward the center of the potential, neutron stars are expected to live in the less-densely populated outskirts, where they are less likely to meet a companion (82). Indeed, numerical simulations suggest a merger rate for BNS in dynamical encounters much lower than that inferred by the LVC. A possible way to form GW190425 in galactic fields is to assume that the secondary object went supernova when the binary was already in a tight orbit, too strongly bound to be ripped apart by the explosion. In this scenario, the resulting BNS would merge in only a few

million years. The fast timescale to merger, together with the significant orbital acceleration of the neutron stars, could also explain why these systems have not been previously observed by galactic radio surveys (80, 83). Others have explored the possibility that GW190425 was in fact a neutron star black hole merger (84, 85). The debate is far from settled, and future detections will certainly illuminate the mass distribution and the origin of neutron stars in binaries.

Cosmology and strong-field tests of general relativity

A fundamental result associated with GW170817 is a new measurement of the present value of the Hubble parameter (henceforth “Hubble constant”) (86, 87). A measurement of the Hubble constant, and hence of the local expansion rate of the universe, requires the knowledge of both the redshift and luminosity distance of astrophysical objects. While EM data naturally provides the former, it does not carry any information about the latter. In the decades preceding the discovery of GW170817, two methods have been used to circumvent this obstacle: building a “cosmic distance ladder” and then using type Ia supernovae as standard candles; or inferring the Hubble constant from the power spectrum of the cosmic microwave background. Those two methods currently yield measurements of the Hubble constant which are in tension at a larger than 5σ level (88). Various explanations have been proposed, from systematic errors to new physics (88), while more recent measurements have not solved the puzzle (89).

Unlike EM waves, GWs encode information about the luminosity distance of their astrophysical sources (86), and could help arbitrate the existing tension. For BNS sources for which an EM counterpart is detected, one can thus obtain a distance estimation from the gravitational waves and a redshift estimation from the host galaxy. This idea was successfully demonstrated with GW170817, obtaining a measurement of the Hubble constant $H_0 = 70.0^{+12.0}_{-8.0} \text{ km s}^{-1} \text{ Mpc}^{-1}$ (90). While this measurement is still too uncertain to help resolve the existing tension, it will improve over the next few years, as more sources are detected, and

might reach percent-level precision in a five-year timescale (91).

Another remarkable outcome of the discovery of GW170817 and the associated GRB is a constraint on the speed of GWs. By using the time-delay between the arrival of the GRB and the GWs here at Earth, it was inferred that the relative difference between the speed of GWs and the speed of light is in the range $[-3 \times 10^{-15}, +7 \times 10^{-16}]$ (57). This constraint represents the most solid piece of evidence that GWs travel at the speed of light, as predicted by GR. More generally, GWs are an optimal tool to test GR, since they probe regions of the universe where space-time curvature and gravitational potential are both large (92). The exquisite precision with which one can measure the phasing evolution of compact binaries allows testing of any non-GR effects that might affect the orbital evolution, e.g. the emission of dipolar radiation, or the propagation of the waves, e.g. the existence of a massive graviton. A measurement of the GWs emitted as a newly formed black hole “rings” down to equilibrium, Fig. 1 panel b, offers an unprecedented and direct access to the dynamic of event horizons.

The GW detections made public to date do not reveal any hints that GR might be violated (93, 94). Both tests that look for specific violations, and unmodeled tests that only check whether the detected signals are consistent with GR, show no issues with the theory. The new upper limit on the mass of the graviton, $m_g \leq 1.76 \times 10^{-23} \text{ eV}/c^2$, is nearly two times tighter than that obtained with solar system experiments (94). Observing the birth of the heaviest black holes in O3a reveals that they ring down just as expected for Kerr black holes in GR (94). No sign of non-tensorial polarizations, or black hole “mimickers” – exotic compact objects such as boson stars (95) – have been found (94). While most researchers on the planet who focus on astrophysics and high-energy physics are looking for hints of new physics, the old theory of GR still suffices to explain all the features observed in the GW data.

The next five years

LIGO and Virgo are currently being upgraded, and are expected to start the fourth observing run in late 2021 or early 2022 (11). Their sensitivity will be such that more than 100 BBHs and many BNSs should be detected every year (11). At the same time, the global GW network will grow, with the addition of the Japanese detector Kagra (96) in the fourth observing run, and LIGO India (97) in the next few years. Work is already ongoing to deliver the next significant boost in sensitivity: the “A+” LIGO configuration will ensure another factor of 2 in sensitivity (98).

Not only will these more sensitive instruments reveal more mergers of binary neutron stars and binary black holes, but they will detect them at larger distances from Earth, and with larger signal-to-noise ratios. The precision with which we will be able to study the birth and deaths of compact objects will increase significantly, as will our ability to test general relativity, probe the behavior of dense nuclear matter, and measure the expansion rate of the universe. Improved sky localization will increase the number of compact binaries for which EM counterparts are found, enabling more systematic studies of the EM emission at all wavelengths, including the energetics and structure of kilonovae.

Meanwhile, new territory will be charted. The discovery of GW190521 has shown that black holes can exist in the supernovae pair-instability mass gap: the next few years will enable measuring the mass distribution of black holes in that region, and illuminate their origin. On the other edge of the mass spectrum, it will become more clear how common massive binary neutron stars such as GW190425 are, and how they form.

The probability of detecting new types of sources will also increase. The most likely new type of source yet to be detected is the merger of a neutron star and a black hole, assuming that GW190814 was not one such system. Rare events such as core collapse supernovae are likely

to be detectable, if they happen in the Milky Way, and could be observable with gravitational waves, photons and neutrinos. It is likely that the stochastic GW signal produced by thousands of weak unresolved binary mergers will be detected in the near future, complementing the analysis of resolved sources. Finally, continuous near-monochromatic GW emission from galactic pulsars might also become observable in the next five years, depending on the ellipticity of neutron stars.

While ground-based detectors will explore the explosions and mergers of stellar-mass objects, another experiment based on Pulsar Timing Array (PTA), might detect gravitational waves from the merger of supermassive black holes. Those mergers might happen after two galaxies collide, and the supermassive black holes at their core form a bound system. PTA uses a network of tens of galactic millisecond pulsars as ultra precise clocks. By monitoring the arrival times of the radio beams from these pulsars, the presence of gravitational waves with frequencies in the nano-Hertz region – that is, with wavelengths of the order of 10 parsecs – can be revealed. Both individual mergers and the stochastic background generated by unresolved sources may be detected. The latest results of the NANOGrav collaboration, obtained with 12.5 years of PTA data, have revealed an excess of power which might be the first hint of the stochastic background GW signal (99).

In 1922, Arthur Eddington mockingly claimed that gravitational waves traveled “at the speed of thought” (100), after noticing that some wave-like solutions to Einstein’s equations could have arbitrary speed. Only five years after the first-ever direct detection of gravitational waves, we now know that in fact they travel at the speed of light, and bring precious information about remote corners of our universe.

Acknowledgments

I would like to thank the LIGO-Virgo-Kagra Collaboration, and in particular C. Berry, S. Biscoveanu, M. Fishbach, P. Fritschel, G. Mansell, L. McCuller, R. Weiss and A. Zimmerman for useful comments and discussions throughout the preparation of this manuscript. I am also grateful to E. Kara, D. Kaiser and P. Shanahan for comments on a mature version of the manuscript. I acknowledge support of the National Science Foundation and the LIGO Laboratory. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0757058. This is LIGO Document P2000387.

References

1. A. Einstein, *Sitzungsber. K. Preuss. Akad. Wiss. I*, 844 (1916).
2. C. W. Misner, K. S. Thorne, J. A. Wheeler, *Gravitation*, Princeton University Press, ed. (2018).
3. A. Einstein, *Sitzungsber. K. Preuss. Akad. Wiss. I*, 154 (1916).
4. D. Kennefick, *Traveling at the speed of thought: Einstein and the quest for Gravitational waves*, Princeton University Press, ed. (2007).
5. J. Weber, *Phys. Rev. Lett.* **22**, 1320 (1969).
6. J. H. Taylor, J. M. Weisberg, *ApJ* **253**, 908 (1982).
7. R. Weiss, *Quarterly Report of the Research Laboratory for Electronics, MIT Report No. 105* (1972).
8. J. Aasi, *et al.*, *Classical and Quantum Gravity* **32**, 074001 (2015).

9. F. Acernese, *et al.*, *Classical and Quantum Gravity* **32**, 024001 (2014).
10. B. P. Abbott, *et al.*, *Phys. Rev. Lett.* **119**, 161101 (2017).
11. B. P. Abbott, *et al.*, *Living Reviews in Relativity* **21**, 3 (2018).
12. B. Abbott, *et al.*, *Phys. Rev. Lett.* **116**, 061102 (2016).
13. B. P. Abbott, *et al.*, *ApJ* **818**, L22 (2016).
14. B. Abbott, *et al.*, *Phys. Rev. X* **6**, 041015 (2016). [Erratum: *Phys. Rev. X* **8**, 039903 (2018)].
15. M. Tse, *et al.*, *Phys. Rev. Lett.* **123**, 231107 (2019).
16. F. Acernese, *et al.*, *Phys. Rev. Lett.* **123**, 231108 (2019).
17. R. Abbott, *et al.*, *arXiv:2010.14527* (2020).
18. T. Venumadhav, B. Zackay, J. Roulet, L. Dai, M. Zaldarriaga, *Phys. Rev. D* **101**, 083030 (2020).
19. B. Zackay, T. Venumadhav, L. Dai, J. Roulet, M. Zaldarriaga, *Phys. Rev. D* **100**, 023007 (2019).
20. A. H. Nitz, *et al.*, *ApJ* **872**, 195 (2019).
21. A. H. Nitz, *et al.*, *ApJ* **891**, 123 (2020).
22. LIGO-Virgo Collaboration, <https://dcc.ligo.org/LIGO-G1901322/public> (2020).
23. L. L. Smarr, R. Blandford, *ApJ* **207**, 574 (1976).
24. V. Kalogera, K. Belczynski, C. Kim, R. W. O’Shaughnessy, B. Willems, *Phys. Rept.* **442**, 75 (2007).

25. E. S. Phinney, S. Sigurdsson, *Nature* **349**, 220 (1991).
26. M. J. Benacquista, J. M. B. Downing, *Living Reviews in Relativity* **16**, 4 (2013).
27. P. Kroupa, *Science* **295**, 82 (2002).
28. S. E. Woosley, *ApJ* **836**, 244 (2017).
29. C. D. Bailyn, R. K. Jain, P. Coppi, J. A. Orosz, *ApJ* **499**, 367 (1998).
30. W. M. Farr, *et al.*, *ApJ* **741**, 103 (2011).
31. F. Özel, D. Psaltis, R. Narayan, J. E. McClintock, *ApJ* **725**, 1918 (2010).
32. B. P. Abbott, *et al.*, *ApJ* **882**, L24 (2019).
33. R. Abbott, *et al.*, *arXiv:2010.14533* (2020). Data release at <https://dcc.ligo.org/LIGO-P2000434/public>.
34. R. Abbott, *et al.*, *Phys. Rev. Lett.* **125**, 101102 (2020).
35. J. E. Greene, J. Strader, L. C. Ho, *arXiv:1911.09678* (2019).
36. R. Abbott, *et al.*, *ApJ* **900**, L13 (2020).
37. R. Abbott, *et al.*, *ApJ* **896**, L44 (2020).
38. S. Vitale, R. Lynch, J. Veitch, V. Raymond, R. Sturani, *Phys. Rev. Lett.* **112**, 251101 (2014).
39. M. Pürrer, M. Hannam, F. Ohme, *Phys. Rev. D* **93**, 084042 (2016).
40. É. Racine, *Phys. Rev. D* **78**, 044021 (2008).
41. P. Schmidt, F. Ohme, M. Hannam, *Phys. Rev. D* **91**, 024043 (2015).

42. G. Ashton, E. Thrane, *MNRAS* **498**, 1905 (2020).
43. Y. Huang, *et al.*, *arXiv e-prints* p. arXiv:2003.04513 (2020).
44. J. Fuller, L. Ma, *ApJ* **881**, L1 (2019).
45. C. L. Rodriguez, C.-J. Haster, S. Chatterjee, V. Kalogera, F. A. Rasio, *ApJ* **824**, L8 (2016).
46. T. A. Apostolatos, C. Cutler, G. J. Sussman, K. S. Thorne, *Phys. Rev. D* **49**, 6274 (1994).
47. R. Abbott, *et al.*, *Phys. Rev. D* **102**, 043015 (2020).
48. J. Roulet, M. Zaldarriaga, *MNRAS* **484**, 4216 (2019).
49. C. L. Rodriguez, *et al.*, *ApJ* **896**, L10 (2020).
50. I. Mandel, T. Fragos, *ApJ* **895**, L28 (2020).
51. D. Gerosa, S. Vitale, E. Berti, *Phys. Rev. Lett.* **125**, 101103 (2020).
52. C. Kimball, *et al.*, *ApJ* **900**, 177 (2020).
53. S. I. Blinnikov, I. D. Novikov, T. V. Perevodchikova, A. G. Polnarev, *Soviet Astronomy Letters* **10**, 177 (1984).
54. E. Berger, *ARA&A* **52**, 43 (2014).
55. B. D. Metzger, *Living Reviews in Relativity* **20**, 3 (2017).
56. B. P. Abbott, *et al.*, *Physical Review X* **9**, 011001 (2019).
57. B. P. Abbott, *et al.*, *ApJ* **848**, L13 (2017).
58. B. P. Abbott, *et al.*, *ApJ* **848**, L12 (2017).

59. D. A. Coulter, *et al.*, *Science* **358**, 1556 (2017).
60. R. Margutti, *et al.*, *ApJ* **856**, L18 (2018).
61. R. Margutti, *et al.*, *ApJ* **848**, L20 (2017).
62. P. S. Cowperthwaite, *et al.*, *ApJ* **848**, L17 (2017).
63. P. A. Evans, *et al.*, *Science* **358**, 1565 (2017).
64. M. Soares-Santos, *et al.*, *ApJ* **848**, L16 (2017).
65. L. Lehner, *Class. Quant. Grav.* **18**, R25 (2001).
66. É. É. Flanagan, T. Hinderer, *Phys. Rev. D* **77**, 021502 (2008).
67. K. Takami, L. Rezzolla, L. Baiotti, *Phys. Rev. D* **91**, 064001 (2015).
68. B. Margalit, B. D. Metzger, *ApJ* **880**, L15 (2019).
69. A. Akmal, V. R. Pandharipande, D. G. Ravenhall, *Phys. Rev. C* **58**, 1804 (1998).
70. K. C. Gendreau, Z. Arzoumanian, T. Okajima, *Space Telescopes and Instrumentation 2012: Ultraviolet to Gamma Ray* (2012), vol. 8443 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 844313.
71. B. P. Abbott, *et al.*, *Phys. Rev. Lett.* **121**, 161101 (2018).
72. F. Özel, P. Freire, *ARA&A* **54**, 401 (2016).
73. K. Yagi, N. Yunes, *Phys. Rep.* **681**, 1 (2017).
74. D. Radice, A. Perego, F. Zappa, S. Bernuzzi, *ApJ* **852**, L29 (2018).
75. G. Raaijmakers, *et al.*, *ApJ* **893**, L21 (2020).

- 76. M. C. Miller, *et al.*, *ApJ* **887**, L24 (2019).
- 77. H. T. Cromartie, *et al.*, *Nature Astronomy* **4**, 72 (2020).
- 78. M. M. Kasliwal, *et al.*, *Science* **358**, 1559 (2017).
- 79. M. R. Drout, *et al.*, *Science* **358**, 1570 (2017).
- 80. B. P. Abbott, *et al.*, *ApJ* **892**, L3 (2020).
- 81. P. Lazarus, *et al.*, *ApJ* **831**, 150 (2016).
- 82. C. S. Ye, *et al.*, *ApJ* **888**, L10 (2020).
- 83. I. M. Romero-Shaw, N. Farrow, S. Stevenson, E. Thrane, X.-J. Zhu, *MNRAS* **496**, L64 (2020).
- 84. K. Kyutoku, *et al.*, *ApJ* **890**, L4 (2020).
- 85. R. J. Foley, *et al.*, *MNRAS* **494**, 190 (2020).
- 86. B. F. Schutz, *Nature* **323**, 310 (1986).
- 87. D. E. Holz, S. A. Hughes, *ApJ* **629**, 15 (2005).
- 88. L. Verde, T. Treu, A. G. Riess, *Nature Astronomy* **3**, 891 (2019).
- 89. K. C. Wong, *et al.*, *MNRAS* **498**, 1420 (2019).
- 90. B. P. Abbott, *et al.*, *Nature* **551**, 85 (2017).
- 91. H.-Y. Chen, M. Fishbach, D. E. Holz, *Nature* **562**, 545 (2018).
- 92. C. M. Will, *Living Reviews in Relativity* **17**, 4 (2014).

- 93. B. P. Abbott, *et al.*, Phys. Rev. D **100**, 104036 (2019).
- 94. R. Abbott, *et al.*, *arXiv:2010.14529* (2020).
- 95. S. L. Liebling, C. Palenzuela, *Living Reviews in Relativity* **15**, 6 (2012).
- 96. T. Akutsu, *et al.*, *Nature Astronomy* **3**, 35 (2019).
- 97. B. Iyer, *et al.*, *LIGO-India Technical Report No. LIGO-M1100296* (2011).
- 98. L. Barsotti, L. McCuller, M. Evans, P. Fritschel, <https://dcc.ligo.org/LIGO-T1800042/public> (2018).
- 99. Z. Arzoumanian, *et al.*, *arXiv:2009.04496* (2020).
- 100. A. S. Eddington, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **102**, 268 (1922).