Physics Letters B 784 (2018) 312–323

Contents lists available at ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Extraction of black hole coalescence waveforms from noisy data

Martin A. Green [a], J.W. Moffat [a,b,*]

[a] Perimeter Institute for Theoretical Physics, Waterloo ON N2L 2Y5, Canada
[b] Department of Physics and Astronomy, University of Waterloo, Waterloo ON N2L 3G1, Canada

ARTICLE INFO

Article history:
Received 19 October 2017
Received in revised form 7 August 2018
Accepted 8 August 2018
Available online 11 August 2018
Editor: M. Doser

Keywords:
Gravitational waves
Black hole coalescence
Signal extraction

ABSTRACT

We describe an independent analysis of LIGO data for black hole coalescence events. Gravitational wave strain waveforms are extracted directly from the data using a filtering method that exploits the observed or expected time-dependent frequency content. Statistical analysis of residual noise, after filtering out spectral peaks (and considering finite bandwidth), shows no evidence of non-Gaussian behaviour. There is also no evidence of anomalous causal correlation between noise signals at the Hanford and Livingston sites. The extracted waveforms are consistent with black hole coalescence template waveforms provided by LIGO. Simulated events, with known signals injected into real noise, are used to determine uncertainties due to residual noise and demonstrate that our results are unbiased. Conceptual and numerical differences between our RMS signal-to-noise ratios (SNRs) and the published matched-filter detection SNRs are discussed.

1. Introduction

The LIGO Scientific Collaboration and Virgo Collaboration (LIGO) have reported six events in which spatial strain measurements are consistent with gravitational waves produced by the inspiral and coalescence of binary black holes [1–7]. In each case, similar signals were detected at the Hanford, Washington (H) and Livingston, Louisiana (L) sites, with a time offset less than the 10 ms inter-site light travel time. The latest event, GW170814, was also detected at the Virgo site, in Italy. The likelihood of these being false-positive detections is reported as: less than 10^{-4} yr^{-1} for GW150914, GW151226, GW170104 and GW170814; less than 3.3×10^{-4} yr^{-1} for GW170608; and 0.37 yr^{-1} for LVT151012. This paper examines the earliest four events, GW150914, LVT151012, GW151226, and GW170104, using data made available at the LIGO Open Science Center (LOSC) [8].

GW150914 had a sufficiently strong signal to stand above the noise after removing spectral peaks and band-pass filtering [9]. For the other events, primary signal detection involved the use of matched filters, in which the measured strain records are cross-correlated with template waveforms derived using a combination of effective-one-body, post-Newtonian and numerical general relativity techniques [10]. Matched filter signal detection treats each template as a candidate representation of the true gravitational wave signal and measures how strongly the template's correlation with the measured signal exceeds its expected correlation with detector noise. Strong correlation indicates a good match, but the true signal may not exactly match any of the templates. The dependence of matched filter outputs on template parameters informs estimates of the physical parameters of the source event [11,12].

Template-independent methods, using wavelets, were used to reconstruct the waveform for GW150914 from the data [9, 12–14], achieving 94% agreement with the binary black hole model. A template-independent search for generic gravitational wave bursts also detected GW170104, but with lower significance than the matched filter detection. For GW170104 a morphology-independent signal model based on Morlet–Gabor wavelets was used, following detection, to construct a de-noised representation of the binary black hole inspiral waveform from the recorded strain data [5]. This was found to have an 87% overlap with the maximum-likelihood template waveform of the binary black hole model, which is statistically consistent with the uncertainty of the template.

In [15], a Rudin–Osher–Fatemi total variation method was used to de-noise the signal for GW150914, yielding a waveform comparable to that obtained with a Bayesian approach in [1]. The same

* Corresponding author.
E-mail addresses: mgreen@perimeterinstitute.ca (M.A. Green), jmoffat@perimeterinstitute.ca (J.W. Moffat).

https://doi.org/10.1016/j.physletb.2018.08.009

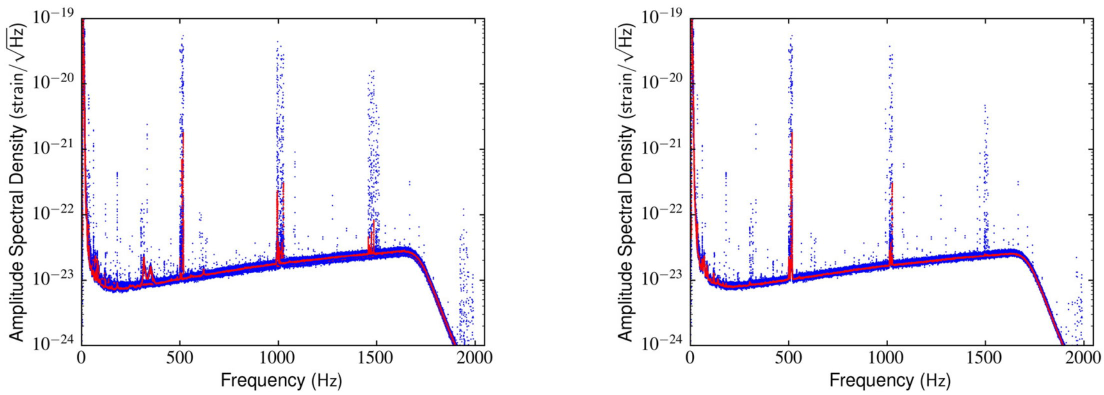

Fig. 1. Amplitude spectral density (ASD $\equiv \sqrt{S(f)}$) of detector noise (blue) and the smoothed baseline $\sqrt{S_b(f)}$ (red) from 4096 s strain records at Hanford (left panel) and Livingston (right panel) at the time of GW150914. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

authors also applied dictionary learning algorithms to de-noise the Hanford signal for GW150914 [16].

Cresswell et al. [17] have independently analyzed the LIGO data for GW150914, GW151226 and GW170104. They report correlations in the residual noise at the two sites—after subtracting model templates obtained from the LOSC—and suggest that a clear distinction between signal and noise remains to be established. Recognizing that a family of template waveforms may “fit the data sufficiently well”, they claim that “the residual noise is significantly greater than the uncertainly introduced by the family of templates”. The claim of correlations in the residuals is contrary to analysis in [10]. We discuss below why we believe [17] is erroneous.

The present work introduces a new method for extracting signal waveforms from the noisy strain records of black hole binary coalescence events provided by the LOSC. Noise that is inconsistent with prior knowledge regarding timing and a reasonable-fit template for an event is selectively filtered out to better reveal the gravitational wave signal. Our method relies on knowledge obtained using the matched-filter techniques, discussed above, for signal detection and identification of a reasonable-fit template. We rely on the (approximate) signal event times given by the LOSC and the broad, time-dependent spectral features of the black hole coalescence templates.[1] In the case of GW150914, we have also done signal extraction without using the template, but assuming the event has the smoothly varying frequency content typical of black hole inspiral, merger and ringdown. Our waveform extraction method does not use the templates' phase or detailed amplitude information—instead such information is obtained directly from the recorded data. The extracted waveforms are compared with similarly filtered templates to determine best-fit amplitude and phase parameters and associated uncertainties.

Our limited objective in this work was to independently analyse data provided by the LOSC in the hope of obtaining clean representations of the black hole coalescence strain signals that could be compared with the provided templates and published results. Our analysis method is *not* designed to *detect* gravitational wave events, and should not be confused with the matched-filter techniques used for such detection. Application of our method to estimation of physical parameters or to events other than black hole coalescence is beyond the scope of the present work.

In Section 2 we describe the characterization of detector noise and identification and removal of spectral peaks due to AC line power (60 Hz and harmonics), calibration signals and other non-astrophysical causes. Band-pass filters are used to remove the high amplitude noise below about 30 Hz and at frequencies higher than expected in the gravity wave events. Statistical analysis gives no indication that the filtered signals differ significantly from band-limited Gaussian noise, with the exception of a few obvious glitches. Also, no significant correlation is found between detectors. Section 3 describes the use of time-frequency bands to further reduce the influence of noise that masks the astrophysical strain signals. This allows determination of the event time, phase and amplitude differences between the two detectors, and construction of a coherent signal once the Hanford signal time, phase and amplitude are adjusted to match Livingston (which is taken as reference). The clean signals are compared with the reasonable-fit templates provided by the LOSC. Analysis of simulated events, with known signals injected into real noise, demonstrates the reliability and uncertainties associated with our signal extraction method. In Section 4 we discuss the relationship of our work to matched-filter signal detection, provide some remarks on signal-to-noise ratios, and comment on correlations of noise and residuals between detectors. The final section provides brief conclusions.

2. Signal cleaning and noise characterization

The LOSC has made available time series records of the measured strain data for each of the reported events. The events are roughly central within short (32 s) and long (4096 s) records, each available at 4096 samples per second (sps) and 16384 sps. The 4096 sps records—used in the present work—were constructed from the 16384 sps records by decimation.

From each long strain record, $s_l(t)$, the power spectral density (PSD), $S(f)$, was constructed using Welch's average periodogram method with overlapping 64 s segments and Planck windows [18] with 50% tapers. This enabled identification of the positions and widths of numerous sharp spectral peaks corresponding to AC powerline harmonics, detector calibration signals, and other deterministic sources. A smoothed PSD baseline, $S_b(f)$ was constructed, corresponding to the PSD with the narrow peaks removed. This was used as the PSD for subsequent analysis of the given event and detector, and for whitening of signals. The square roots of the PSDs and baselines (i.e., the amplitude spectral densities (ASDs)) found for GW150914 are shown in Fig. 1.

Order 2 Butterworth notch filters were used to remove the deterministic signals from the 32 s records, $s(t)$, leaving a clean

[1] It is noted at https://losc.ligo.org/ that the provided numerical relativity template waveforms are consistent with the parameter ranges inferred for the observed events but were not tuned to precisely match the signals. “The results of a full LIGO-Virgo analysis of this BBH event include a set of parameters that are consistent with a range of parameterized waveform templates.” While a full LIGO-Virgo analysis may combine analyses of many templates, each consistent with the data, the variation of the time-dependent spectral features of the reasonable-fit templates will be inconsequential for the present work.