MNRAS **000**, 1–?? (0000) Preprint 7 December 2016 Compiled using MNRAS LATEX style file v3.0

On similarity of binary black hole gravitational-wave skymaps: to observe or to wait?

Salvatore Vitale[1⋆], Reed Essick[1], Erik Katsavounidis[1], Sergey Klimenko[2] and Gabriele Vedovato[3]
[1] *LIGO, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2] *University of Florida, P.O.Box 118440, Gainesville, Florida, 32611, USA*
[3] *INFN, Sezione di Padova, via Marzolo 8, 35131 Padova, Italy*

arXiv:1611.02438v2 [gr-qc] 25 Nov 2016

7 December 2016

ABSTRACT
Localization estimates for GW150914, the first binary black hole detected by the LIGO instruments, were shared with partner facilities for electromagnetic follow-up. While the source was a compact binary coalescence (CBC), it was first identified by algorithms that search for unmodeled signals, which produced the skymaps that directed electromagnetic observations. Later on, CBC specific algorithms produced refined versions, which showed significant differences. In this paper we show that those differences were not accidental and that CBC and unmodeled skymaps for binary black holes will frequently be different; we thus provide a way to determine whether to observe electromagnetically as promptly as possible (following a gravitational-wave detection), or to wait until CBC skymaps become available, should they not be available in low latency. We also show that, unsurprisingly, CBC algorithms can yield much smaller searched areas.

1 INTRODUCTION

The first-ever gravitational-wave detection, the black-hole binary (BBH) GW150914, was identified in low-latency by algorithms that search for unmodeled signals (henceforth, burst algorithms)(Abbott et al. 2016c): `cWB` (Klimenko et al. 2016) and `LIB` (Lynch et al. 2015). These algorithms produced skymaps in low-latency and these were circulated to partner facilities for electromagnetic (EM) and neutrino follow-up with a latency of ∼2 days (Abbott et al. 2016f; Adrián-Martínez et al. 2016) (All the GCN pertinent to GW150914 can be found at http://gcn.gsfc.nasa.gov/other/GW150914.gcn3). Skymaps were later generated by a fast sky localization algorithm for compact binary coalescence (CBC) sources called `BAYESTAR` (Singer & Price 2016; Abbott et al. 2016f) and by a full parameter estimation algorithm that assumed a CBC model, `lalinference` (Veitch et al. 2015; Abbott et al. 2016d), and found to be different from the burst results. For example, the `lalinference` skymap released on January 13 2016 and had a 90% confidence level area of 620 deg^2, less than half of which overlapped with the 90% area of the burst skymaps (Abbott et al. 2016f,c). Many groups have already reported on their follow-up observations of GW150914 (Connaughton et al. 2016; Savchenko et al. 2016; Smartt et al. 2016; Soares-Santos et al. 2016; Evans et al. 2016; Ackermann et al. 2016; Tavani et al. 2016; Hurley et al. 2016; Palliyaguru et al. 2016; Gando et al. 2016; Díaz et al. 2016).

The skymap provided by `lalinference` is considered by the LIGO-Virgo collaboration to provide the definitive sky localization for this (and every) CBC source (Abbott et al. 2016f) and for this reason we will use such maps for the purpose of the study. The reconstruction of these two burst algorithms together with the `lalinference` CBC skymap was reported in Abbott et al. (2016c)[1]. Two types of differences are visible: from one side different parts of *the same* sky-ring (defined as the locus of points with fixed time-of-arrival difference between the LIGO instruments (Abbott et al. 2016a)) have posterior support; from the other, slightly different rings can have posterior support (Abbott et al. 2016f). Until the CBC skymaps were provided, EM partners observed regions of the sky where only the burst algorithms had support.

We note that only burst algorithms found GW150914 in low-latency because CBC specific algorithms did not search for BBH at that time (Abbott et al. 2016f), since BBH are not traditionally expected to be luminous in the EM. Later in the run, a low-latency BBH version of the CBC search algorithm `gstlal` (Cannon et al. 2012; Privitera et al. 2014) was enabled, which can be followed-up by the low-latency CBC sky localization algorithm `BAYESTAR` (Singer & Price 2016). It is thus plausible that from the second observing run onward only CBC skymaps will be released for BBH events, with `BAYESTAR` estimates available in close to real time (Singer & Price 2016) and with the definitive `lalinference` maps obtained typically within a few days (Veitch et al. 2015). However, one may envision situations where skymaps obtained by the low-latency burst

[1] FITS files for all skymaps can be found at LIGO Collaboration (2016).

© 0000 The Authors

searches become available prior to the CBC ones. This can happen if the astrophysical signal deviates from the waveform template used by the CBC search algorithms. For example, spin-induced amplitude and phase modulation might be present in the signal, which is not captured by the spin-aligned waveforms used by CBC searches (Abbott et al. 2016b; The LIGO Scientific Collaboration et al. 2016), causing a loss of signal-to-noise ratio (SNR) (Abbott et al. 2016b).

Efforts to introduce fully spinning CBC searches are ongoing, but not yet implemented (Harry et al. 2016). Given their unmodeled nature, burst algorithms would not be affected in the same way. In this example, CBC skymaps could be produced with higher latency, after fully precessing spin parameter estimation analyses are completed. Situations in which the burst skymaps are generated earlier could also be produced by technical issues in processing the data or race conditions that lead to latency offsets. Additionally, because of the different strides used by each of the burst and CBC low latency searches, it is plausible that their (temporal) overlaps may not be perfect (Abbott et al. 2016c,b). Although we do not expect these situations to be common, they certainly warrant consideration.

Potential differences in the time at which skymaps may become available can have significant implications for EM observing strategies designed to maximize the chance of capturing an EM counterpart. It is thus important to verify if the degree of difference between CBC- and burst-reconstructed BBH skymaps similar to GW150914 would be typical and, overall, quantify the relation of such skymaps. In this paper we consider a large set of BBH simulations and compare their skymaps produced by the most recent version of the low-latency burst algorithms that ran in the first observing run (O1), `cWB` and `LIB` (Klimenko et al. 2016; Lynch et al. 2015) and by the CBC parameter estimation algorithm `lalinference` (Veitch et al. 2015).

2 SIMULATIONS

The sky localization performance of burst algorithms was thoroughly analyzed in Essick et al. (2015), where several morphologies traditionally used to simulate GW bursts were investigated, including BBH signals generated by heavy black holes. Essick et al. (2015) focuses on two burst algorithms, considering a two-interferometer (Hanford, WA and Livingston, LA, denoted as "HL") network with the two LIGOs at early 2015 sensitivity (roughly achieved during O1), and a three-interferometer (HLV, where "V" denotes Virgo) network with the two LIGOs and Virgo at their early sensitivity (Abbott et al. 2016a). In this paper we extend that study to include skymaps from the CBC parameter estimation algorithm `lalinference` (Veitch et al. 2015). Of the original BBH events considered in Essick et al. (2015) we only kept sources with component masses larger than $20M_{\odot}$, because heavy BBH are more compact in the time-frequency domain and thus more easily detectable by burst algorithms with high confidence (For example, the second BBH detection, GW151226, was not found with high significance by burst algorithms (Abbott et al. 2016e)). This selection leaves 237 simulated events for the HL network and 195 simulated events for the HLV network.

We still consider the HL network because the first few months of the second observing run (O2) will only comprise the two LIGOs, with a sensitivity not too different from O1. We therefore expect the HL runs of Essick et al. (2015) to be representative of the events one could detect in the first part of O2, without Virgo. Similarly, the HLV events could represent the typical situation for the late part of O2, with Virgo online. The simulated events were added to simulated Gaussian noise (identical to Essick et al. (2015)) and their parameters estimated with the nested sampling version of `lalinference` (Veitch et al. 2015). Berry et al. (2015) showed that using Gaussian noise yields the same statistical distribution of skymaps one would obtain with real interferometric noise.

We used a spin-aligned template family (`SEOBNRv2ROM` (Pürrer 2014, 2016)) to decrease the computational time. The simulated signals were also spin-aligned, although from a different waveform family (`IMRphenomB` (Ajith et al. 2011)). Because the typical detection with advanced detectors will have a small inclination angle (Schutz 2011), as indeed was the case for both GW150914 and GW151226 (Abbott et al. 2016d; The LIGO Scientific Collaboration et al. 2016; Abbott et al. 2016e), which reduces the visible effects of eventual spin precession (Apostolatos et al. 1994; Vitale et al. 2014). This spin-aligned study thus represents a reasonable approximation. Furthermore, it has been shown that neglecting mild spin precession does not significantly affect the reconstructed skymaps (Farr et al. 2016; Raymond et al. 2009). We allowed for calibration uncertainties in the data, using the same method described in (Abbott et al. 2016d). For the amplitude calibration error prior, we used a zero-mean Gaussian with standard deviation of 10%, while for the phase we used a standard deviation of 5 degrees. This is comparable with what was achieved during O1 (The LIGO Scientific Collaboration & Abbott 2016; The LIGO Scientific Collaboration et al. 2016).

The burst algorithms skymaps originally released in Essick et al. (2015) were regenerated using the most recent versions of `cWB` and `LIB`. The burst algorithms do not currently marginalize over calibration uncertainties.

3 RESULTS

3.1 Fidelity

We first address the question of how similar two skymaps are. Several figures of merit can be used to quantify similarity. In this section we use the Fidelity, defined as follows. Given two skymaps p and q we calculate the Fidelity as:

$$F(p,q) \equiv \sum_i \sqrt{p_i q_i} \in [0,1] \tag{1}$$

where p_i and q_i are the probabilities assigned to the ith pixel by each skymap. The Fidelity would be exactly 1 for two identical skymaps, whereas it would be 0 for skymaps which do not have any common support. In Fig. 1 we report the cumulative distribution of the Fidelity for pairs of algorithms for the HL (solid lines) and HLV (dashed lines) networks.

Let us consider the HL network first. We see that the

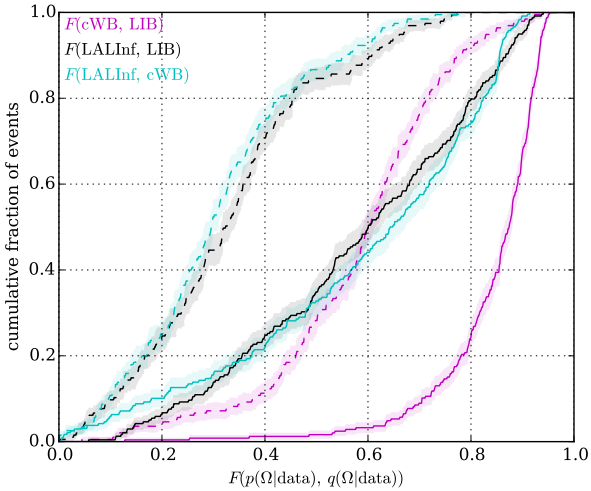

Figure 1. The cumulative distribution of the Fidelity of skymaps for pairs of algorithms in the HL (solid) and HLV (dashed) analyses.

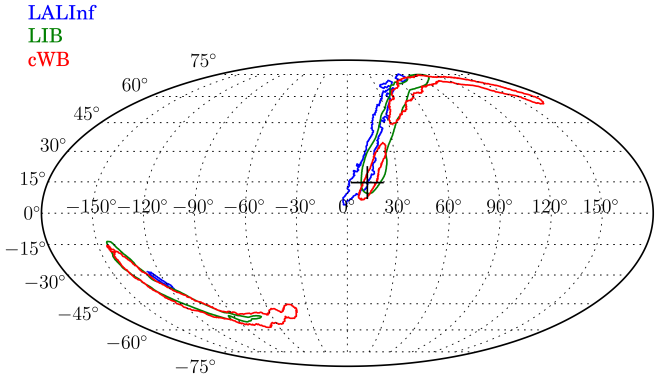

Figure 2. HLV event 968666271. Burst and CBC pick a different ring. The true position is indicated with a black "+". The Fidelity of the burst skymaps with `lalinference` is below 0.35.

two burst algorithms generally agree very well with each other, with 50% of skymaps having a Fidelity larger than 0.85. The `cWB-LIB` Fidelity for GW150914 (0.55 (Abbott et al. 2016c)) was significantly lower than this; which we explain later. Both burst algorithms have a similar degree of Fidelity with `lalinference`. For 50% of the events, the Fidelity of `cWB` skymaps with `lalinference` skymaps is larger than 0.65, while for `LIB` this number is 0.6. These numbers are consistent with what found for GW150914 (0.51 for `cWB-lalinference` and 0.68 for `LIB-lalinference`).

Let us now consider the HLV network, which might realistically be online in early 2017 (Abbott et al. 2016a). These are the dashed lines in Fig. 1. Comparing with the full lines we see a general degradation of the Fidelity. In fact, the median Fidelity across burst algorithms reduces to roughly 0.65, while it is above 0.8 for the HL network. The degradation is even more significant for the Fidelity of each burst algorithm with `lalinference`, with medians at around 0.3. This is because the HLV skymaps are known to break up into several islands of probability, for both CBC (Singer et al. 2014) and burst algorithms (Essick et al. 2015). Because `lalinference` can better match CBC signals in the data, it makes better use of information from Virgo and fragments first. This produces the significant differences between burst and `lalinference` localizations with three detectors. As mentioned earlier, the skymaps from two classes of algorithms might be in tension because they select different parts of the same sky-ring, or because they prefer different rings. An example of a source for which different skymaps pick different rings is given in Fig. 2.

Can the two effects be disentangled? If different rings are selected, one will find that the posteriors on the time-of-flight between two detectors (e.g. Hanford and Livingston) are different. To verify how often this happens, we calculate the Fidelity for pairs of one-dimensional posteriors on the HL time-of-flight (Fig. 3). We see that with the HL network (solid lines), for 50% of the skymaps the Fidelity is above 0.8 for CBC-burst skymaps, and above 0.9 for `LIB-cWB` skymaps. The CBC-burst Fidelity is a bit lower for the HLV network, while the `LIB-cWB` median Fidelity stays close to 0.9. The two burst algorithms thus select the same sky ring very often. Strong correlation exists between the Fidelity of the full skymap and that of the time-of-flight, implying that when slightly different rings are selected (low Δt_{HL} Fidelity) the overall Fidelity is strongly reduced. Thus, the disagreement we found while comparing the full skymaps, Fig. 1, is determined by both the placement of probability around the ring and the position of the ring itself, with the latter happening less frequently but having a large impact when present.

Fig. 4 shows the clear correlation between angular offsets and decreased Fidelity using data from the HL network, as well as the corresponding values from GW150914 (stars), which fall within the scatter observed in simulations. We note that while there is quite a bit of scatter, the general trends seem to be independent of which algorithms are actually compared. This further suggests that low Fidelity is driven by differences in the selected triangulation rings, because all algorithms produce triangulation rings.

We mentioned above that the GW150914 Fidelity for `LIB-cWB` was much lower than average. We can now understand why: the `LIB` and `cWB` skymaps for GW150914 have support on different sky rings (see Abbott et al. (2016f)). They thus belong to this rarer class of reconstructed skymaps for which there is a significant drop of Fidelity.

Furthermore, one can compute the distribution of the angular distance between rings. We find that the distributions look unimodal, with mean smaller than 0.2° for all pairs of algorithms, and standard deviations in the range $[2.5° - 4°]$ ($[2.5° - 7°]$) for the HL (HLV) network.

3.2 Searched Area

So far we have focussed on the degree of similarity of skymaps, which was naturally prompted by the fact that burst and CBC skymaps were so different for GW150914. However that is probably not the most relevant figure of merit to assess the chances of a successful EM follow-up. The fact that a skymap is larger or different might not lead to wasted telescope observation time as long as most of the probability is at the true position.

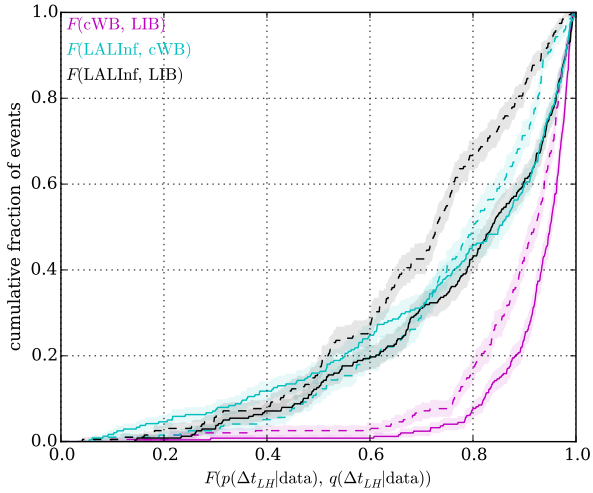

Figure 3. The cumulative distribution of time-of-flight Fidelity for pairs of algorithms in the HL (solid) and HLV (dashed) analysis.

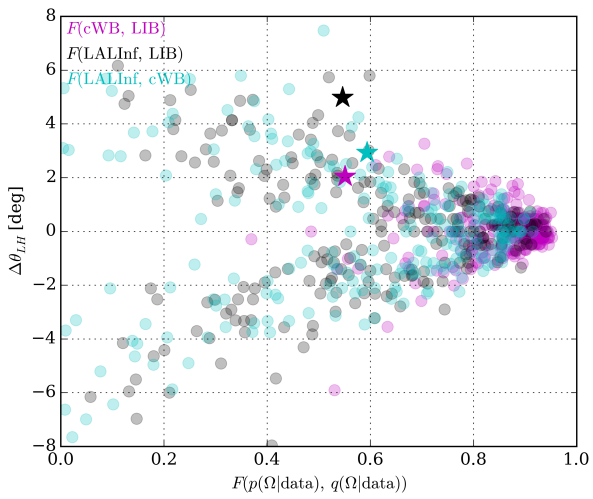

Figure 4. A scatter plot of the fidelity using the full skymap and the angular separation between triangulation rings. GW150914's values are shown as stars.

In this section we report the searched area found by the burst and CBC algorithms. We define the searched area as the amount of sky one needs to image, starting from the most likely skymap pixel and going down, until the true position is found. This is shown in Fig. 5 for HL (solid lines) and HLV (dashed lines). Once again, we first focus on the HL network. As already seen in Essick et al. (2015), the two burst algorithms perform similarly, with 50% of events found with searched areas of roughly 200 deg^2. We also find that `lalinference` can do a factor of 2 better, with 50% of events having searched areas slightly below 100 deg^2. This is perhaps not surprising since `lalinference` can match the signal in the data nearly exactly[2]. Considering the HLV results, we see that while burst skymaps improve only by a factor of 2, the CBC median searched area reduces by an order of magnitude.

[2] Nearly exactly but not exactly since we used a different waveform families to simulate and recover the BBH signals.

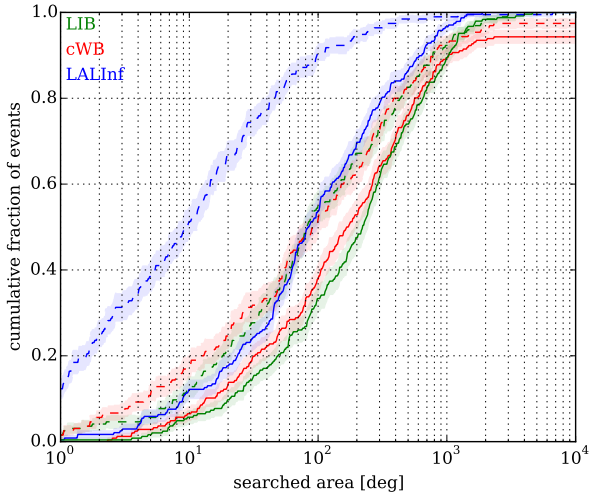

Figure 5. The cumulative distribution of the searched area for each algorithm in the HL (solid) and HLV (dashed) analyses.

We should note that burst searches can be tuned to specific astrophysical systems, including CBC. Such tuning can yield improved search result, including better localization. In this study we only considered skymaps obtained with the most unconstrained versions of the burst algorithms. We have verified that imposing polarization constraints on `cWB` can reduce the searched area by a factor of ∼4 for the HLV events.

4 DISCUSSION AND CONCLUSIONS

Four skymaps for GW150914 were sent to partner EM and neutrino facilities: two produced by unmodeled burst algorithms and two by CBC parameter estimation algorithms. They showed significant differences and large patches of sky where only one of the skymaps had posterior support. In this paper, we addressed the question of whether such differences were just accidental or will rather be frequently encountered.

We considered both a two detector (HL, made of the two LIGO instruments) and a three detector (HLV, made of the two LIGO and Virgo) network. We used sensitivities comparable to what is expected in the next observing run, starting in the fall of 2016. Simulating roughly 200 BBH for each network, we compared the skymaps produced by the two low-latency burst analyses that ran during the first observing run, and the skymap produced by the CBC parameter estimation algorithm `lalinference`. We quantified the degree of similarity between pairs of skymaps using the Fidelity, which is 1 for identical probability distributions and 0 when no overlap exists. We found that the two burst algorithms typically agree with each other, with median Fidelity of ∼0.85 for the HL network and ∼0.6 for the HLV network. The skymaps are less similar for the HLV network.

This is due to the different way each algorithm deals with sub-threshold events in Virgo, which is less sensitive than either LIGO. The agreement between each of the burst algorithms and the CBC result is lower, around 0.6 for HL and 0.3 for HLV. We then verified that the differences can

arise from a combination of two factors: different parts of the same sky rings can be selected by different algorithms, or different sky rings can be preferred in the first place. This latter effect is less common for HL networks, but significantly decreases the Fidelity. We find that the mean angular distance between rings is below 0.2 degrees for all pairs of algorithms, with standard deviations of $\sim 2^\circ$ to 7°. Finally, we focused on the searched area, i.e. the amount of sky that must be imaged before the true position of the source is found.

We found that for the HL network the two burst algorithms lead to similar results, with median searched area of $\sim$200 deg^2. The CBC algorithm does a factor of $\sim$2 better. When the HLV network is considered, we found that using CBC skymaps can lead to a dramatic reduction in searched area. In fact, while the median searched area shrinks by only a factor of $\sim$2 for the burst skymaps, it reduces by an order of magnitude for `lalinference`, for which the median searched area is only 10 deg^2 with an HLV network. We underline that `lalinference` will not produce skymaps with minute-scale latency. However, it has been shown that the low-latency algorithm `BAYESTAR` produces skymaps which agree well with the `lalinference` ones (Singer et al. 2014), at least for the HL network. Nonetheless, we used the `lalinference` skymaps as representative of the definitive skymaps for a BBH source, which was the LIGO and Virgo collaborations' approach in O1 (Abbott et al. 2016f,d). Furthermore, while `BAYESTAR` needs the output of a CBC search algorithm to run (Abbott et al. 2016f), `lalinference` does not, and can thus refine skymaps for events found by either burst or CBC search algorithms.

Our study shows that the level of tension observed between burst and CBC skymaps for GW150914 is not exceptional (although the level of disagreement of `LIB` and `cWB` was larger for GW150914 than on average for BBH). Furthermore, we conclude that not only will low-latency burst and CBC skymaps be different, but the latter will be significantly more accurate. Waiting for skymaps from CBC analysis, if not available in low latency, may delay EM observations. These delays should be considered depending on the EM facility and the possible astrophysical source. Commencing observations promptly is the only way to capture an EM signature on timescales less than a day, typically expected for the high energy emission and UV/O/IR (Metzger & Berger 2012). Alas, this may add uncertainty and hinder success due to the limitations of prompt localization. On the other hand, EM facilities targeting emission on the scale of days to months to years (primarily in the radio (Metzger & Berger 2012)) may opt to wait for subsequent communications from the LIGO-Virgo collaborations regarding localization information.

5 ACKNOWLEDGMENTS

The authors would like to thank C. Berry, T. Dent, V. Kalogera and S. Eikenberry, as well as the MNRAS referee for useful comments and suggestions. The authors acknowledge the support of the National Science Foundation and the LIGO Laboratory. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0757058. S. K thanks the National Science Foundation for support under grant PHY-1505308. The authors would like to thank the Albert Einstein Institute in Hannover, supported by the Max-Planck-Gesellschaft, for use of the Atlas high-performance computing cluster. This is LIGO Document P1600300

REFERENCES

Abbott B. P., et al., 2016a, Living Reviews in Relativity, 19
Abbott B. P., et al., 2016b, Phys. Rev. D, 93, 122003
Abbott B. P., et al., 2016c, Phys. Rev. D, 93, 122004
Abbott B. P., et al., 2016d, Physical Review Letters, 116, 241102
Abbott B. P., et al., 2016e, Physical Review Letters, 116, 241103
Abbott B. P., et al., 2016f, ApJ, 826, L13
Ackermann M., et al., 2016, ApJ, 823, L2
Adrián-Martínez S., et al., 2016, Phys. Rev. D, 93, 122010
Ajith P., et al., 2011, Physical Review Letters, 106, 241101
Apostolatos T. A., Cutler C., Sussman G. J., Thorne K. S., 1994, Phys. Rev. D, 49, 6274
Berry C. P. L., et al., 2015, ApJ, 804, 114
Cannon K., et al., 2012, ApJ, 748, 136
Connaughton V., et al., 2016, ApJ, 826, L6
Díaz M. C., et al., 2016, ApJ, 828, L16
Essick R., Vitale S., Katsavounidis E., Vedovato G., Klimenko S., 2015, ApJ, 800, 81
Evans P. A., et al., 2016, MNRAS, 460, L40
Farr B., et al., 2016, ApJ, 825, 116
Gando A., et al., 2016, ApJ, 829, L34
Harry I., Privitera S., Bohé A., Buonanno A., 2016, Phys. Rev. D, 94, 024012
Hurley K., et al., 2016, The Astrophysical Journal Letters, 829, L12
Klimenko S., et al., 2016, Phys. Rev. D, 93, 042004
LIGO Collaboration 2016, `https://losc.ligo.org/events/GW150914`
Lynch R., Vitale S., Essick R., Katsavounidis E., Robinet F., 2015, preprint, (`arXiv:1511.05955`)
Metzger B. D., Berger E., 2012, ApJ, 746, 48
Palliyaguru N. T., et al., 2016, ApJ, 829, L28
Privitera S., et al., 2014, Phys. Rev. D, 89, 024003
Pürrer M., 2014, Classical and Quantum Gravity, 31, 195010
Pürrer M., 2016, Phys. Rev. D, 93, 064041
Raymond V., van der Sluys M. V., Mandel I., Kalogera V., Röver C., Christensen N., 2009, Classical and Quantum Gravity, 26, 114007
Savchenko V., et al., 2016, ApJ, 820, L36
Schutz B. F., 2011, Classical and Quantum Gravity, 28, 125023
Singer L. P., Price L. R., 2016, Phys. Rev. D, 93, 024013
Singer L. P., et al., 2014, ApJ, 795, 105
Smartt S. J., et al., 2016, MNRAS, 462, 4094
Soares-Santos M., et al., 2016, ApJ, 823, L33
Tavani M., et al., 2016, ApJ, 825, L4
The LIGO Scientific Collaboration Abbott B. P., 2016, preprint, (`arXiv:1602.03845`)
The LIGO Scientific Collaboration et al., 2016, Phys. Rev. X, 6, 041015
Veitch J., et al., 2015, Phys. Rev. D, 91, 042003
Vitale S., Lynch R., Veitch J., Raymond V., Sturani R., 2014, Phys. Rev. Lett., 112, 251101