5 Signal Extraction

The signals that come out of LIGO are noisy. This is due to both "fundamental" noise limits, like quantum and Brownian noise, but also because of residual seismic noise and more practical sources, for example, electronics. There is generally no easy way to see by looking at them if a gravitational wave is present or not. In order to pull signals from the noisy data, an extensive data analysis method is used. The analysis contains multiple steps and uses fairly simple statistical tests as well as sophisticated template fitting procedures; the LIGO experimental team includes experts in relativity and computation who create templates of the types of source that are expected to emit gravitational waves. These templates are compared to the experimental data, and under certain conditions, a match is declared. This is a somewhat simplified description of the process, but some of its details will be explored in the remainder of this section. The creation of templates, along with numerical techniques to compare them with data, are as essential to the success of LIGO as are the optical and mechanical structures that have been described previously.

5.1 Simulating the Astrophysical Sources

The first step in computing a template signal is computing the gravitational wave signature of different astrophysical sources. These waveforms are computed by various means depending on the astrophysical system being modeled. Some require solving the Einstein equation numerically. Some use analytic or semi-analytic models. The *Geometry and Gravity for Weak Fields* document in the first part of this course looked at some of the simple analytic solutions of the Einstein equation for the case of weak fields and no sources. As shown there, these solutions include waves. The sources of these waves can be in regions where the gravity is so strong and strongly time dependent that numerical solutions of the Einstein equation are the only way to describe them. The merging black hole sources detected so far are examples. However, other sources of gravitational waves are more quiescent. These would include merging neutron stars, for example, but none of these kinds of systems had been detected at the time this document was written.

As a reminder, the Einstein equation relates spacetime curvature to the mass/energy distribution.

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{5.1}$$

The left hand side of this equation describes the curvature of spacetime. It is generally written simply as $G_{\mu\nu}$, the Einstein tensor. However, here we have expanded it to emphasize its connection to spacetime curvature, which is described by the metric tensor, $g_{\mu\nu}$, the Ricci tensor $R_{\mu\nu}$, (a contraction of the Riemann tensor $R^{\alpha}{}_{\mu\beta\nu}$ on its first and third indices) and the Ricci scalar, R, which is a contraction of the Ricci tensor. The right hand side describes the distribution of mass/energy in spacetime via the stress-energy tensor, $T_{\mu\nu}$. G is the gravitational constant.

The Einstein tensor is the most general second-rank tensor related to spacetime curvature that shares the same symmetries as the stress-energy tensor. Thus, the Einstein equation is the simplest equation one can write that relates spacetime curvature, in other words, gravity, to matter and energy, the source of gravity. If this is not familiar, and if you are curious about the equation, see *Geometry and Gravity for Weak Fields* from the first LIGO course. The Einstein equation is explained there in more detail.

It is not obvious from Equation 5.1, but the Einstein equation is a non-linear, second order partial differential equation: Both $R_{\mu\nu}$ and $G_{\mu\nu}$ contain partial derivatives of the spacetime coordinates and their derivatives. What's more, gravity has energy and is therefore a source for itself - unlike the electromagnetic field. This makes solving the gravitational field equations much more difficult than solving Maxwell's equations for electromagnetism. The Einstein equation is really a set of ten coupled equations (in contrast to eight for the Maxwell equations), and it can only be solved analytically under very limited circumstances. The rapidly time-evolving, high-curvature systems that are strong emitters of gravitational waves do not fall into this small sample of analytic solutions.

5.2 Simulating the LIGO Signals

The first step in creating useful waveform templates is knowing what to simulate. Obviously, there could be unknown types of sources that we cannot possibly simulate - because we have not thought of them! The sources that have been considered break down into two basic types: short duration and continuous sources. These are exactly what they sound like they should be. Short duration events are one-off, or discrete events, though they could in principle repeat at some interval. Continuous sources produce gravitational waves from some mechanism that persists for some time. In the following discussion we briefly describe the general aspects of the astrophysical sources that LIGO might detect. A fuller discussion (with references for further reading) can be found in Abbott & et al. (2009) and in the document about astrophysical sources from the Summer 2015 LIGO course.

An example of a short-duration source would be the merger of two compact objects into a single object, most likely a black hole. This is exactly the type of event that was detected in both September and December, 2015. This type of source is probably the easiest to model. In the initial phase, there are simply two slowly in-spiraling objects that are emitting gravitational waves. In fact, if we had much more sensitive gravitational wave detectors, these would be considered continuous sources in the time before the last few seconds of the merger, not short-duration ones. Every orbiting system emits gravitational waves, but only the most massive objects, like neutron stars and black holes, emit them in directly detectable strengths, and even then, the emitted waves are only detectable in the moments just before the objects merge[11]. However, two binary neutron star systems are

[11]This is true for LIGO. Future space-based interferometers will have much longer baselines and much greater sensitivity. They will be able to detect much weaker sources at much lower frequencies than ground-based experiments like LIGO.

already known to be emitting gravitational waves as a result of their in-spiraling orbits.

The first of these in-spiraling systems to be discovered is called PSR 1913+16. It was discovered in 1974 by Russell Hulse and Joseph Taylor, and is often called simply the Husle-Taylor pulsar. The two researchers noticed an anomaly in the arrival times of the radio pulses from the source, which would arrive slightly earlier than anticipated for a while, and then arrive slightly later than anticipated. The pattern repeated with approximately an 8 hour period, an indication that the pulsar was part of a binary system. The companion is also a neutron star, but no pulsations have ever been detected from it. This could be because of its orientation, or perhaps because it does not emit pulses.

Pulsars are extremely precise clocks (more so even than atomic clocks), and so the orbital motion of PSR 1913+16 could be studied in great detail. After tracking the pulsar for years, its orbit was seen to change. The orbital period was decreasing steadily with time. In 1982, Hulse and his collaborator Joel Weisberg showed that the changes in the orbital period matched the prediction of general relativity for energy loss via gravitational radiation for a binary neutron star system. Figure 5.1 shows a plot of their data for PSR1913+16.

Continued observations of the system over the subsequent decades have borne out their interpretation, and in 1993 Hulse and Taylor were awarded the Nobel Prize in Physics for their discovery. Studies of this system, and another system like it, are properly considered to be the first confirmation, though indirect, of the existence of gravitational radiation.

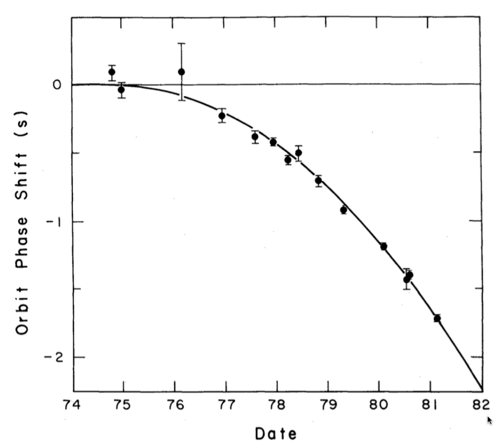

Figure 5.1: This plot shows the time evolution of the orbital period, plotted as a phase shift on the vertical axis, vs. the time, plotted in years on the horizontal axis. The curve shows the prediction from general relativity assuming loss of orbital energy due to the emission of gravitational radiation. From Taylor & Weisberg (1982)

Because the in-spiraling phase of a merger event is quite simple[12], it can be modeled semi-analytically. At some point the full description of the Schwarzschild metric has to be used, but that is still analytic.

The final part of the merger, when the objects coalesce, is a different matter. In the final coalescence, the gravity is so strong and highly time dependent that only numerical solutions of the Einstein equation can be used. However, it is still possible to specify the properties of the system in terms of the masses, spins and orientations of the two components. The solution can then be calculated to high precision, keeping in mind the finite computational time available. In practice, the way this is done is to calculate grids of templates that have different masses and spins. This is done in a manner that uses the limited computation time efficiently. These families of templates can then be compared to the data from the LIGO interferometers.

Another type of short-duration source that is not so easy to model is the collapse of a stellar core during a supernova. The collapse should produce a burst of gravitational waves, just as it produces a burst of other kinds of radiation. However, the details of the collapse, its symmetry or lack thereof, the size of the collapsing core, etc, will all affect the output pattern of gravitational waves. It is not as easy to imagine all the possible permutations for these systems, and so there is less certainty in the models that can be computed for them.

Continuous gravitational wave sources also fall into two basic kinds, one easy to model, the other not. The easy-to-model sources are asymmetric, rapidly spinning neutron stars. For instance, if a neutron star had a "mountain" on its surface, which would certainly be much smaller than the corresponding object on Earth, it would have a time-variable quadrupole mass moment, and would emit gravitational waves. Such an object is fairly easy to model. The possible differences between them can be described by the height of the mountain and its latitude. If additional mountains are desired, they can be added. The waves generated by these systems are described by simple formulae, and grids of templates can be produced, just as in the case of in-spiraling compact objects.

In contrast, the gravitational wave signal generated during the first moments of the big bang by Inflation is not easy to model. The spectrum and strength of these waves depend upon the details of Inflation, and these are, to say the least, not known. In fact, detection of such waves would be strong confirmation of the Inflationary theory, which stands now as not much more than an Ansatz, a working assumption, though one with some compelling explanatory power. A tentative announcement of primordial gravitational waves was made in March, 2014 by the BICEP2 experiment, which measures polarization of the Cosmic Microwave Background (CMB). However, the claim was withdrawn about a year later when better data from the Planck CMB mission were analyzed. The Planck data showed that the BICEP2 signal was likely caused by intervening cosmic dust in our

[12] It becomes much more complicated as the point of coalescence is approached, requiring sophisticated numerical treatment.

galaxy, not by gravitational waves. Thus far, no gravitational waves of this sort have been detected. While it is possible that LIGO and similar ground-based instruments will see them, it is far more likely that their detection, assuming they even exist, will have to await space-based interferometers with much longer baselines, or by experiments designed specifically to measure CMB polarization.

While template matching is a powerful way to extract a gravitational wave signal from the noise, it only works for sources that can be easily modeled. For other cases, for example, stellar core collapse, another method is needed. This is where having two observatories is absolutely essential. While having two antennas is helpful, and sometimes needed to verify sources found by template matching, it is not absolutely necessary; the gravitational waves from GW150914 were loud enough to be seen by either LIGO antenna individually, though seeing it in both increased the statistical significance of the detection. But by looking for excess noise in the signals from each detector, and then comparing arrival times, it is possible in principal to detect waves from a single short-duration source that is not well-modeled. The more instruments, the better this technique works. So with currently just the two LIGO detectors online, the method is somewhat hampered.

In coming years, the "excess-power" method for detecting gravitational waves will become much more powerful. This is because LIGO will soon have company. There is already a similar instrument to LIGO, called VIRGO, in Italy near Pisa. The development of VIRGO follows closely that of LIGO. It was commissioned in 2000, and its first phase construction was completed in 2003. It operated in test mode for a number of years, similar to the first-generation LIGO detector. It was then decommissioned in 2011 in order to be upgraded. The new version, called Advanced VIRGO, is set to begin operations in conjunction with LIGO in 2016. In fact, the LIGO and VIRGO research groups have pooled their resources to share technology and expertise since 2007, such that they are currently operating almost as a single experimental group. This makes sense given the advantages gained by using all three instruments together. In addition, another gravitational wave detector is currently under construction in Japan, at the Kamiokanda mine. Unlike LIGO and VIRGO, KAGRA, as it is called, is located underground. Otherwise it will be a similar L-shaped interferometer-based instrument. Both VIRGO and KAGRA have arms 3 km long. Still another gravitational wave detector is in the planning stages, though not yet under construction. This will be another version of LIGO, using an extra set of optics from the Hanford site. It will be built in India. KAGRA is set to begin operations in or around 2018, and LIGO-India will join the four other detectors sometime after that.

The "excess power" method is much enhanced when several detectors are employed. With it, the data streams from each observatory are searched for signals that are not easily accounted for by the noise characteristics of that particular instrument. When such interesting signals are found, corresponding signals, using an appropriate time window, are searched for in the data from the other observatories. Essentially, the signals from the different detectors are cross-correlated with each other. Since the noise in each experiment is uncorrelated with that of the others, a real signal should give a large spike in the correlation statistic. Complicating the analysis is the sensitivity of the detectors to the direction to the source, which can weaken the signal in one observatory relative to the others. But this effect can be taken into account and poses no serious problem. The method should be quite robust, even with just the two currently operating LIGO observatories. Thus far (in June, 2016), only one gravitational wave source (GW150914) has been strong enough to be detected using this method. As the additional observatories in Italy, Japan and India come online, the method will become much more powerful.

Another method to find gravitational waves is to look for signals that coincide with events that are visible using other means, and that should also emit gravitational waves. For example, core-collapse supernovae or long timescale gamma-ray bursts (GRB), which are caused by the collapse of a supermassive stellar core, are observed by telescopes sensitive to electromagnetic radiation. Both should produce gravitational waves. Likewise, short GRB events are thought to be the result of a neutron star - black hole merger, or a neutron star - neutron star merger. Soft gamma repeaters are another source that could emit gravitational waves. In these sources, a highly magnetized neutron star (a magnetar) suffers "star quakes" as its magnetic filed relaxes. The resulting rearrangement of the mass of the neutron star should produce gravitational waves that coincide with the gamma-ray emission. Combing through the LIGO data at times coincident with any of these events is one strategy for finding gravitational wave signals that might otherwise go unnoticed. As with the other non-template detection methods, no sources have yet been found using any external trigger. For additional information on the methods used to search for signals in LIGO data look at the Science Summary pages for GW151226 and GW150914 on the LIGO website.

5.3 Matching Templates to Signals: LIGO

The models used to compare to LIGO data are generally physically motivated. Einstein's Theory of General Relativity provides a framework for constructing such models in a very specific sense: The Einstein field equation. The models compared to the LIGO data are called *phenomenological models*, and they are fit to numerical simulations of systems created by solving the Einstein equation on supercomputers. These simulations are computationally quite expensive, and so it is not possible to create one for every system of interest. Instead, a smaller number are computed, and then an analytic model is created that links one simulation to another. A set of freely adjustable parameters is used with these models that allow them to match all of the available numerical simulations and to interpolate between them. There are several models that are used for this. Each uses a slightly different method, and so they produce slightly different waveforms. The implied properties of the modeled systems also differ as a result, but only in small ways. For example, in the case of GW150914 the differences between various model-derived values are within the uncertainties expected from the relevant models.

LIGO provides the data to compare to the numerically-derived models. However, it turns out that matching a template to the data is only the first (or sometimes the second) step. Things are more complicated than just template matching. The remainder of this section spells out some of the details. It is based upon a discussion with Kipp Cannon, a professor at the University of Tokyo. Professor Cannon contributes to one of the detection pipelines (automated analysis software) that are used to analyze the data from LIGO, and he provided comments about those processes to help us understand them better.

To understand why template matching alone isn't good enough, you have to consider the type of data that LIGO produces. It is a continuous stream of data full of all sorts of noise, as we have already discussed. These data are first checked for large departures from the expected noise profile. The nominal sampling rate of the data is 16.384 kHz, which produces 16,384 samples per second. These data are resampled at 2048 kHz, thus reducing the amount of data by a factor of 8, and, thus, also the time and expense of the related computations. The data are then compared to approximately 250,000 template waveforms. This analysis produces around 500 million samples of signal-to-noise ratio (SNR) *for every second of data collected* for each LIGO interferometer. So between them the two antennas produce a *billion* individual SNR data samples each second. Most of them are noise, and in such a large collection of noise it is very likely to get many spurious matches, just from random fluctuations in the data. As a result, the pipeline software has to do much more work.

The first step in the process is to apply a "SNR" filter[13], taking only the samples above some threshold value of the signal-to-noise ratio. In the most recent LIGO data reductions, the threshold was 4 SNR, or four SNR above the typical SNR value. Anything below that value was cut from further consideration. This was a way to very quickly reduce the number of data points that required further analysis. Figure 5.2 shows the three strongest candidates of the first observing run, along with the noise for that run. They all easily pass the threshold cut, as we would expect.

In the LIGO data, it turns out that if one sample has a high SNR, then it is usually surrounded by neighbors that also have high SNR. Most of these are false positives. So LIGO performs yet another cut, keeping only the single highest SNR candidate in each such cluster. This further reduces the size of the dataset. After that, the software looks for coincidence between the two LIGO antennas.

Each step of this process is progressively more difficult computationally. A SNR cut is quite simple, finding the highest SNR in a cluster is a bit harder, and looking for coincidence between the two machines is much harder still. However, at each step the amount of data is reduced, so despite the increasing complexity of the computations, the computers are able to perform each task. For example, in data from the most recent science run, the SNR threshold and clustering analysis reduced the number of data samples from ~ 500 million per antenna per second, to "just" 100 billion per interferometer for the entire run.

[13] This SNR is related to, but not the same as, the standard deviation from the mean, σ. You can sort of think of it as σ if you like, and you won't be too far off. We won't worry about its exact definition here.

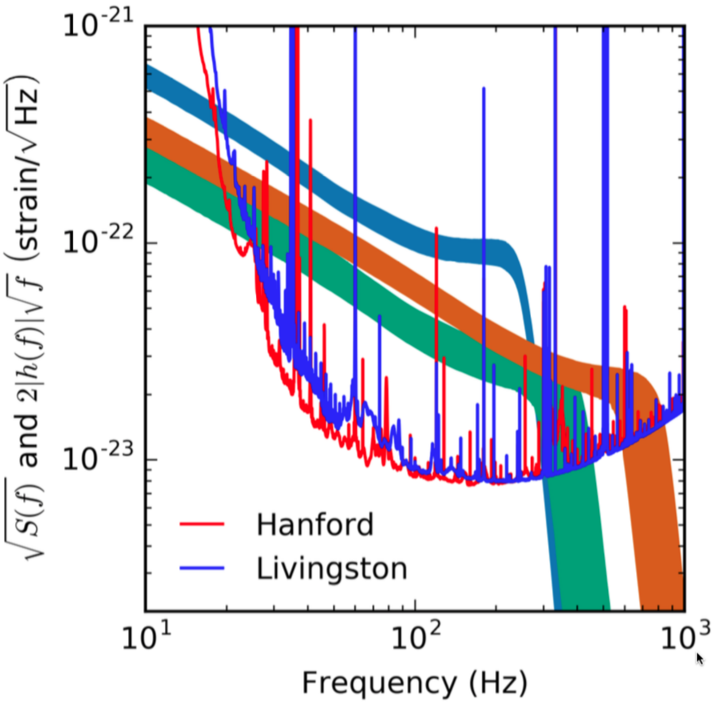

Figure 5.2: This plot shows the strain noise from both the Hanford and Livingston antennas vs. frequency. Also plotted are the waveform strengths for GW150914 (blue, top), GW151226 (orange, middle) and LVT151012 (green, bottom). The linear phase represented by the sloping line at low frequencies is the system in-spiral. The plateau is the merger event, and the abrupt drop is the ring down. The plot gives an idea of the SNR for each event. The finite width of each event indicates the spread in models that can describe it adequately. From LSCollaboration (2016)

Applying the coincidence check reduced the sample further. Ninety-nine percent of the data after the first two cuts were spurious, and only coincidence matching between the two antennas allowed them to be discriminated. Still, even after all these cuts and slices, there were a *billion* candidate gravitational wave sources. Again, almost all of these were noise. To reduce the sample still further, another test must be devised.

At this point, each candidate is described by a set of numbers. There is the SNR derived from each antenna, there is the template matched for the candidate, the χ^2 residuals from the template fitting, the sensitivity of each antenna at the time of the candidate given the noise spectrum in that antenna, and the time of each candidate in each antenna. From this is computed a statistic called the *log likelihood ratio*, or LLR[14]. The LLR is the natural log of the ratio of the probability that a true signal will produce a candidate with some exact set of properties to the probability that noise alone would produce a candidate with exactly that same set of properties. The larger the LLR for a candidate, the more likely it is to be caused by a signal rather than noise, and vice versa. The most important aspect of LLR is that it can be used to compare all candidates, regardless of the template they matched, the time of day they were observed or other individual differences. Ranking the candidates by LLR allows a determination to be made about which are likely caused by signals and which by noise.

When a real signal is present in the data, it is generally surrounded by a large number of candidates which are the result of matches by similar-shaped templates to the matching one. This has been determined by many runs in which simulated waveform data have been injected into the data stream to test the software. This allows candidates to be clustered, similar to the way the SNR threshold peaks were clustered; in this case, if a candidate falls within 4 seconds of another candidate that has a higher LLR, the weaker candidate is discarded. It turns out that there is a high probability of finding a low-ranked candidate next to a high ranked one, and so this method successfully trims many low-ranked candidates from the sample. However, for candidates with LLR larger than about 6, this clustering method is not effective because the probability of two such candidates being within 4 seconds of one another so low: the recent data run produced one candidate like this about every 5 minutes, or about $10,000$ of them for the entire run.

The total sample in the dataset has now been reduced from 500 million per second to only $10,000$ total, a small enough sample that it is possible to perform detailed statistical testing on each one of them. Most are probably (though not necessarily) still noise, but the only way to know is to look. This is done by modeling each candidate as a trial in the noise, assuming that each trial is statistically independent of the others. This allows the determination of how many of the candidates would be expected above some particular LLR value. Professor Cannon likens the process to flipping a coin repeatedly. Knowing the probability of each trial (coin flip in that case, GW candidate for LIGO) it is possible

[14] If you like, you can read the paper on how this statistic is computed. It is located at `http://arxiv.org/abs/1504.04632`

to say wether a certain number of "heads" in a row is too likely to be caused by chance, just as it is possible to determine (given the probability distribution of the LIGO noise) to determine if some particular LLR value is too large to be caused by chance.

The criterion used for LIGO data is that, if a candidate has an LLR value so high that it would require 1 million repetitions of the experiment to obtain the result just due to chance, then the signal is real. This is the standard 5σ rule use by scientists. A result is considered real if it is 5 standard deviations larger than the mean fluctuation, or in other words, if it has a one-in-a-million chance of happening due to random effects alone. To determine this probability, it is necessary to compute the distribution of likelihood of the number of events at or above some LLR as a function of LLR. That is what is shown by the blue line in Figure 5.3. As professor Cannon notes, a 5σ result is the likelihood that a coin could be flipped 20 times in a row and produce heads each time. That is one example that illustrates, in an intuitive way, the probabilities involved.

Figure 5.3 shows a plot of number of events expected vs. $\ln \mathcal{L}$ (what we refer to as LLR in the text). From the plot you can see that LVT151012 is seen at ~ 100 times the likelihood of being produced by mere chance. This is roughly 2σ, not significant enough to claim a detection. GW151226 is detected at $\sim 10^6$ times the likelihood of being purely chance. This barely meets the 5σ detection threshold. An interesting aspect of the analysis is that GW150914 is so strong that it skews the background. If it is not removed from the data, then GW151226 falls just below 3σ. However, removing GW150914 produces a more realistic background, and that brings GW51226 to just under 5σ according to the LLR cut. A complementary analysis pipeline to LLR finds GW151226 to be just over 5σ when GW150914 is removed. Together, the two analyses indicate that GW1512126 is barely detected at the 5σ significance level. LVT151012 is only 2σ in both analyses, regardless of whether GW150914 is removed or not. Even removing GW151226 has no effect on the statistical significance of LVT151012, but GW151226 is not a particularly strong detection, so that is not really surprising. Those wishing to know more about these three signals should read LSCollaboration (2016).

To summarize this process, a sample of about a billion data points *per second* for 30 days was cut to around 100 billion total through SNR threshold filtering and clustering cuts. This was then cut further, to about 1 billion, by looking for coincidence between the two detectors. Then, by cutting all log-likelihood-ratio candidates below $\text{LLR} = 6$, the total number was cut to $\sim 10,000$. Further statistical analysis found that only two of these met the 5σ threshold to be declared real, while a third was close (2σ), but not strong enough to be considered a true detection. In fact, Dr. Cannon points out, ~ 7 other candidates were likely due to signals, but their statistical significance was too low to be considered as such.

LIGO is a machine of many parts, and building the hardware to collect the measurements is only the first step to success. In addition, sophisticated computational modeling is needed, and the ability to match those models to the data. After that, careful statistical

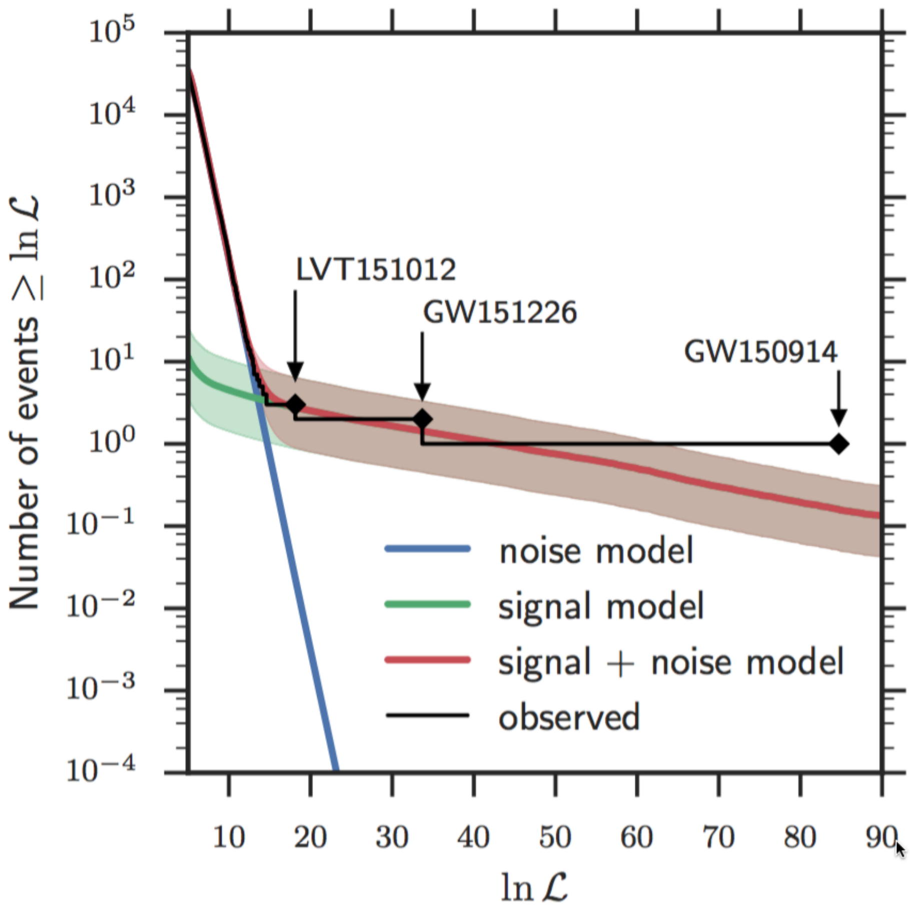

Figure 5.3: This plot shows the number of events expected from noise (blue) and signals (red) as a function of $\ln \mathcal{L}$. The three black squares are, from left to right, LVT151012, GW151226 and GW150914, respectively. Notice that the three candidates are all consistent with being signals, but that LVT151012 is only a factor of ~ 100 above the expected noise for that value of $\ln \mathcal{L}$. GW151226 barely meets the 10^{-6} detection threshold, and GW150914 is strongly detected - though the plot cuts off those values of the noise line. From LSCollaboration (2016)

analysis of these matches are required. Without any of these parts, it would not be possible to extract the signals of gravitational waves hidden in the data.

5.4 Additional Resources

The LIGO Open Science Center (LOSC) allows anyone to download and interact with LIGO data. We have collaborated with the creators of the LOSC site to build interactive activities for this class. They are linked on the Moodle site for this course at the bottom of Section Five. Go to the Moodle and run through the LOSC activities, which comprise the final graded homework problem for Section 5.

Acknowledgements

We would like to thank several members of the LIGO Science Collaboration for their helpful comments on this text. David Shoemaker of MIT and Marc Favata of Montclair State University both provided feedback on the general text. Kipp Cannon of Tokyo University gave us extensive comments about the data analysis process in Section 5. Andreas Freise, of the University of Birmingham in the UK, provided guidance on the use of the GW Optics website. This document was greatly improved by the input we received from these commenters, and any remaining errors or omissions are solely the responsibility of the authors.