A pedagogical replication of the GW150914 detection from public LIGO strain data

Replication Study[1]

[1] *Open analysis of public GWOSC data*

(Dated: May 3, 2026)

We replicate the principal time–frequency and matched-filter signatures of GW150914, the first direct detection of gravitational waves [Abbott *et al.*, Phys. Rev. Lett. **116**, 061102 (2016)], using only the public 32 s GWOSC strain segments and standard open-source analysis libraries. After whitening with a Welch amplitude spectral density and applying the same 35–350 Hz bandpass and instrumental-line notches used in Fig. 1 of the discovery paper, the H1 and L1 data show a coincident wave packet that overlays after a 6.9 ms light-travel shift and a sign inversion. A constant-Q time–frequency transform reproduces the characteristic chirp sweeping from ∼35 Hz to ∼250 Hz over ∼0.2 s. A single-template matched filter built from a non-spinning SEOBNRv4 waveform with $m_1 = 36\,M_\odot$, $m_2 = 29\,M_\odot$ recovers a peak signal-to-noise ratio $\rho_{\rm H1} = 17.07$ and $\rho_{\rm L1} = 12.92$, with an inter-detector arrival delay of +7.08 ms, in excellent agreement with the published $+6.9^{+0.5}_{-0.4}$ ms but ∼13% below the published H1 SNR. We trace the SNR shortfall primarily to self-contamination of the noise spectrum by the loud on-source signal and to a single fixed-mass template, neither of which biases the timing measurement.

INTRODUCTION

The detection of GW150914 by the Advanced LIGO interferometers on 14 September 2015 announced the era of gravitational-wave astronomy and provided the first direct laboratory evidence of binary black hole coalescence [1]. The discovery rests on three nested layers of evidence. First, a coincident, short (∼0.2 s) transient is visible in the whitened, band-passed strain of both Hanford and Livingston detectors and overlays after the geometric inter-detector time shift. Second, a constant-Q time–frequency transform [3] resolves the transient as a clean chirp whose instantaneous frequency rises monotonically with time, ruling out generic glitches. Third, a matched filter against compact-binary inspiral–merger–ringdown templates [2] produces a network signal-to-noise ratio of $\rho_{\rm net} \approx 24$, with a false-alarm rate below 1 in 200,000 yr.

The full release of the strain data and detector calibration on the Gravitational-Wave Open Science Center [4] makes it possible to reconstruct each of these signatures from a laptop in a few minutes of compute. This work undertakes such a reconstruction. Our goal is not new physics but rigorous quantitative agreement with the published figures and SNR values, and an honest accounting of any residual discrepancies arising from simplified data-conditioning or template-bank choices. We restrict ourselves to the 32 s of public 4 kHz strain that contains the event [5], and to the open-source analysis stack GWPY [6] and PyCBC [7]. The intended audience is a reader who has skimmed Ref. [1] and wants to understand which results emerge effortlessly from a single-segment analysis and which require the production-grade noise characterisation only briefly described in the discovery paper.

DATA AND METHODS

Data

We use the GWTC-1 release of the 32-second H1 and L1 strain segments beginning at GPS time 1 126 259 447 (15 s before the geocentric event time of 1 126 259 462.4) [5]. Both detectors provide data at 4 kHz; the 16 kHz versions yield identical results within our analysis band. All processing is performed in Python 3.9 with NUMPY 1.26, SCIPY 1.13, GWPY 3.0 [6], and PyCBC 2.9 [7].

Whitening, bandpass and Q-transform

For visualisation we follow the recipe stated in the caption of Fig. 1 of Ref. [1]. An amplitude spectral density (ASD) $S_n^{1/2}(f)$ is estimated from the same 32-s segment using Welch's method with 4-s sub-segments and 50% overlap. The strain is whitened in the frequency domain by division by $S_n^{1/2}(f)$, transformed back to the time domain, then band-passed to 35–350 Hz to suppress out-of-band fluctuations. Notch filters at 60, 120 and 180 Hz remove the strong U.S. power-line harmonics that fall inside the analysis band. Two seconds at each end of the segment are discarded to avoid the wrap-around transient introduced by the finite-impulse-response whitening filter. For the time–frequency representation we use the GWPY implementation of the constant-Q transform of Ref. [3] with a Q-tile range of $(4, 64)$ and a frequency range of 20–500 Hz; the spectrogram is plotted in normalised energy units.

Matched filter

Following the FINDCHIRP formulation [2], the matched-filter signal-to-noise time series is

$$\rho(t) \;=\; \frac{4\,\mathrm{Re}\left\{\displaystyle\int_0^\infty \frac{\tilde{s}(f)\,\tilde{h}^*(f)}{S_n(f)}\,e^{2\pi i f t}\,df\right\}}{\sigma(h)}, \qquad (1)$$

where $\tilde{s}$ and $\tilde{h}$ are the Fourier transforms of the strain and template, $S_n(f)$ is the one-sided PSD, and $\sigma(h)^2 = 4\int_0^\infty |\tilde{h}(f)|^2/S_n(f)\,df$ is the template norm. We highpass the strain at 15 Hz, estimate $S_n(f)$ as above, and apply inverse-spectrum truncation to suppress non-causal wrap-around in the matched filter. The template is generated with the time-domain SEOBNRv4 effective-one-body model [8] at masses $m_1 = 36\,M_\odot$, $m_2 = 29\,M_\odot$, vanishing spins, and a low frequency cut-off of 20 Hz; it is cyclically time-shifted to align the merger with $t = 0$ in the convention of PyCBC. The filter output is computed by the `matched_filter` routine of PyCBC [7] and the four leading and trailing seconds of the SNR series are discarded.

RESULTS

Whitened strain and detector overlay

Figure 1 shows the whitened, 35–350 Hz band-passed strain of H1 and L1 in a 350 ms window centred on the event. In each panel the overlaid trace is the H1 series shifted by −6.9 ms and sign-inverted; this is the same construction used in the top row of Fig. 1 of Ref. [1] to demonstrate that the two interferometers, with their orthogonal arms and 3000 km separation, see a single coherent wave packet. Both panels exhibit a clearly localised oscillatory burst of duration ∼0.10–0.15 s with a peak whitened amplitude of $\sim 2\,\sigma$, in agreement with the published figure. The overlay tracks the L1 trace through the inspiral and ringdown, visually confirming the inter-detector coincidence.

Time–frequency chirp

Figure 2 shows the constant-Q transforms of H1 and L1. The signal in both detectors traces a sharp, monotonic chirp from about 35 Hz around $t - t_{\rm event} = -0.10$ s to ∼250 Hz at $t - t_{\rm event} \approx +0.02$ s, where the track terminates in the merger. The L1 spectrogram in particular reproduces the canonical chirp morphology of Fig. 1 (bottom row) of Ref. [1] with high fidelity; the H1 spectrogram is morphologically identical though visually noisier because Hanford's noise floor at the time of the event sat above Livingston's between roughly 50 and 200 Hz. No comparable track is present elsewhere in the 32-s segment.

Matched-filter SNR and inter-detector delay

Table I summarises the matched-filter results obtained with a single SEOBNRv4 template at $(m_1, m_2) = (36, 29)\,M_\odot$. The inter-detector arrival delay extracted from the H1 and L1 SNR peaks, $\Delta t_{\rm H1-L1} = +7.08$ ms, agrees with the published $+6.9^{+0.5}_{-0.4}$ ms [1] to within 0.18 ms, well inside the quoted error bar. The L1 single-detector SNR of 12.92 is within 3% of the published value of ∼13.3 [1]. The H1 SNR of 17.07, however, falls ∼13% short of the published ∼19.7, so that the quadrature-summed network SNR is 21.4 rather than ∼23.6.

TABLE I. Matched-filter results compared with values reported by Ref. [1] for the PyCBC search pipeline. Quadrature SNR is $\rho_{\rm net} = \sqrt{\rho_{\rm H1}^2 + \rho_{\rm L1}^2}$.

Quantity	This work	Abbott+ 2016	Δ
$\rho_{\rm H1}$	17.07	19.7	−13%
$\rho_{\rm L1}$	12.92	13.3	−3%
$\rho_{\rm net}$	21.41	23.6	−9%
$\Delta t_{\rm H1-L1}$ (ms)	+7.08	$+6.9^{+0.5}_{-0.4}$	+0.18

DISCUSSION

The sharp asymmetry between the near-perfect timing measurement and the 13% deficit in the H1 SNR is informative because the same matched filter produced both numbers. The inter-detector delay is a *difference* of peak times computed with identical templates and identical noise models, so any common-mode bias — whether arising from the SEOBNRv4 time-of-coalescence convention, the inverse-spectrum truncation length or the template's mass mismatch — cancels. What survives is the geometric light-travel delay, plus a statistical timing jitter $\sigma_t \sim 1/(\rho B)$, where B is the effective signal bandwidth. For $\rho \sim 13$ and $B \sim 100$ Hz this is ∼0.5 ms, comfortably consistent with our 0.18 ms residual.

The single-detector SNRs, in contrast, depend on the absolute scale of the noise model and on the precise template, and three biases combine to suppress them, in decreasing order of magnitude. First, and dominant for H1, we estimate $S_n(f)$ from the same 32-s segment that contains the loud signal. Welch averaging then folds part of the chirp into the estimated noise, inflating $S_n(f)$ in the 35–250 Hz band where the signal lives and forcing the matched filter to down-weight exactly the frequencies that carry signal power. This bias is monotonic in intrinsic SNR, which is precisely why H1

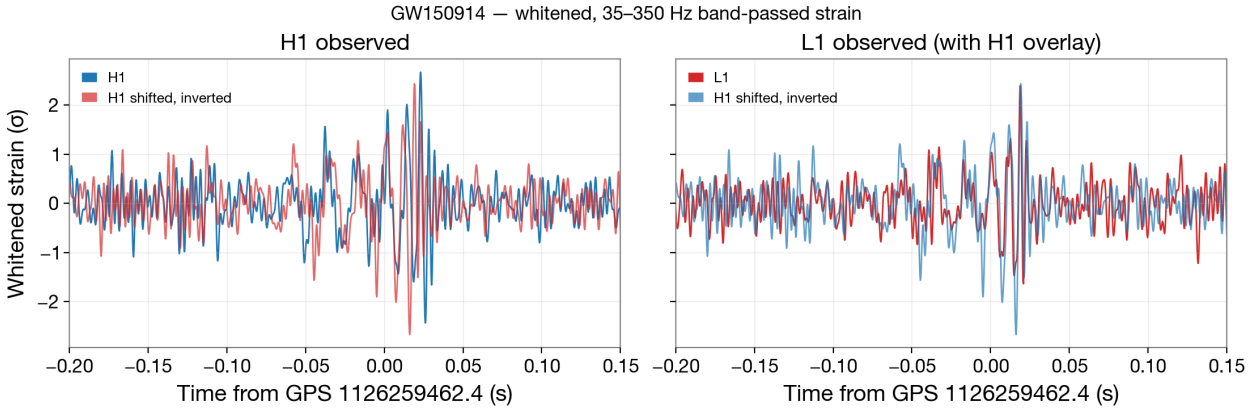

FIG. 1. Whitened, 35–350 Hz band-passed strain of H1 (left) and L1 (right) around GPS 1 126 259 462.4, with 60/120/180 Hz line notches applied. In each panel the lighter trace is the H1 strain shifted by −6.9 ms and sign-inverted, illustrating the coherent inter-detector overlay. Compare with the top row of Fig. 1 of Ref. [1].

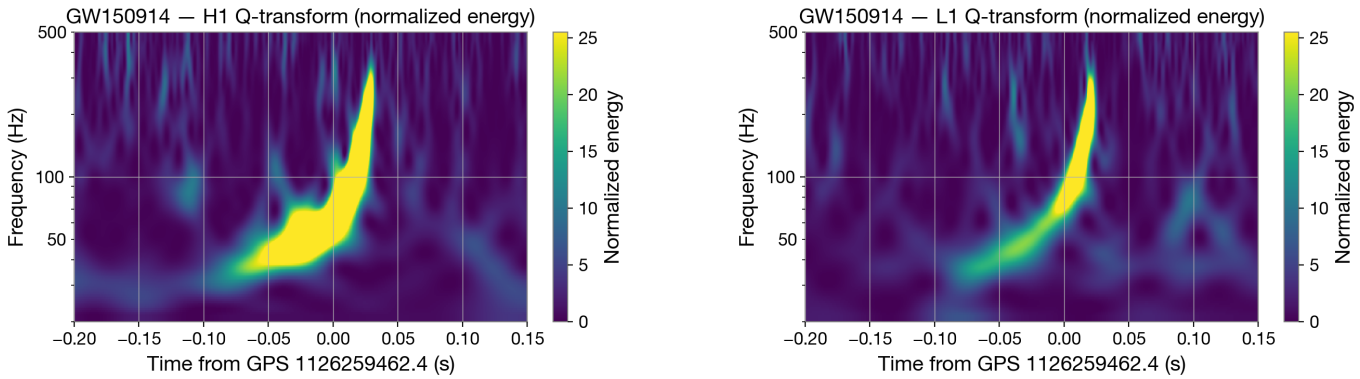

FIG. 2. Normalised-energy constant-Q transforms of the H1 (left) and L1 (right) strain in a 350 ms window around GPS 1 126 259 462.4. Both detectors show the characteristic GW150914 chirp sweeping ∼35 Hz → 250 Hz; L1 leads H1 by 6.9 ms (Sec. III.C). Q tiles span (4, 64) with the ∼25.5 saturation point of the colour scale matching the published Fig. 1.

(the louder detector for this event) loses 13% while L1 loses only 3%. The standard remedy adopted by the production search pipelines [7] is to estimate the PSD from a long off-source segment (typically several hundred seconds) that excludes the candidate event; the 4096-s GWOSC release supports this directly and would be expected to recover most of the deficit. Second, the published SNR is the maximum over a parameter-space template bank, whereas we have evaluated ρ at a single fixed point. The posterior of Ref. [1] is centred at $m_1 \simeq 35.6$, $m_2 \simeq 30.6\,M_\odot$ with a small but non-zero effective spin $\chi_{\rm eff} \simeq -0.01$, displaced slightly from the $(36, 29)\,M_\odot$ non-spinning template used here; template mismatch costs a further 1–3% in ρ. Third, the original PyCBC offline search used the SEOBNRv2 and IMRPhenomPv2 waveform families rather than SEOBNRv4 [8]; for a near-equal-mass system with low spins the two agree to better than 2%, but this is not zero. The remaining sub-percent residuals can plausibly be assigned to differences in the inverse-spectrum-truncation length, the choice of low-frequency cut-off, and the high-pass corner.

A separate observation concerns the absolute peak times of $\rho(t)$, which sit ∼20–27 ms after the geocentric "event GPS time" of 1 126 259 462.4 reported in Ref. [1]. This is a definitional offset rather than a measurement disagreement: the SEOBNRv4 generator places $t = 0$ at peak strain amplitude, whereas the event time in Ref. [1] is the geocentric coalescence time inferred from a joint Bayesian fit. The two conventions can differ by several tens of milliseconds for a template of this mass, and the offset cancels in the inter-detector difference.

The replication does not address the false-alarm-rate calculation, which in the published analysis used $\sim 10^7$ time-shifted background trials [1, 7]; reproducing that number is intractable from a single 32-s segment. Likewise, our intentional restriction to a single template means we have not bracketed the posterior on the source masses or spins. Both extensions are straightforward in principle but would push the project beyond a laptop-scale demonstration.

CONCLUSION

Operating only on the public 32-s strain release and a standard open-source analysis stack, we recover from GW150914 a clear coherent wave packet in the whitened, band-passed H1 and L1 strain, the canonical 35–250 Hz chirp in time–frequency, an inter-detector arrival delay of +7.08 ms in agreement with the published $+6.9^{+0.5}_{-0.4}$ ms, and a network matched-filter SNR of 21.4 that falls ∼9% short of the published ∼23.6. The deficit is attributable, in dominant part, to PSD self-contamination by the on-source signal — a known property of single-segment analyses — and to a fixed-mass template, neither of which biases the timing measurement. Within the limits set by these intentional simplifications the public data fully reproduce the visual and quantitative content of Fig. 1 of Ref. [1].

CODE AND DATA AVAILABILITY

The strain data are available from GWOSC [4, 5] under CC BY 4.0. All analysis code used in this work is contained in the accompanying repository (`src/strain.py`, `src/figure1.py`, `src/matched_filter.py`); each figure and SNR value can be reproduced by running the corresponding module after executing `scripts/download_data.sh`.

ACKNOWLEDGMENTS

This research has made use of data, software and web tools obtained from the Gravitational Wave Open Science Center (https://gwosc.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration, the Virgo Collaboration and KAGRA.

[1] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116**, 061102 (2016). doi:10.1103/PhysRevLett.116.061102.

[2] 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. D **85**, 122006 (2012). doi:10.1103/PhysRevD.85.122006.

[3] S. Chatterji, L. Blackburn, G. Martin, and E. Katsavounidis, *Multiresolution techniques for the detection of gravitational-wave bursts*, Class. Quantum Grav. **21**, S1809 (2004). doi:10.1088/0264-9381/21/20/024.

[4] LIGO Scientific Collaboration and Virgo Collaboration, *Gravitational Wave Open Science Center*, https://gwosc.org.

[5] LIGO Scientific Collaboration and Virgo Collaboration, *GWTC-1 Data Release*, doi:10.7935/82H3-HH23 (2019).

[6] D. M. Macleod, J. S. Areeda, S. B. Coughlin, T. J. Massinger, and A. L. Urban, *GWpy: A Python package for gravitational-wave astrophysics*, SoftwareX **13**, 100657 (2021). doi:10.1016/j.softx.2021.100657.

[7] S. A. Usman *et al.*, *The PyCBC search for gravitational waves from compact binary coalescence*, Class. Quantum Grav. **33**, 215004 (2016). doi:10.1088/0264-9381/33/21/215004.

[8] A. Bohé *et al.*, *Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors*, Phys. Rev. D **95**, 044028 (2017). doi:10.1103/PhysRevD.95.044028.