Prepared for submission to JCAP

Investigating the noise residuals around the gravitational wave event GW150914

Alex B. Nielsen, [a,b,1] **Alexander H. Nitz,** [a,b] **Collin D. Capano** [a,b] **Duncan A. Brown** [c]

[a] Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany
[b] Leibniz Universität Hannover, D-30167 Hannover, Germany
[c] Department of Physics, Syracuse University, Syracuse NY 13244, USA

E-mail: alex.nielsen@aei.mpg.de, alex.nitz@aei.mpg.de, collin.capano@aei.mpg.de, dabrown@syr.edu

Abstract. We use the Pearson cross-correlation statistic proposed by Liu and Jackson [1], and employed by Creswell et al. [2], to look for statistically significant correlations between the LIGO Hanford and Livingston detectors at the time of the binary black hole merger GW150914. We compute this statistic for the calibrated strain data released by LIGO, using both the residuals provided by LIGO and using our own subtraction of a maximum-likelihood waveform that is constructed to model binary black hole mergers in general relativity. To assign a significance to the values obtained, we calculate the cross-correlation of both simulated Gaussian noise and data from the LIGO detectors at times during which no detection of gravitational waves has been claimed. We find that after subtracting the maximum likelihood waveform there are no statistically significant correlations between the residuals of the two detectors at the time of GW150914.

[1] Corresponding author.

Contents

1 Introduction **1**

2 Residuals after subtracting a maximum-likelihood waveform and their statistical significance **3**

3 Robustness of the analysis **7**

4 Whitening vs bandpassing the data **8**

5 Conclusions **10**

1 Introduction

The direct detection of gravitational waves from the binary black hole merger GW150914 [3] was a groundbreaking discovery in physics and astronomy. An event of this importance warrants broad community scrutiny using a variety of different techniques. To that end, the LIGO Scientific and Virgo Collaborations (LVC) have released a number of data products related to GW150914 through the Gravitational Wave Open Science Center (GWOSC) [4, 5]. This includes 4096 seconds of calibrated strain data [6–8] centered on GW150914, along with the figure data of Ref. [3].

One of the groups to take on the challenge of independently verifying the LVC results has raised some concerns about the reliability of the detection [2]. A key feature of the concerns of Ref. [2] is the possibility of significant excess correlations (after the signal is subtracted) between the data streams of the two LIGO detectors, one of which is in Hanford, WA and the other in Livingston, LA. Since independence of the background noise between the two LIGO detectors is a key assumption in assigning statistical significance to GW150914, the concerns of Ref. [2] have received considerable attention. Here, we use the methods of Ref. [2] to look for correlations in the LIGO data after subtracting the GW150914 signal. We show that there are no statistically significant correlations that might call into question the claim of detection of GW150914.

The analysis of Ref. [2] is fundamentally different from the analyses performed by the LVC to produce their scientific results. The LVC analyses that assign statistical significance to GW150914 are discussed in Ref. [3] and further detailed in a number of companion [9–11] and methodological papers [12–17] (and references therein). These analyses include both unmodeled and modeled searches. The unmodeled searches make minimal assumptions about the types of signals to be searched for [9, 14]. The modeled searches use binary black hole template waveforms that are based on general relativity (GR) to filter the detector data [10, 12, 13, 15–17]. Final results of both of these techniques rely on careful monitoring of the detectors' performances to exclude data during known terrestrial disturbances [10]. The modeled searches bounded the significance of GW150914 to higher values than the unmodeled searches [3]. This is expected: using templates based on general relativity helps to exclude unknown instrumental and terrestrial noise in the detector data. A check for any residual signal after subtraction of a GR-based template was performed by the LVC in Ref. [18]. That analysis found that the Gaussian-noise hypothesis was preferred over any coherent residual signal between the two detectors, suggesting that all the measured power is well represented by the GR prediction for the signal from a binary black-hole merger.

The independent analysis of Ref. [2] is based on calculating the Pearson correlation coefficient [1]

$$C(\tau;t,\omega)=\int_{t}^{t+\omega}\frac{H(t'+\tau)}{\sigma_H}\frac{L(t')}{\sigma_L}\mathrm{d}t' \tag{1.1}$$

between the data streams of the two LIGO detectors, $H(t)$ and $L(t)$. Here, $\sigma_{H,L}$ are the standard deviations of the Hanford and Livingston data, respectively; t gives the time stamp of the data to analyze, while ω is a window specifying how much data to use in the correlation. As in Ref. [2], we choose ω to be 20 ms. Some of the results of Ref. [2] made use of the source data for Fig. 1 of Ref. [3], which is available for download from GWOSC. This figure consists of three elements. The uppermost row shows the strain data of the two LIGO detectors around the time of GW150914. A bandpass filter was applied to this data to remove high (above 350 Hz) and low (below 35 Hz) frequencies, where the detectors are less sensitive. A notch filter was also applied to remove a number of narrow frequencies. The second row of Fig. 1 of Ref. [3] shows credible regions for template-based and unmodeled wavelet-based reconstructions of the signal along with an example waveform simulated by numerical relativity. The timing, amplitude, and phase of the numerical waveform was tuned to match the data by visual inspection for illustrative purposes. The third row of the figure shows the result of subtracting the numerical relativity waveform of the second row from the detector strain data of the first row to produce so-called residuals, which we will refer to as the "150914 PRL Residuals".

It is straightforward to compute the Pearson correlation coefficient using the 150914 PRL Residuals. We use the PyCBC analysis toolkit [19] to reproduce this result here in our Fig. 1; this is qualitatively similar to the result of Ref. [2] (their Fig. 8). A particular concern of Ref. [2] was that the correlation coefficient of the 150914 PRL Residuals is peaked at a time-shift between the Hanford and Livingston data streams that is very close to the claimed delay time of $6.9^{+0.5}_{-0.4}$ ms for GW150914 [3]. A region encompassing the timing uncertainty is shown in the plot as a shaded band at 7 ± 0.5 ms in Fig. 1.

The authors of Ref. [2] claim that this is concerning enough to call into question the validity of the detection of GW150914. Although reanalyzing this result is the main topic of this work, here we point out that a correlation of residuals was not the method employed by the LVC to make the significance claim for GW150914 in Ref. [3]. The LVC significance claims for GW150914 are based both on matched-filter searches – which use GR model waveforms to match against the data – and unmodeled searches that look for excess coherent power between the two detectors. Both of these searches empirically measure the probability of obtaining chance coincidence between the detectors by time-shifting the detector data millions of times and reanalyzing. Neither of these methods assumes that the underlying noise is Gaussian distributed, as the data is known not to be Gaussian or stationary over long periods of time [11]. Times where one of the detectors was known to be operating incorrectly were removed from the analysis ahead of time, as described further in Ref. [11]. From this background analysis, both the unmodeled and modeled searches obtained a high statistical significance for GW150914.

Our analysis here differs in several ways from Ref. [2]. We use data corresponding to an updated calibration model released by the LVC in October 2016, labeled as "v2" at the GWOSC [5] and "C02" by the LVC, which is slightly different from that plotted in Fig. 1 of Ref. [3]. We use a different waveform to produce the residual data. Instead of using

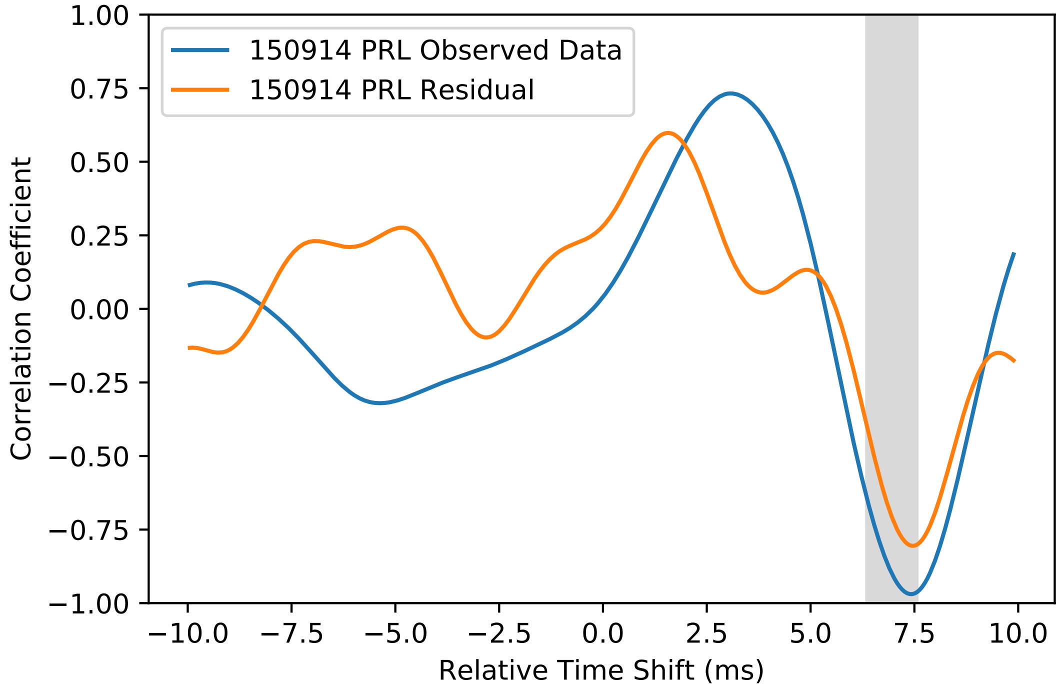

Figure 1. Correlations between timeseries data of the Hanford and Livingston detectors that are displayed in Fig. 1 of Ref. [3]. The correlations shown are for both the data containing the GW150914 signal (in blue) and the example residuals produced after subtracting a numerical relativity GR waveform (in orange). As found in Ref. [20] and Ref. [2], both the data containing the GW150914 signal and the example residuals show a significant correlation at a time shift corresponding to the time delay between the signal arriving at the Livingston detector and then arriving at the Hanford detector. The grey shaded region shows a region from 6.5 ms to 7.5 ms that contains the time shift between the Hanford and Livingston detectors for the signal observed by LIGO.

the numerical waveform of [4], we use a maximum likelihood (ML) model waveform from Ref. [21, 22]. We estimate the significance of our result using a distribution of uncorrelated Gaussian noise samples and samples from the real detector noise in time segments that do not include the GW150914 event. We also investigate the effect of using whitening rather than bandpass and notching on the correlations. Whitening of the data more closely follows the methodology used by the LVC to estimate the significance of GW events.

2 Residuals after subtracting a maximum-likelihood waveform and their statistical significance

The waveform used to produce the residuals of Fig. 1 of Ref. [3] was a numerical relativity simulation which approximately fits the data [23]. Numerical binary simulations take a long time to run (typically several weeks on a dedicated computer cluster). They cannot yet be used for a detailed exploration of the likelihood surface for the parameters involved [24], such as the individual masses and spins of the coalescing black holes. Instead, model waveforms that

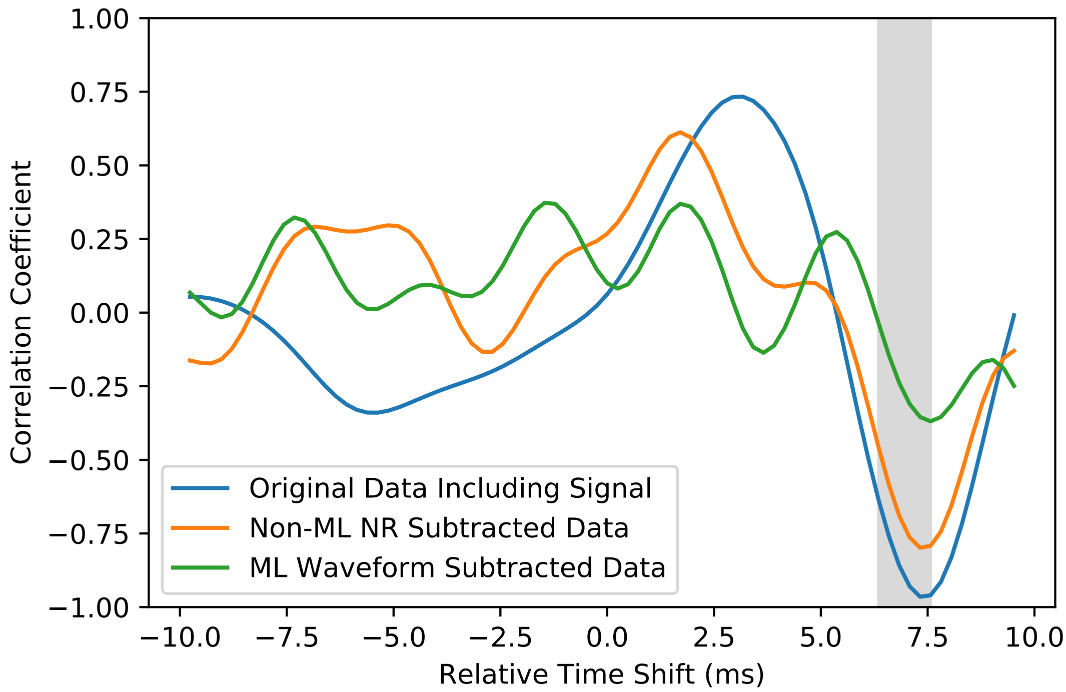

Figure 2. Correlations between the Hanford and Livingston data released at Ref. [5] that has been bandpassed and notched in the same way as used to produce Fig. 1 of Ref. [3]. This data has an updated calibration compared to data used in Ref. [3]. Therefore the orange and blue curves are not exactly identical (although very similar) to those plotted in Fig. 1. The figure also contains the correlations for residuals constructed by subtracting the maximum likelihood model waveform found in Ref. [21, 22]. The correlations for these residuals are seen to be significantly lower than for the residuals constructed from the example numerical waveform from Fig. 1 of Ref. [3].

combine both analytical and numerical relativity information are used which are considerably faster to generate. These model waveforms have been shown to agree well with numerical relativity simulations of GW150914 [24]. The LVC used waveform models to estimate the source parameters of GW150914 [25]. Parameters of a maximum likelihood (ML) model waveform for GW150914 from such a parameter exploration were released with Ref. [21, 22]. These values apply to the IMRPhenomPv2 family of effective precessing spin waveforms [26] that is freely available through the LALSuite package [27]. Although the result of Ref. [21] is not identical to those of the LVC [25] (Ref. [21] did not include a marginalisation over detector calibration uncertainty), the differences are sufficiently small for our purposes. The parameters of this waveform lie well within the confidence intervals given in Ref. [25]. Using this waveform, suitably projected onto the two detectors for their given location and orientation, we produce the residuals released as a supplement to this work [28]. Due to the difference in data (we use data with an updated calibration from [5]) and the waveform subtracted (we use the maximum likelihood waveform from [21, 22]), our residuals are not identical to the 150914 PRL Residuals.

Although the waveform we subtract was selected by maximising a coherent Gaussian

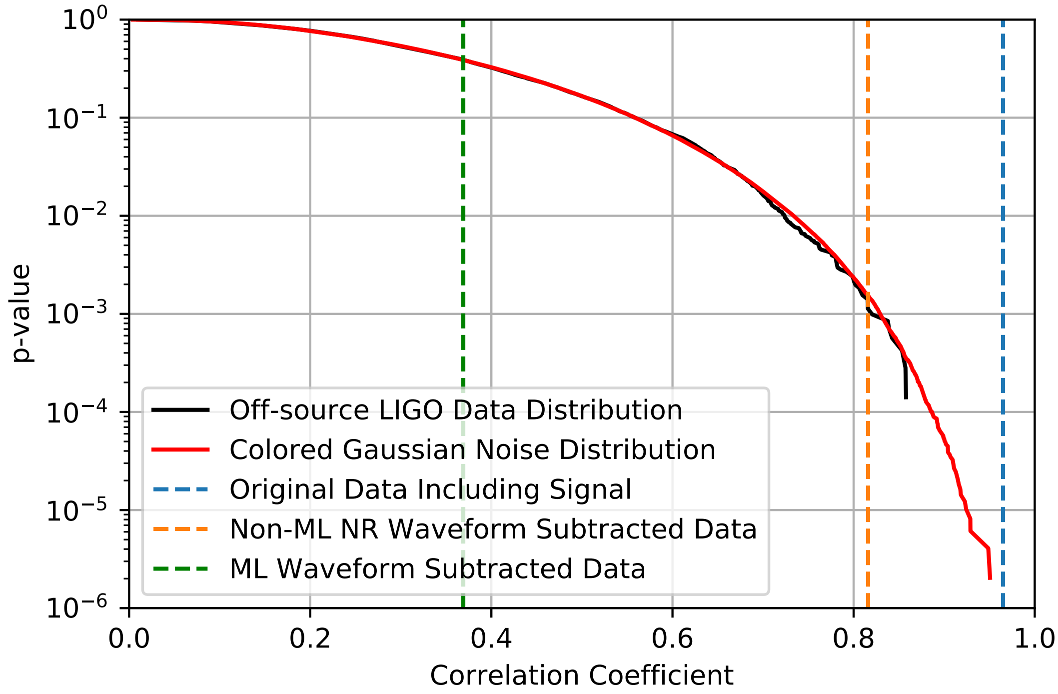

Figure 3. Cumulative distribution of correlations in simulated colored Gaussian noise and detector data away from GW150914, using the same bandpass and notching, and correlation time window settings as used to produce Fig. 2. The minimum correlation values near the time of flight difference in Fig. 2 are plotted as vertical lines for reference.

likelihood function for the two detectors, subsequently subtracting it from the data will not necessarily produce uncorrelated residuals, since the coherence algorithm assumes a consistent GR source signal with amplitude and phasing at the detectors appropriate for a tensorial gravitational wave. It is possible that there are other correlations between the detectors that are not coherent or consistent with GR binary mergers. Checking the resulting residuals with the Pearson correlation coefficient is a consistency test as to whether such incoherent or non GR-like correlations exist.

We calculate the Pearson correlation coefficient using the same method as Ref. [2]. The result is shown in Fig. 1 for the data released for Fig. 1 of Ref. [3], and in Fig. 2 for the updated calibration data available from GWOSC [5]. Both blue curves depict the correlations with the signal still included. The orange curves show the correlations after the NR waveform from Ref. [3] is subtracted. The green curve in Fig. 2 shows the correlations for our residuals after subtracting the maximum-likelihood waveform of Ref. [21].

Comparison of the blue and orange curves in Fig.1 and Fig.2 shows that the effect of the updated calibration on the correlation results is minimal. However, in Fig.2 the maximum anti-correlation near the detected time offset of 6.9 ms is reduced from 0.799 for the non-ML waveform, to 0.369 for the maximum-likelihood waveform of Ref. [21]. This result is due to the use of more accurate intrinsic parameters (such as the masses and spins of the black

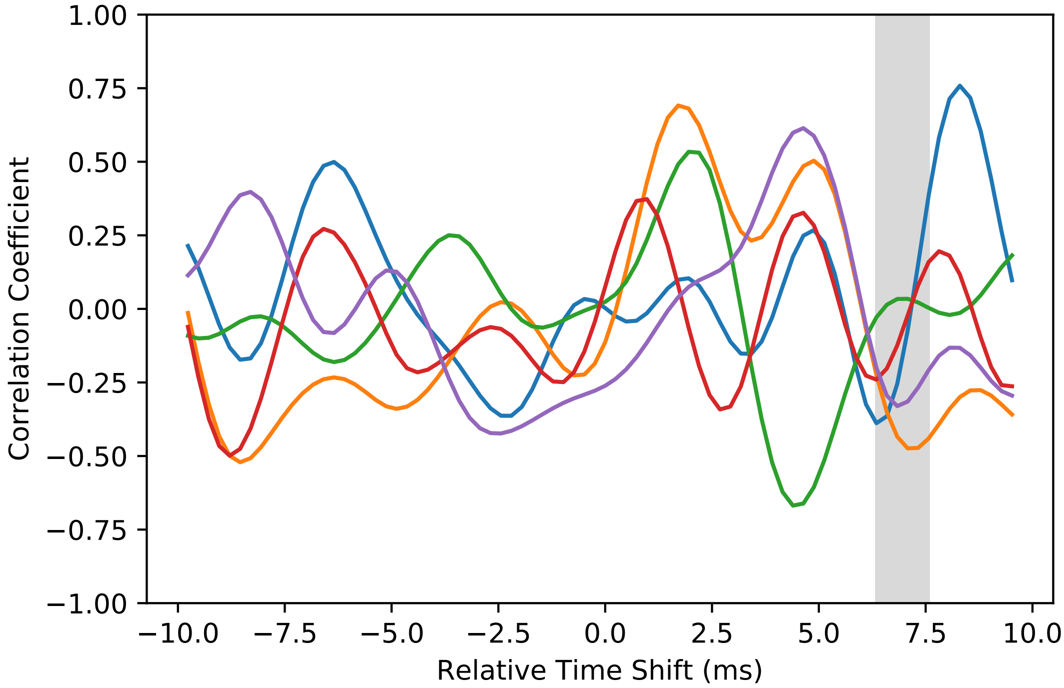

Figure 4. Five examples drawn at random from simulations of uncorrelated colored Gaussian noise. The same grey shaded region as Fig. 1 is shown. This is the same region that is used to produce the curves plotted in Fig. 3.

holes) and extrinsic parameters (that describe how the source is related to the detectors) of the subtracted template waveform.

It is important to ask how significant any measured correlations are. To answer this question we generated simulated colored Gaussian noise, from which we calculated $O(10^5)$ independent Pearson correlation coefficients. We performed the same test with $O(10^4)$ samples from the maximum-likelihood subtracted data, in segments of 61 ms over periods of 256 seconds both before and after GW150914. For this choice of segments, the 20 ms period of data corresponding to the GW150914 peak correlation is not included. The Gaussian noise was colored using a PSD estimated from the real detector data and was then processed in the same way as the data used in Fig. 2, with a bandpass and notch filter. The p-values are calculated by finding the largest correlation or anti-correlation around a time offset between the detectors of 7 ± 0.5 ms. (In practice this is extended to the nearest complete time sample on either side, so that the actual range runs from 6.35 ms offset to 7.57 ms with sampling at 4096 Hz.) The results for this background are displayed in Fig. 3. Five example realizations, chosen at random from the distribution, are additionally displayed in Fig. 4.

Figure 3 shows that the p-value for the correlation of the non-ML waveform residuals is ~ 0.001, whereas the p-value for the ML subtracted residuals is only ~ 0.4. While the residuals of Fig. 1 of Ref.[3] contain statically significant correlations at the 0.001 level, the correlations in the residuals using the ML waveform of Ref.[21] are not statistically significant.

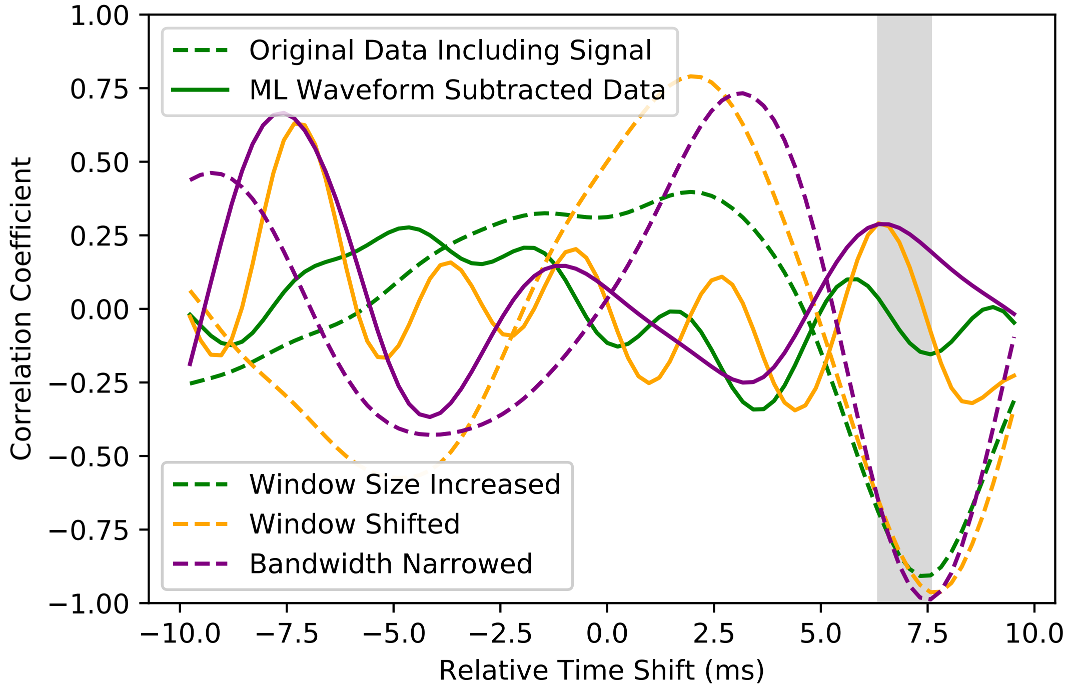

Figure 5. Plot showing the effect of changes to the correlation parameters. The dotted lines are data with the signal in, all of which peak strongly at the same value of time shift which is largely unaffected by the changes. All the solid lines are for the maximum-likelihood waveform subtracted residual data and the effect of the changes has a noticeable effect on the value of the correlation near the offset time of 7ms.

The data with the GW150914 event still included are of course highly significant compared to the Gaussian background.

It is also clear from Fig. 3 that the simulated Gaussian background and the background estimated in real data from times away from GW150914 do not differ greatly at the level studied here. A Kolmogorov–Smirnov test on the two distributions has a p-value of 0.95. This value is consistent with the null hypothesis that they are drawn from the same distribution. This does not imply that the real data is strictly Gaussian, but gives a sense of the differences over the time stretch considered (512 seconds of data centered around GW150914).

3 Robustness of the analysis

Our results in Sec. 2 and those of Ref. [2] were generated with a particular choice of the correlation parameters. The frequency bandwidth from 35 Hz to 350 Hz was chosen to capture the majority of the signal power for display in Fig. 1 of Ref. [3]. We used the simple notching and list of spectral lines taken from the tutorial of Ref. [5]. The width and start time of the correlation window corresponds roughly to the four loudest peaks visible in Fig. 1 of Ref. [3]. None of these choices are set in stone and other choices are possible. Entirely different methods are used to derive the LVC results [9, 10, 18, 25].

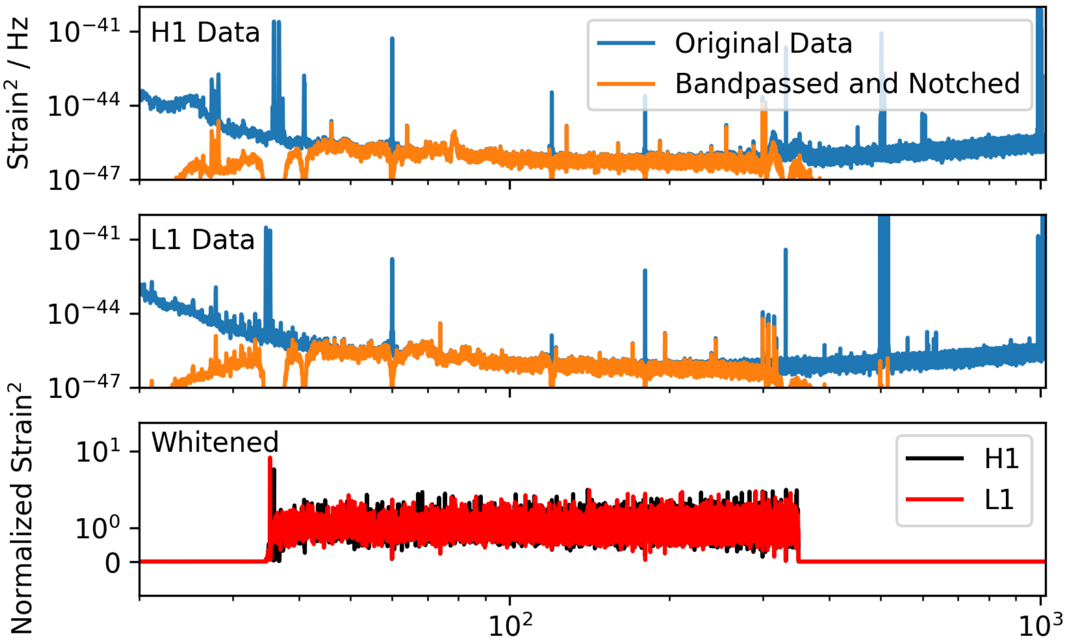

Figure 6. The power spectral density of the data before and after applying the band-pass and notching procedure, used to produce Fig. 1 of Ref. [3]. It is apparent that not all of the spectral lines are removed by the simple choices made to produce Fig. 1 of Ref. [3]. The power spectral density after applying the whitening procedure employed in this section is also shown for comparison. The remaining peak in the power spectral density of the whitened data is due to incomplete down-weighting of the calibration lines.

In this section, we investigate the effect of modifying some of the correlation parameter choices. We separately try tripling the window size, shifting the start time of the window earlier by 5 ms, and narrowing the bandpass filter to the range 60–220 Hz. As can be seen in Fig. 5, these changes do not have a significant effect on the location of the peak correlation for data containing the GW150914 signal, but they do have a noticeable effect on the value of the correlation statistic for the maximum-likelihood waveform subtracted residual data. Indeed, at the low levels of correlation seen for these residuals, relatively small changes in the correlation parameters can shift the location of the anti-correlation peak and even turn an anti-correlation into a correlation near 7ms time separation. However, none of these correlations in our residual data are at a level that is statistically significant when compared to simulated Gaussian noise.

4 Whitening vs bandpassing the data

The LIGO detectors are only sensitive to gravitational waves within a certain frequency range. At low frequencies, the detectors' sensitivities are dominated by seismic and thermal noise and at high frequencies by shot noise of the laser light. The detectors' sensitivities are not

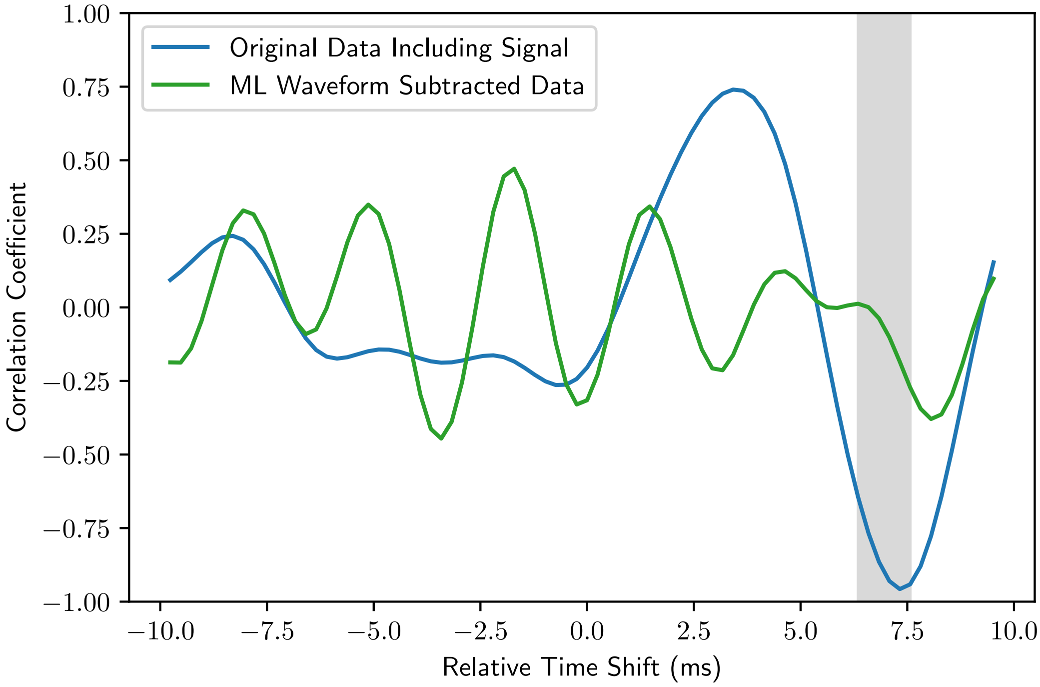

Figure 7. Correlations of residuals after whitening and subtracting the maximum-likelihood model waveform. The maximum and minimum peaks in the ML residuals (green line) are now away from the time of flight delay of ~ 7 ms.

white in the frequency domain [29]. To make the GW150914 signal visible in Fig. 1 of [3], a band pass between 35 and 350 Hz was applied to the data.

The LIGO detector data also contain strong spectral lines, small regions of frequency space where the noise is particularly high. These occur at, for example, the 60 Hz line of electrical power transmission, the violin modes of the mirror suspensions, and at calibration frequencies used to calibrate the detectors. In certain cases, these spectral lines will be explicitly correlated between the Hanford and Livingston detectors for a given offset. To produce Fig. 1 of Ref. [3], in addition to bandpassing between 35 Hz and 350 Hz, a number of these loud spectral lines were removed by notching, in particular, the 60 Hz line and its harmonics at 120 and 180 Hz. A list of these notched spectral lines was made available at Ref. [5]. However, to illustrate the signal in Fig. 1 of Ref. [3], it was not necessary to remove all spectral lines. The result of this bandpassing and notching is shown in the upper two rows of Fig. 6. Although the loudest spectral lines have been removed, some noticeable lines remain.

The majority of LVC analyses use whitening of the data rather than bandpassing in producing their results. This includes the pipelines used to assess significance, both with GR templates [10] and without [9]; to estimate the binary black hole parameters [25]; and to test the robustness of GR templates in describing the data [18]. Whitening has the advantage that it can suppress spectral lines without a prescribed list. Figure 6 shows a comparison of the whitened data and the bandpassed and notched data. We see that there remain several lines in the bandpassed and notched data, and that the power spectrum varies significantly over the frequency band. Without needing an explicit list, the whitening procedure has down-weighted the majority of spectral lines.

If instead of bandpassing and notching the detector data – as done for Fig. 1 of Ref. [3] and used for the residuals analysed by Ref. [2] – the data is whitened and the maximum-likelihood waveform (also suitably whitened) is subtracted, then we obtain the Pearson correlation coefficient shown in Fig. 7. Similar to the results of the bandpassing and notching procedure in Sec. 2, the anti-correlation observed around 6.9 ms is not statistically significant.

5 Conclusions

We have investigated the question of whether there are statistically significant correlations between the data of the LIGO Hanford and Livingston detectors around the time of GW150914. Detailed examinations of potential sources of detector correlations have also been carried out by the LVC [11, 30–32]. We have focused on 512 seconds of data around the time of GW150914 and employed the Pearson correlation coefficient test to examine whether the results of Ref. [2] pose a potential problem with the detection of GW150914. Our residuals are obtained by subtracting the maximum-likelihood general-relativity based template found in Ref. [21] from the data. Both our residuals and maximum-likelihood waveform are available at `www.github.com/gwastro/gw150914_investigation`. We compute statistical significance by calculating the Pearson correlation from samples of uncorrelated colored Gaussian data. We reproduce the result of Ref. [2] for LIGO data containing the GW150914 signal, but find no statistically significant correlations after subtracting the maximum-likelihood waveform from the data. Although our test statistic and residual data are different, our findings are consistent with the study of residual data in Ref. [18]. Furthermore, a similar analysis was also performed in Ref. [33] which likewise concluded that the residual data was consistent with noise. We therefore find no reason to doubt the significance statements reported in Refs. [3, 34, 35] and the conclusion that GW150914 is a gravitational-wave signal.

Acknowledgments

We thank Sylvia Zhu, Sebastian Khan, Peter Shawhan, Martin Green, and John Moffat for their comments. We acknowledge the Max Planck Gesellschaft for support. ABN and DAB thank Andrew Jackson, Hao Liu and Pavel Naselsky for helpful discussions and the 2017 Kavli Summer Program in Astrophysics at the Niels Bohr Institute in Copenhagen and DARK University of Copenhagen for support during this work. The 2017 Kavli Summer Program program was supported by the the Kavli Foundation, Danish National Research Foundation (DNRF), the Niels Bohr International Academy and DARK. DAB thanks Will Farr for helpful discussions and NSF award PHY-1707954 for support.

References

[1] H. Liu and A. D. Jackson, *Possible associated signal with GW150914 in the LIGO data*, *JCAP* **1610** (2016) 014 [1609.08346].

[2] J. Creswell, S. von Hausegger, A. D. Jackson, H. Liu and P. Naselsky, *On the time lags of the LIGO signals*, *JCAP* **1708** (2017) 013 [1706.04191].

[3] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116** (2016) 061102 [1602.03837].

[4] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein and B. Stephens, *The LIGO Open Science Center*, J. Phys. Conf. Ser. **610** (2015) 012021 [1410.4839].

[5] Gravitational Wave Open Science Center (GWOSC). https://doi.org/10.7935/K5MW2F23.

[6] LIGO Scientific collaboration, B. P. Abbott et al., *Calibration of the Advanced LIGO detectors for the discovery of the binary black-hole merger GW150914*, Phys. Rev. **D95** (2017) 062003 [1602.03845].

[7] C. Cahillane, J. Betzwieser, D. A. Brown, E. Goetz, E. D. Hall, K. Izumi et al., *Calibration uncertainty for Advanced LIGO's first and second observing runs*, Phys. Rev. **D96** (2017) 102001 [1708.03023].

[8] A. Viets et al., *Reconstructing the calibrated strain signal in the Advanced LIGO detectors*, Class. Quant. Grav. **35** (2018) 095015 [1710.09973].

[9] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *Observing gravitational-wave transient GW150914 with minimal assumptions*, Phys. Rev. **D93** (2016) 122004 [1602.03843].

[10] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *GW150914: First results from the search for binary black hole coalescence with Advanced LIGO*, Phys. Rev. **D93** (2016) 122003 [1602.03839].

[11] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914*, Class. Quant. Grav. **33** (2016) 134001 [1602.03844].

[12] B. Allen, χ^2 *time-frequency discriminator for gravitational wave detection*, Phys. Rev. **D71** (2005) 062001 [gr-qc/0405045].

[13] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown and J. D. E. Creighton, *FINDCHIRP: An Algorithm for detection of gravitational waves from inspiraling compact binaries*, Phys. Rev. **D85** (2012) 122006.

[14] S. Klimenko et al., *Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors*, Phys. Rev. **D93** (2016) 042004 [1511.05999].

[15] S. A. Usman et al., *The PyCBC search for gravitational waves from compact binary coalescence*, Class. Quant. Grav. **33** (2016) 215004 [1508.02357].

[16] T. Dal Canton et al., *Implementing a search for aligned-spin neutron star-black hole systems with advanced ground based gravitational wave detectors*, Phys. Rev. **D90** (2014) 082004 [1405.6731].

[17] C. Messick et al., *Analysis Framework for the Prompt Discovery of Compact Binary Mergers in Gravitational-wave Data*, Phys. Rev. **D95** (2017) 042001 [1604.04324].

[18] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *Tests of general relativity with GW150914*, Phys. Rev. Lett. **116** (2016) 221101 [1602.03841].

[19] A. H. Nitz, I. W. Harry, J. L. Willis, C. M. Biwer, D. A. Brown, L. P. Pekowsky et al., "PyCBC Software v1.12.4." https://github.com/gwastro/pycbc, 2018. 10.5281/zenodo.1479480.

[20] P. Naselsky, A. D. Jackson and H. Liu, *Understanding the LIGO GW150914 event*, JCAP **1608** (2016) 029 [1604.06211].

[21] C. M. Biwer, C. D. Capano, S. De, M. Cabero, D. A. Brown, A. H. Nitz et al., *PyCBC Inference: A Python-based parameter estimation toolkit for compact binary coalescence signals*, 1807.10312.

[22] C. M. Biwer, C. D. Capano, S. De, M. Cabero, D. A. Brown, A. H. Nitz et al., "v1.0 data release for the PyCBC Inference paper." https://github.com/gwastro/pycbc-inference-paper, 2018. 10.5281/zenodo.1492649.

[23] G. Lovelace et al., *Modeling the source of GW150914 with targeted numerical-relativity simulations*, *Class. Quant. Grav.* **33** (2016) 244002 [1607.05377].

[24] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *Directly comparing GW150914 with numerical solutions of Einstein's equations for binary black hole coalescence*, *Phys. Rev.* **D94** (2016) 064035 [1606.01262].

[25] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *Properties of the Binary Black Hole Merger GW150914*, *Phys. Rev. Lett.* **116** (2016) 241102 [1602.03840].

[26] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme et al., *Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms*, *Phys. Rev. Lett.* **113** (2014) 151101 [1308.3271].

[27] R. A. Mercer et al., "LIGO Algorithm Library v6.49." https://git.ligo.org/lscsoft/lalsuite, 2018.

[28] A. B. Nielsen, A. H. Nitz, C. Capano, and D. A. Brown, "GW150914 Investigation Supplementary Material." www.github.com/gwastro/gw150914_investigation, 2018.

[29] B. P. Abbott et al., *Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy*, *Phys. Rev.* **D93** (2016) 112004 [1604.00439].

[30] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *Effects of data quality vetoes on a search for compact binary coalescences in Advanced LIGO's first observing run*, *Class. Quant. Grav.* **35** (2018) 065010 [1710.02185].

[31] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *Upper Limits on the Stochastic Gravitational-Wave Background from Advanced LIGO's First Observing Run*, *Phys. Rev. Lett.* **118** (2017) 121101 [1612.02029].

[32] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *All-sky search for short gravitational-wave bursts in the first Advanced LIGO run*, *Phys. Rev.* **D95** (2017) 042003 [1611.02972].

[33] M. A. Green and J. W. Moffat, *Extraction of black hole coalescence waveforms from noisy data*, *Phys. Lett.* **B784** (2018) 312 [1711.00347].

[34] Virgo, LIGO Scientific collaboration, B. P. Abbott et al., *Binary Black Hole Mergers in the first Advanced LIGO Observing Run*, *Phys. Rev.* **X6** (2016) 041015 [1606.04856].

[35] A. H. Nitz, C. Capano, A. B. Nielsen, S. Reyes, R. White, D. A. Brown et al., *1-OGC: The first open gravitational-wave catalog of binary mergers from analysis of public Advanced LIGO data*, 1811.01921.