Evidence of gravitational waves, or evidence of confirmation bias?

Michael Mozina (http://orcid.org/0000-0001-7306-4492)
President - Emerging Technologies - P.O. Box 1539 - Mount Shasta, CA 96067 - michael@etwebsite.com - (800) 729-4445

Abstract:

On February 11th 2016 the LIGO Scientific Collaboration announced the discovery of gravitational waves. Their announcement has been met with a great deal of excitement and enthusiasm by the scientific community. However, a careful and detailed analysis of the published papers, and other internal LIGO documents, reveal critical scientific methodology problems and unresolved questions surrounding the published materials which tend to undermine the veracity of the discovery claim and which could suggest a pattern of confirmation bias. For instance, the sigma calculation that is provided by LIGO does not provide any statistical assessment as to the likelihood that gravitational waves are the specific source of this signal. The published papers report that no data quality vetoes were active within an hour of the signal, but this account of events conflicts with other internal LIGO documents. The wave form, frequency range and duration of GW150914 are remarkably similar to common blip transient events which are routinely observed by both LIGO detectors. The exclusion of a blip transient as the probable cause of GW150914 is based upon a questionable assumption when considering the significant sensitivity upgrades that had just been completed prior to engineering run eight (ER8). The lack of any visual or neutrino confirmation of a celestial event at the time of the signal offers the LIGO Scientific Collaboration a viable elimination method related to their own gravitational wave claims which must be implemented in this case and all future claims of gravitational wave discovery to avoid any possibility of confirmation bias.

1. Introduction

All discoveries in physics require a high standard of evidence, and a very high degree of statistical confidence in the claim that is being made, typically in the form of a mathematical calculation. A 5+ sigma level of mathematical confidence is usually required before the announcement of a new discovery in physics. This quantified sigma figure is supposed to directly support the claim itself, and estimate the likelihood that the observation in question is directly related to the presumed source or cause of the observation rather than being caused by a random event. Claims of scientific discovery also require a methodical way to rule out all other, potentially more probable causes of the observation, and they require a clear method to verify or eliminate the claim that is being made. The information which is presented in the published papers and various materials must be complete and accurate for anyone outside of the original authors to be able to properly assess the accuracy of the claim. None of these standards of evidence would appear to have been met with respect to the claim that GW150914 was caused by a gravitational wave. [1]

2. Fuzzy Sigma

"The significance of the GW150914 event was measured to be greater than 5.1 σ, corresponding to a false-alarm rate of less than 1 event per 203000 years."[2,Page 10]

The sigma assessment method that is provided by the LIGO team does systematically eliminate ordinary statistical anomalies in the data set which might be produced by typical (non-vetoed) background noise, and their calculation does statistically eliminate the likelihood of false-alarms just as the LIGO team suggests. However, that sigma figure does not calculate the likelihood that this signal is specifically caused by gravitational waves rather than by any other potential cause of the signal.

The original test data that is used to create their sigma calculation only includes a data set period of 16 quiet days of non-vetoed times which were hand selected from a total of over 40 days of raw environmental noisy data.[2, Page 11] The time-shifting technique that is described in the sigma calculation method by LIGO therefore cannot be used all by itself to determine the likelihood that all potential environmental factors, and all possible raw environmental data from all 40 days, are incapable of producing a similar signal. There was no attempt made to eliminate all possible environmental effects as a potential cause of this signal based exclusively upon the time-shifting method which is described in their papers. Since vetoed times, and heavy environmental signal days were selectively removed from the test data, it is not possible to determine if all environmental influences, including noisy days, and all influences from all vetoed noise events, could reproduce a similar signal using the time-shift technique. That probability calculation was never tested by this specific sigma calculation method.

This effectively means that there is no actual sigma calculation included in this presentation which specifically determines the likelihood of this signal being directly related to gravitational waves, nor is there any sigma calculation that is provided by LIGO which can quantifiably exclude all other possible environmental influences from consideration. Any confidence figure that one might try to assign to the claim that GW150914 is likely to be related to gravitational waves, rather than from other potential environmental sources must therefore be based upon the premise that we can confidently rule out all other potential causes of this signal by some means other than the aforementioned sigma calculation method alone.

3. Elimination Of Known And Identified Environmental Influences

There is no direct visual or neutrino confirmation of this signal which might provide us with clear evidence that the candidate signal is likely to be celestial in origin rather than terrestrial or solar system in origin. The claim of gravitational wave discovery must therefore be based upon some non-mathematically quantified 'process of elimination' arguments and methods which are presented in the papers. We must be able to statistically rule out every other potential cause of this signal between the two LIGO detectors on Earth, and two merging black holes which were presumably located over a billion light years from Earth in a galaxy far, far away.

LIGO's methodical attempt to eliminate environmental influences is based upon the installation and use of a very comprehensive series of external auxiliary hardware channels, and data quality veto software methods, which are specifically designed to look for, identify and eliminate all known environmental background noises which might adversely influence the different LIGO detectors. Some 200,000 auxiliary hardware input channels have been designed and installed to check for a variety of common potential noise sources that are associated with each detector and their unique environments, including known sources such as earthquakes, maintenance and facility related noise, and electrical discharges from local thunderstorms. [2, Page 12]

The rationale that was used to ensure that GW150914 was not caused by any known environmental influence is based on the premise that most identified environmental factors would be likely to be picked up and recorded in one or more of the various auxiliary hardware input channels, and systematically vetoed by their software. Based upon this process of elimination methodology, The LIGO team presumes that they can confidently identify and systematically rule out many other potential environmental factors because they state that: "No data quality vetoes were active within an hour of the event."[2, Page 23]

However, that published account of events is inconsistent with other internal LIGO documents which report that this specific signal was originally vetoed and rejected by the external hardware and software veto methods which were in place at the time of the signal, with a level of "high confidence", within 18 seconds of the signal being uploaded to the (GraceDB) database. [3, Page 10] They furthermore state that this specific data quality veto remained in place throughout the first two and a half hours after the event until the data quality veto was eventually overridden by hand:

"LLO – September 14, 2015, 09:53:51 UTC – Alex Urban, Reed Essick:

The Coherent WaveBurst (cWB) data analysis algorithm detected GW150914. An entry was recorded in the central transient event database (GraceDB), triggering a slew of automated follow-up procedures. Within three seconds, asynchronous automated data quality (iDQ) glitch-detection follow up processes began reporting results. Fourteen seconds after cWB uploaded the candidate, iDQ processes at LLO reported with high confidence that the event was due to a glitch. The event was labeled as "rejected" 4 seconds afterward. Automated alerts ceased.

Processing continued, however. Within five minutes of detection, we knew there were no gamma-ray bursts reported near the time of the event. Within 15 minutes, the first sky map was available.

At 11:23:20 UTC, an analyst follow-up determined which auxiliary channels were associated with iDQ's decision. It became clear that these were un-calibrated versions of h(t) which had not been flagged as "unsafe" and were only added to the set of available low latency channels after the start of ER8. Based on the safety of the channels, the Data Quality Veto label was removed within 2.5 hours and analyses proceeded after restarting by hand. "

If GW150914 was originally vetoed automatically by the auxiliary channel hardware and software with high confidence as reported by LIGO magazine, then the presence of that high confidence data quality veto would tend to suggest that the signal was most likely caused by one or more of the various environmental factors which are being monitored by the external auxiliary hardware rather than caused by anything distant from Earth.

If a data quality veto of this specific signal did take place, a full and detailed explanation of the specific data quality veto which reported a high confidence rejection must be provided. It should identify the source code of the specific veto method that was involved, include a full explanation of the specific cause of the veto, and include an explanation as to the high confidence rejection level which was originally assigned to the veto. It should also identify and specify which of the auxiliary hardware channels and software veto routines triggered the veto, and it should include a complete and thorough explanation as to how and why the veto was later deemed "safe". All such critical information would be necessary to determine which specific environmental factors warrant further scrutiny by anyone external to the original LIGO authors. No specific information was given about the data quality veto in question. In fact, the published material specifically stated that no data quality veto took place within an hour after the event. It's not clear which of these conflicting accounts of events is correct.

If the signal in question originally failed the data quality veto test at the time of the signal, it would tend to call the entire environmental methodology into question, effectively rendering it moot. The LIGO team methodically filtered out all other data quality veto times and events from the time-shifted test set, and they claimed to have kept the equipment and presumably the veto software in the same configuration throughout the data gathering period as it was at the time of the signal. [2, Page 11] If they left everything exactly as it was at the time of the signal, and they followed the very same automated procedures with respect to this specific signal and the automated veto event, GW150914 would also have been systematically removed from the data set before starting the time-shift technique. There would be no non-vetoed signal to compare anything to in the first place, rendering the whole time-shifting technique irrelevant.

If the signal was originally vetoed with high confidence as reported in LIGO magazine, and that data quality veto was later changed from a high confidence rejection, to a high confidence "safe" category, exactly how safe was that change from a quantified mathematical perspective? Was it 80 percent safe, 95 percent safe, or 5.1 sigma safe? Since there is no quantified sigma figure associated with any such human intervention to be found in the published papers, it's unclear how anyone could externally corroborate or mathematically quantify that specific human decision.

There are simply too many unanswered questions surrounding the circumstances of the data quality veto event of this specific signal for anyone to independently verify the overall effectiveness of LIGO's attempt to eliminate known environmental sources as the cause of the signal.

4. Eliminating Blip Transients And Other Probable Non-Vetoed Environmental Causes

In addition to known and identified environmental influences which effect the LIGO detectors, LIGO reports that both detectors have routinely recorded a signal called a blip transient which shares remarkably similar frequency and duration characteristics with GW150914[2], and which also shares similarities with short burst electrical discharge events in the Earth's atmosphere.[4] Since blip transients have typically been observed in just one detector at a time, these events are presumed to be environmental in origin, but they cannot be traced to any specific auxiliary hardware input channels, and no effective veto method exists to automatically filter out blip transient events. [2, pages 14-15]

It is important to note that any limitation related to blip transients being observed by a single detector at a time was only necessarily true prior to the most recent upgrades which were being tested during engineering run eight (ER8). The most recent round of enhancements to LIGO have significantly improved the sensitivity of both detectors:

"The lessons learned during Initial LIGO's operation led to a complete redesign of LIGO's instruments, which were subsequently rebuilt between 2010 and 2014. This redesign made LIGO's new interferometers 10 times more sensitive. A 10-fold increase in sensitivity means that the new and improved LIGO will ultimately be able to listen for gravitational waves 10 times farther away than Initial LIGO. This is an enormous improvement since listening 10 times farther away will give LIGO access to 1000 times more *volume* of space (volume increases with the cube of the distance)." [5]

LIGO's rationale for eliminating blip transients as the probable cause of this specific signal is based in part on the fact that LIGO had not previously seen blip transients in both detectors simultaneously. However, in ER8 they are using significantly upgraded and enhanced equipment for the very first time. During ER8 they are testing and calibrating their upgraded detectors and presumably using ER8 to verify that blip transient events are never detected by both detectors simultaneously (speed of light). Since LIGO had not yet completed the full engineering run associated with these hardware sensitivity improvements, which are 10 times more sensitive in terms of distance and 1000 times more sensitive in terms of volume space, it is difficult to understand how the LIGO team had yet determined if their new and improved detectors are still incapable of observing any blip transients in both detectors simultaneously. How exactly did they make that determination even before the end of the engineering run, and without identifying the cause of blip transients? This highly subjective choice could be interpreted as an example of confirmation bias.

It seems just as possible, if not more probable, that the recent hardware upgrades to LIGO simply provide the improved detectors with far better sensitivity to blip transient events. As they enhanced the sensitivity of each detector by 1000 times in terms of volume space, they may have significantly enhanced their ability to observe at least some blip transient events in both detectors simultaneously. If blip transients are simply long distance, high voltage electrical discharge events in the Earth's atmosphere, or unusual power grid related phenomenon which may have only been observable a few hundred miles from previous (less sensitive) detectors, they might now be observable over thousands of miles in the upgraded equipment. It's certainly conceivable that the improved sensitivity of the detectors might now provide some overlap to blip transient sensitivity between each detector and thereby enable both detectors to occasionally pick up the same blip transient signal. This scenario cannot be ruled out, particularly since the engineering run wasn't even complete yet to even make such a determination, and the actual cause of blip transients remains unknown.[2] GW150914 is located within the typical blip transient frequency range and duration pattern. The LIGO papers do in fact cite blip transient events as being one the most common types of false-alarm trigger signals. They also tend to look like gravitational wave signals. [2 – Page 20] In fact, LIGO cites an example of one blip transient event which most closely fits a mathematical model of a neutron star/ black hole merger pattern. [2 - Figure 12]

5. Confidence from black hole calculations

In the final analysis, the cosmological mathematical match to a specific LIGO signal amounts to a curve fitting exercise with an extremely high degree of freedom in the variables, such as the distance from Earth, the mass of the two objects, their spin rates, their merging speed, the merging angle with respect to Earth, etc. Because of the relative "freedom" of the variables, such mathematical models can be modified to fit a large variety of different signal patterns and waveforms, which might involve anything from the merger pattern of two neutron stars, to the merger signal of two black holes. Such mathematical models therefore provide some evidence and some high level of confidence that gravitational waves are a possible explanation for the signal in question, but such models alone do not offer us much confidence in determining the probability of the signal being directly related to gravitational waves.

This inherent limitation related to describing any probability calculations from mathematical merger models would not necessarily apply were LIGO to visually identify a celestial origin of this or any similar signal at the same time of the event. If the LIGO Scientific Collaboration had been able to visually confirm that an observed celestial event could be triangulated and correlated back to this specific LIGO signal, it's hypothetically possible to mathematically calculate the odds that a specific potential gravitational wave signal might be related to a specific celestial event, rather than from a random series of events. For instance, we might be able to use redshift measurements from a celestial event to calculate the approximate distance to the event, and compare that distance figure to the distance figures and triangulation methods which were estimated by the various LIGO merger models.[6] With three or more LIGO detectors online, we might be able to triangulate the candidate signal to a very small region of the sky, and show the mathematical correlation to a specific celestial event in terms of the exact triangulated location and the approximate distance measurements associated with that celestial observation. In such a scenario, it may indeed be possible to calculate the odds that a specific signal which is observed by LIGO is mathematically related to a gravitational wave from a specific celestial event rather than from any other potential environmental cause. Without any visual confirmation of a celestial source however, and with no additional information to work with, there is no logical or reliable way to calculate the likelihood of any given signal observed by LIGO as being specifically caused by a gravitational wave rather than caused by any other possible environmental source.

No visual or neutrino confirmation was ever found in spite of the event presumably releasing the energy equivalent of three full solar masses in gravity waves alone, in the span of about a quarter of a second.[6] A charged black hole merger might conceivably be expected to release more energy in the form of EM radiation than from gravitational waves during the same merger.[7] Other types of black hole mergers might produce significantly less EM radiation of course, but without any visual confirmation of a celestial origin of the signal, such mathematical models cannot speak to the probability that any given signal was related to gravitational waves. Without a visual confirmation to work with, these mathematical merger models can only demonstrate the possibility that a candidate signal could be related to gravitational waves but they cannot address the probability issue.

In the absence of any visual confirmation, and without any identification method or elimination method whatsoever being applied to an infinite number of potential gravitational wave claims, the discovery claim by LIGO appears to be a classic example of confirmation bias.

6. Confirmation Bias

Throughout their methodology, the LIGO team systematically identified, checked for, and then eliminated various potential environmental sources of the signal by introducing some type of external auxiliary information, or external resource which was then used to either verify or systematically eliminate the potential cause in question. For instance, if there was no auxiliary hardware channel input associated with ground motion and earthquakes during the time of the signal, they eliminated earthquakes from further consideration. If there was no significant external auxiliary hardware input associated with magnetic field changes during the time of the event, those potential “causes” of the signal were also eliminated from consideration. If they checked other external environmental resources, and they found nothing to confirm the idea in those external resources, they used that lack of external confirmation to systematically rule out various options as a possible cause of the signal.[2]

In the case of their gravitational wave claim, the external auxiliary hardware is represented by a plethora of different telescopes on the ground and in space which are certainly capable of observing high energy celestial events. However, in the case of gravitational wave claims, this external auxiliary information was not ever used to either verify or to eliminate the celestial origin claim in the very same way and method that external resources had been used to eliminate every other potential cause of the candidate signal.

The gravitational wave claim was uniquely exempted from the process of elimination method entirely, and no category for “unknown cause” was even allowed for in their categorization method.

The external hardware check, and the subsequent elimination process was applied to every other potential cause of the signal, but in the specific case of the gravitational wave claim, there is no external resource that is used as an elimination process for that one specific claim. Only the gravitational wave claim is specifically exempted from elimination based on a complete lack of external corroboration. If the LIGO team had followed their own methodology consistently, the external auxiliary input from various telescopes on the ground and in space, should have been used to eliminate the gravitational wave option from consideration, just as external crosschecks and resources had been used to eliminate other types of potential causes of the signal. All other potential explanations for this signal were eliminated based on a lack of external corroboration. In the specific case of gravitational wave claims, the LIGO team ignored the external input of their telescope network altogether. They assumed instead that the signal must still be caused by a gravitational wave in the absence of any external support, and in spite of their very methodical attempt to verify a celestial origin of the signal. That is a textbook example of confirmation bias.

7. Discussion

There is a completely different standard of evidence and a completely different process of elimination method that is being applied to gravitational wave claims than is being applied to every other potential environmental factor as a proposed potential cause of the same LIGO signal. There appears to have been a very concerted effort made to visually confirm a celestial origin of this signal, including attempts to locate signs of a celestial merger using the ICE CUBE neutrino detector.[8] All that monumental effort went to waste. That lack of external confirmation can and should have been used as a method to eliminate gravitational wave claims from further consideration just as various external resources had been used to eliminate other potential environmental causes. The GW150914 signal should have ultimately ended up in the unknown cause category, rather than the gravitational wave discovery category had the LIGO team followed their own process of elimination methodology consistently. The simple fact that there is no provision for a category of “unknown cause” casts a highly suspicious light on this claim.

Even the order of the ‘process of elimination’ method that was used by LIGO plays an important role in their bias. Had the process of elimination been applied to the gravitational wave claim first, before checking any other potential causes of the signal, gravitational waves claims would have been eliminated first based on a lack of any additional hardware confirmation. The remaining potential environmental causes would also have been systematically eliminated due to a lack of external confirmation, and we would end up with a signal without any identified cause. By choosing instead to eliminate all other environmental variables first, and then changing the methodology at the very end of the check list, it appears as though the signal is related to gravitational waves, but only because the process of elimination method is not being applied to the very last item to be checked, and no unknown cause category is possible. If we changed the order of elimination, and made earthquakes the last item to be checked and eliminated, and we removed the need for external corroboration for earthquakes, we would end up with a positive confirmation for earthquakes instead of for gravitational waves. If we made electromagnetic interference the last option to be checked and we removed the need for external corroboration for electromagnetic interference, electromagnetic interference would have been the default cause as well.

For these reasons, LIGO's claim of the discovery of gravitational waves appears to have all the earmarks of confirmation bias.

The process of elimination method that is described in the published papers automatically favors the gravitational wave claim by making it the default option throughout the process of elimination method. The method that they used is biased against every other possible local environmental explanation as to the potential cause of the signal, and it is heavily biased in favor of all gravitational wave claims by its haphazard and inconsistent use of external process of elimination methods. This serious confirmation bias problem is ultimately the Achilles heel of all claims of the discovery of gravitational waves which are devoid of visual or neutrino confirmation, and which provide no process of elimination mechanism related to the gravitational wave claim. This is a fundamental flaw in the scientific methodology that is described in the paper which is very disturbing.

It is necessary in science to take a good hard skeptical look at all claims that are being made, including, and most especially any claims involving a scientific discovery of significant magnitude. The LIGO Scientific Collaboration's sigma calculations do not technically even address or calculate the likelihood of this signal being directly caused by a gravitational wave. The conflicting testimonies from LIGO in terms of the data quality veto events surrounding this exact signal also needs to be fully explained by the LIGO team, including the source code of the veto method and a full explanation as to the "high confidence' figure that was originally assigned to the veto before any logical assessment of the potential influence of environmental factors by anyone other than the LIGO team can occur. Because that specific veto information was not provided in the published papers, a full external (to LIGO) assessment of potential environmental causes is not currently feasible. LIGO's subjective choice to eliminate blip transients as a potential cause of this signal is a very questionable and subjective choice due to the sensitivity enhancements which had just been implemented, and the fact that LIGO had not even completed their engineering run yet to determine if any blip transients could or might be observed in both detectors simultaneously after the upgrades. The real deal breaker of the claim of the discovery of gravitational waves is fundamentally due to the lack of any visual confirmation of this signal or any future signals, and the lack of any process of elimination method being applied to the gravitational wave claim as was applied to all other potential claims related to environmental factors. Unfortunately, due to the numerous aforementioned problems in methodology, and unresolved questions surrounding the veto events of September 14th 2015, the LIGO Scientific Collaboration simply did not meet the standard of evidence that is necessary to claim the discovery of gravitational waves.

With the addition of more LIGO detectors in the future, and continued improvements in hardware, LIGO will be better able to triangulate potential candidate signals to a very small region of space. Such improvements may enable us to then confirm or eliminate a celestial origin of a specific candidate signal. With hard work, and a little luck from a visual confirmation of a celestial event, we may be able to say with a high level of mathematical confidence that a specific observed celestial event can be mathematically correlated to a specific LIGO candidate signal. However, without any type of visual confirmation, there is fundamentally no possible way to calculate the probability of any given LIGO signal being directly related to gravitational waves, and therefore the lack of a visual or external confirmation of a cosmological event should be used to eliminate the gravitational wave claim to eliminate any potential for confirmation bias.

8. References

1. Abbott, B.P. et al. "Observation Of Gravitational Waves From A Binary Black Hole Merger". *Physical Review Letters* 116.6 (2016): doi:10.1103/PhysRevLett.116.061102.
http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.116.061102
https://arxiv.org/abs/1602.03837

2. Abbott, B P et al. "Characterization Of Transient Noise In Advanced LIGO Relevant To Gravitational Wave Signal GW150914". *Classical and Quantum Gravity* 33.13 (2016): 134001. doi:10.1088/0264-9381/33/13/134001
http://iopscience.iop.org/article/10.1088/0264-9381/33/13/134001/meta
https://arxiv.org/abs/1602.03844

3. "LSC - LIGO Scientific Collaboration - Magazine". *Ligo.org* - Issue 8 - March/2016
(Page 10 – Timeline Of GW150914)
http://www.ligo.org/magazine/
http://ligo.elte.hu/magazine/LIGO-magazine-issue-8.pdf

4. Kim, Hyomin et al. "Narrow-Band Extremely Low Frequency (ELF) Wave Phenomena Observed At South Pole Station". *Geophysical Research Letters* 33.6 (2006) doi:10.1029/2005GL023638
http://onlinelibrary.wiley.com/doi/10.1029/2005GL023638/full

5. LIGO Caltech
https://www.ligo.caltech.edu/page/facts

6. Abbott, B. P. et al. "Properties Of The Binary Black Hole Merger GW150914". *Physical Review Letters* 116.24 (2016) doi:10.1103/PhysRevLett.116.241102
http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.116.241102
https://arxiv.org/abs/1602.03840v2

7. Zilhão, Miguel et al. "Collisions Of Oppositely Charged Black Holes". *Physical Review D* 89.4 (2014): doi: 10.1103/PhysRevD.89.044008
http://journals.aps.org/prd/abstract/10.1103/PhysRevD.89.044008

8. Adrián-Martínez, S. et al. "High-Energy Neutrino Follow-Up Search Of Gravitational Wave Event GW150914 With ANTARES And Icecube". *Physical Review D* 93.12 (2016): doi:10.1103/PhysRevD.93.122010
http://journals.aps.org/prd/abstract/10.1103/PhysRevD.93.122010
https://arxiv.org/abs/1602.05411