Finding Gravitational Waves

Moritz Lange
moritz.lange@helsinki.fi
University of Helsinki
Helsinki, Finland

ABSTRACT

The first successful detection of gravitational waves in 2016 was a very important milestone for physics that even led to the Nobel prize for the involved researchers. This achievement would not have been possible without the use of methods from the field of data science to process the measured signals.

This paper will give an overview over the process of finding gravitational waves, as seen from the data science perspective. The main aspects treated here are the detection and validation of events within noisy data measured by the observatory. Those methods are not explained in great mathematical detail; instead this work aims at giving an intuitive understanding of the underlying ideas and puts them into the context of gravitational wave detection.

Many of the described methods are, however, general for signal detection and can be applied in other fields like natural sciences, engineering or finance as well.

KEYWORDS

gravitational waves, signal detection, matched filtering, LIGO

1 INTRODUCTION

The Nobel prize for physics in 2017 was awarded to three researchers and their teams who were the first to succeed in detecting gravitational waves. These waves had been predicted long ago already by Albert Einstein. The research team measured two clear signals with the LIGO observatory, each a wave through space time caused by the merging of two black holes somewhere in the universe. It has been an important milestone, since the detection of gravitational waves might lead to unforeseen insights into the structure of our universe, the very early beginnings of it, and perhaps even unexpected and unpredicted new physical phenomena.

Their achievement displays how the Nobel laureates overcame big challenges in the areas of physics and engineering. When looking closer, it becomes apparent that in addition to physics, this project also contained a good share of data science and statistics tasks related to signal detection and classification. This paper will explore those aspects of gravitational wave detection.

In section 2, gravitational waves will be described, as well as observatories built to detect them and what kind of data the LIGO observatory in particular generates. Section 3 will then look at how gravitational waves can be detected in noisy data streams based on the signal to noise ratio and matched filtering. Section 4, finally, will have a short look at how parameters for the astrophysical sources of those waves can be derived from the measurements.

2 MEASURING GRAVITATIONAL WAVES

Gravity is the weakest of the four fundamental forces known in physics. Much like electromagnetic radiation, energy transport in gravitational fields happens, according to Einstein's general theory of relativity, through waves propagating through space time. Over the past half century, different experiments have tried to measure these theoretically predicted gravitational waves (GW) by different means. One approach has been to use large masses, such as the Weber bar [7], which are supposed to exhibit a resonance when excited by a gravitational wave. This approach was so far unsuccessful, but some researchers are still experimenting with this method.

A recent and more promising approach are laser interferometers, which measure the change in distance between objects, in this case the mirrors of the interferometer, while the gravitational wave is passing through. A range of different observatories have been and are still being built, some on the ground, some up in space. LISA is a project run by the ESA, using an interferometer comprised of three satellites in a solar orbit, trailing the earth. On earth, amongst others, the Virgo detector has been built in Italy and the LIGO project was set up in the United States. LIGO was the first interferometer to detect gravitational waves in 2015.

2.1 Gravitational Waves

A gravitational wave is in principle generated by any accelerating mass. However, only GWs produced in the most energy-intense processes in our universe are detectable with the finest instruments.

These processes include primarily *binary systems*, i.e. two massive objects, like black holes or neutron stars, rotating around each other. Black hole mergers, where the black holes eventually merge into one, are the most commonly studied source of gravitational waves. Non-symmetric, i.e. deformed, rotating single neutron stars or supernovae also emit gravitational waves which are, due to the smaller energy of the process, weaker than those of binary mergers.

Finally, another source of gravitational waves might be processes in the very early universe. The waves they generate are considered *stochastic* because they are rather static signals generated by many events at once. Detecting these waves would allow insights into the earliest processes happening in the universe, back to 10^{-22} seconds after the big bang. This lies beyond the horizon of the cosmic microwave background radiation since GW propagation, unlike electromagnetic radiation, does not require transparency of the universe [2].

Since a GW is a wave, its basic form is a sine. However, the more specific form depends on the source. Stochastic GWs have a rather static, stationary waveform as opposed to binary mergers. With binary mergers, as the two objects spiral inwards with increasing

---

DSNS '19, Spring 2019, Helsinki, Finland

speed, the frequency and amplitude of the sine wave increase dramatically, only to sharply drop off when the objects merge and settle into an equilibrium [1]. Because of these characteristics, this form of GW is called a *chirp*.

The frequency range of binary merger GWs is around 10^{-4} Hz to 10^{3} Hz. This is in the range detectable by interferometer-based observatories [5]. GWs propagate at the speed of light.

2.2 LIGO

The LIGO interferometer, in its current form called advanced LIGO, consists of two Michelson interferometers located in Hanford, Washington, and Livingston, Louisiana. It is run by Caltech in collaboration with the MIT. Each of LIGO's two interferometers consists of two arms of 4km length each. In an interferometer a coherent laser beam is split and sent down both arms, then reflected by suspended mirrors at their ends. When the beams arrives at the detector, both are combined again to form an interference pattern, which will change if the mirrors in the arms have moved.

With this setup, movements of only 10^{-18} m in the mirrors can be measured despite the distance between the mirrors of about 5.66km [6]. The frequencies it can measure are in the range of 10^{1} Hz to 10^{4} Hz. LIGO is the most sensitive of these experiments in operation today [1]. It is naturally of great importance to isolate the devices from environmental influences as much as practicable, so that the GW signal is as clear as possible compared to the environmental noise.

The orientation of the two LIGO sites is not the same, one being rotated 90° with respect to the other one [2]. This enables detection of GWs with different polarisations.

2.3 Measurements

The two interferometers at LIGO measure phase shifts in the interference pattern of the laser over time, allowing to derive a strain in space. The data always contains noise, as well as possible gravitational events. Both are additive, leading to a total signal $d(t)$ of

$$d(t) = h(t) + n(t)\,,$$

where $h(t)$ is the gravitational wave and $n(t)$ is the noise. In the subsequent analysis the noise is assumed to be Gaussian, although this estimate is not entirely true due to deterministic terrestrial events such as seismic activity. A range of sensors at the LIGO sites measure environmental influences and data is discarded if terrestrial events influence the measurements [1]. Correlating the measurements of two interferometers which are sufficiently independent of each other helps to differentiate between noise and astrophysical signals. The terrestrial noise will not appear in both detectors, while a gravitational wave is detected with both devices.

3 EVENT DETECTION

A vital part in finding gravitational waves is the detection and extraction of events in the stream of noisy data. This is in essence a signal detection task in time series data.

How this task is approached depends on the form of the gravitational wave. There are several different forms of GWs which require different methodologies [2].

The most obvious GW signal is deterministic and stems from binary mergers. In this case the expected waveform is predicted by the general theory of relativity and known in advance, up to a number of parameters depending on the masses of the binaries, a phase and delay of the signal and its origin in space [1]. This allows for the use of a method called matched filtering, where the data is filtered for the expected signal. For spinning single neutron stars the waveforms can equally be determined in advance. This form of signal detection is performed offline.

However, the LIGO team also seeks to measure unpredicted events where the waveform is not known in advance. Methods used for those waveforms are often similar, but matched filtering cannot be applied out of the box here. Approaches here can differ and will not be further explained in this paper.

A third form of GW are the stochastic or background gravitational waves caused by processes in the early universe, or a potential multitude of unresolved sources distributed over large areas of the sky. Those do not allow for the use of matched filtering either, but the signal to noise ratio, a quantity indicating the statistical significance of a potential event, is used similarly to the case of binary mergers.

3.1 Signal to Noise Ratio

The *signal to noise ratio* is a quantity describing how large the GW signal $h(t)$ in a given time frame is compared to the noise $n(t)$. The SNR is the scalar product $\sqrt{(h|h)}$, where the scalar product is defined on the frequency domain as

$$(x|y) = \int_0^\infty \frac{x(f)\,y^*(f)}{S(f)}\mathrm{d}f\,.$$

Here $S(f)$ describes the spectral density of the noise $n(t)$ [4]. The intuition of this formula is that with a large GW signal the SNR will increase, while a large noise signal will lower the SNR. A higher SNR means a more dominant signal and hence a higher probability of detection of a signal. Building detectors less susceptible to noise leads to better SNRs during the experiments.

The SNR serves as an offset in matched filtering, to account for the probability of detection of a certain predicted waveform $h(t)$.

3.2 Matched Filtering

Matched filtering is the technique commonly used for finding gravitational wave forms in data streams. In matched filtering, a template $h(t)$ for a waveform, i.e. a prediction for a GW, is compared to the actual measured data. This is done using the log-likelihood ratio

$$\log(\Lambda) = \log\left(\frac{P(d|H_1)}{P(d|H_0)}\right) = -(d|h) + \frac{1}{2}(h|h)\,,$$

where $P(d|H_1)$ is the probability that there is a GW signal in the data, and $P(d|H_0)$ is the probability that the data only contains noise [4, 5]. H_0 is therefore a null-hypothesis and H_1 is the corresponding alternative hypothesis. The rightmost term provides a good intuition on how the matched filtering works. The term $(d|h)$ can be seen as the correlation between data d and template h, giving high values if both correlate and low values otherwise. The term $(h|h)$ on the other hand is the squared SNR, providing an offset. This means that the likelihood will be influenced by the SNR and therefore the probability of detecting template h.

Before matched filtering can be applied, templates for $h(t)$ need to be found for different parameter configurations. Hence this method depends on the availability of a theoretical model of the processes that emit GWs. In the case of black hole mergers, the complete expected form of $h(t)$ during the entire merging process was only discovered shortly before LIGO went online. Most templates that are used for matched filtering are those predicted by Einstein's general theory of relativity, although some templates for events contradicting this theory have been explored as well. However, no events were found that would be in disagreement with Einstein's theories.

Since many parameters in the model are continuous and not discrete, there cannot be a template for every possible merger process. To tackle this, the templates are positioned in the parameter space in such a way that each possible event is close enough to some template . This guarantees that the matched filtering search will find a corresponding template for each event.

4 EVENT CLASSIFICATION

For the classification of a detected event the first step is to evaluate the possibilities of terrestrial rather than astronomical origin. When a terrestrial source has been ruled out, posterior estimates of the actual parameters of the astronomical event are calculated.

At the very beginning of the process, the signal needs to be separated from the noise. So once an event is found, it is extracted from the data by subtracting the noise. In order to do this, the noise is estimated either based on previous measurements or on theoretical approximations, for example postulating Gaussian noise.

4.1 Validation

Not only natural processes could lead to terrestrial signals that get extracted from the noise, the researchers working at LIGO were also worried about deliberately faked signals of GWs. However, the possibility of faked signals is now considered negligible due to the complexity of the experiment and the exorbitant effort required to set up such a fake. Natural origins, on the other hand, are harder to tackle.

One approach is to estimate occurrence rates of terrestrial events that the filter would detect and those of GW events. By comparing these two estimates, a probability can be compiled for the detected event to be of astrophysical origin.

A more sophisticated approach that is usually applied is called the chi-squared veto [3]. Here the signal is binned into different frequency bands. Then the difference between signal and template over the bands will exhibit a χ^2 distribution. If signal and template do not match, non-central parameters are introduced which grow with increasing mismatch. For certain non-central parameters it can be deduced that the signal and template have a mismatch indicating a terrestrial rather than astronomical source.

4.2 Inference

If the event is believed to originate in space, the next step is to estimate the astrophysical parameters of the process that generated the GW. These parameters include the masses of the merging objects and phases and delay of the signal containing information about distance of the event and position in the sky.

Since there was a prior assumption, given by the template for the matched filtering, and now evidence in form of a signal has been found, Bayesian probability calculus applies for deriving the posterior parameters.

Most commonly *Markov Chain Monte Carlo (MCMC)* methods are used to efficiently sample from the posterior parameter space and finally derive a suitable model [6]. In MCMC, parameters for the process are guessed repeatedly and those guesses then rejected or accepted based on the probability of them being true given the measured event. In the end, the accepted guesses will follow the true posterior probability distribution for the parameters in question.

5 CONCLUSION

Detection of gravitational waves is a complex task involving many data science aspects. Challenges range from storing and distributing the data efficiently, to extracting events out of a noisy data stream, to verifying conclusions and inferring knowledge. Of large importance for the physics community is, furthermore, inference based on detected and validated events. Especially derived quantities are of importance. These include for example merger rates specifying how often black hole binaries merge, or sky maps showing where in the universe mergers take place. This paper, however, has mostly focused on explaining the extraction of signals from the measured data and their validation, while touching on how to infer parameters of the underlying astrophysical processes.

It is noteworthy that the methods described here are largely applicable to any signal detection task that includes noisy time series data. The finance industry would be an example for another domain, with its stock market or exchange rate data streams that contain information about certain market behaviours. Another example is radio transmission in electrical engineering, where the antennae receive a noisy signal in which the original message from the sender is hidden.

In all those contexts, the SNR is a useful measure. Whenever a specific pattern is searched in time series data, matched filtering can be applied. Finally, MCMC methods are generally a very powerful tool to estimate probability densities, even outside the field of signal detection.

REFERENCES

[1] B. P. Abbott et al. 2016. Binary Black Hole Mergers in the first Advanced LIGO Observing Run. *Phys. Rev.* X6, 4 (2016), 041015. https://doi.org/10.1103/PhysRevX.6.041015,10.1103/PhysRevX.8.039903 arXiv:gr-qc/1606.04856 [erratum: Phys. Rev.X8,no.3,039903(2018)].

[2] Bruce Allen. 1997. The stochastic gravity-wave background: sources and detection. In *Relativistic Gravitation and Gravitational Radiation, Proceedings of the Les Houches School of Physics, held in Les Houches, Haute Savoie*, Vol. 26. 373–418.

[3] Bruce Allen, Warren G. Anderson, Patrick R. Brady, Duncan A. Brown, and Jolien D. E. Creighton. 2012. FINDCHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries. *Phys. Rev. D* 85 (Jun 2012), 122006. Issue 12. https://doi.org/10.1103/PhysRevD.85.122006

[4] Piotr Jaranowski and Andrzej Królak. 2012. Gravitational-wave data analysis. Formalism and sample applications: the Gaussian case. *Living Reviews in Relativity* 15, 1 (2012), 4.

[5] Chris Pankow. 2016. The statistics of ground-based interferometric gravitational-wave detection and astrophysics. Presentation. https://dcc.ligo.org/public/0128/G1601914/002/PhyStatNu.pdf

[6] Chris Pankow, Will M. Farr, and Ben Farr. 2017. Statistical Methods of Gravitational-Wave Detection and Astrophysics. http://bulletin.imstat.org/2017/02/statistical-methods-of-gravitational-wave-detection-and-astrophysics/

[7] J. Weber. 1968. Gravitational-Wave-Detector Events. *Phys. Rev. Lett.* 20 (Jun 1968), 1307–1308. Issue 23. https://doi.org/10.1103/PhysRevLett.20.1307